Timothy P. Quinn, Jolene Splett, Joseph D. McColskey, James Dawson, David J. Smith, Adam Himes, and Daniel Cooke

The Reproducibility of a Proposed Standard Fatigue Test for Cardiac Device Leads

Citation

ABSTRACT
The Transvenous Cardiac Leads Working Group of the Cardiac Rhythm Management Devices Committee of the Association for the Advancement of Medical Instrumentation is developing a fatigue performance standard for cardiac device leads. The proposed standard would calculate a figure-of-merit (FOM) that is based on a life prediction using a Bayesian framework. The framework uses distributions for bending fatigue strength, patient age, patient activity level, and in vivo bending. The benchtop fatigue testing portion of the standard is based on the unsupported bending of the lead at multiple alternating curvature levels to generate fatigue fracture data in low-cycle and high-cycle regimes. To estimate the interlaboratory reproducibility of the benchtop testing methodology, a lead mock-up was constructed from a bifilar MP35N coil in a thin-walled polyurethane tube. Four laboratories each tested 48 specimens and produced fatigue life curves based on the results. To compare the data, the FOM...
that is proposed for the standard was calculated using initial in vivo curvature data of a well-performing lead. The reproducibility standard deviation of the FOM across the four laboratories was 0.17% of the grand mean.

**Keywords**
cardiac device leads, fatigue testing, reproducibility, life prediction, figure-of-merit (FOM)

**Introduction**

The purpose of this study is to estimate the reproducibility of the benchtop testing portion of a proposed new standard that addresses failure from fatigue of cardiac device leads that are used in resynchronization therapy. In 2007, the U.S. Food and Drug Administration recalled a lead (Recall Event ID 45403) for unacceptably high failures in fatigue that affected hundreds of thousands of patients. This prompted industry and the Food and Drug Administration to examine the standards\(^1\) used for testing the leads in fatigue. After review, the appropriate standards body (the Transvenous Cardiac Leads Working Group of the Cardiac Rhythm Management Devices Committee of the Association for the Advancement of Medical Instrumentation, or AAMI CRMD WG1) was tasked with developing a standard for fatigue that would be able to distinguish between lead designs that will give adequate service in use from those designs that might not.

Cardiac leads connect a pulse generator to a location in the heart. Applications include sensing the heart rhythm, providing low-energy pulses for pacing, and providing high-energy pulses for defibrillation. Leads usually consist of a conducting coil and (at times) wound wire cables that are threaded through lumens in an insulating sheath.\(^2\) The coil serves a dual purpose: as a conductor and, during lead insertion, as a wire guide. A stylet is introduced into the coil so that the surgeon can place the lead in the proper position in the heart and venous system. The stylet is removed once the lead is positioned. The leads are then connected to a device (pacemaker/defibrillator) that is implanted subdermally, usually in the shoulder region. The relative position of the heart and device changes with heartbeat, body position, and shoulder movement. Slack must be provided in the lead so that it can remain anchored in the heart. This results in a high-fatigue environment for the lead in bending.\(^2-3\) Some torsional loading about the long axis of the lead may be present, but Baxter and McCulloch\(^4\) found the resultant stress in a simple model of the lead to be two orders of magnitude lower than the bending stress. Liu et al.\(^5\) found the torsional stress to be significant but modeled the end condition as fixed whereas the coil is actually free to move within the lumen.

Since the bending load of the lead is proportional to the curvature, the AAMI CRMD WG1 has based the new standard on applied cyclic curvature. The curvature of the lead in vivo can be measured with biplanar x-rays\(^4-6\) to determine the use conditions of the leads. The proposed standard calls for benchtop testing to
determine the fatigue strength by applying cyclic curvature to the lead as it bends in free space. This fatigue strength measured on the benchtop, together with the use conditions measured from a proposed in vivo study, would be the inputs to a life-prediction scheme that uses a simplified form of the Bayesian algorithm detailed by Haddad, Himes, and Campbell. The proposed standard takes the fifth percentile of 10-year survival of the lead and uses it as a figure-of-merit (FOM). The FOM would be used to assess the acceptability of the lead design.

Here, we conducted a round-robin study of the methods used to determine fatigue strength using a simplified, standard fatigue specimen that consisted of a bifilar coil inside a thin-walled polyurethane tubing. Four operators conducted fatigue strength testing. In order to estimate the reproducibility of the testing, we used a simulated use condition and calculated the FOM for each operator.

Methods

The bifilar coil diameter was 0.91 mm and was made from 0.23-mm-diameter MP35N wire; the outer diameter of the 55D polyurethane tube was 1.27 mm with a 1.02-mm inner diameter. The coil design is consistent with the reference test coil described in standard EN 45502-2-1:2003, Annex CC, Section 23.5. A test specimen had a minimum length of 200 mm. All test specimens had the coil fixed to the insulating tubing at one end of the specimen to prevent migration of the coil in the tubing during transportation and mounting. All coil and insulation materials were taken from the same batch. Each operator conducted fatigue tests on 48 specimens.

The testing method is to apply curvature to a freestanding lead (unsupported except at the grips) by either buckling the lead or by simple bending. Clamping methods were also at the discretion of the operators. All testing was conducted at $23^\circ C \pm 5^\circ C$, and the cyclic frequency was 5 Hz. The load ratio $R = k_{\min}/k_{\max}$ was $0.3 \pm 0.03$, with $k_{\min} =$ minimum curvature and $k_{\max} =$ maximum curvature; curvature amplitude is given by $k_a = 0.5 \cdot (k_{\max} - k_{\min})$. In order to reduce the size of the confidence bounds on the failure curve, the proposed standard calls for testing across the range of possible alternating curvatures. Therefore, the operators were asked to test at various curvature levels but were allowed to submit any valid test data. The operators labeled each submission with the level they were attempting to achieve.

All fatigue tests started with a run-in period of 1,000 cycles to account for possible curvature changes during the first cycles. Curvature change can be caused by plastic deformations of insulation or conductors at high curvatures (or both). Minimum and maximum curvatures of each sample were measured after the run-in period and were used for all subsequent analyses provided that $R$ was within tolerance ($R = 0.3 \pm 0.03$). Operators were permitted to readjust minimum and maximum curvatures within the first 100 cycles of the run-in period. All operators tested the specimens either to failure or five million test cycles. A limit of five million cycles was chosen because it adequately characterizes the fatigue strength model,
FIG. 1 At the operator’s discretion, the curvature was applied with (A) a “U”-shaped bend or (B) an “S”-shaped bend (buckle).

TABLE 1 Desired curvature levels and sample size.

<table>
<thead>
<tr>
<th>Curvature Amplitude Level</th>
<th>Curvature Amplitude $k_a$</th>
<th>Sample Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.78 cm(^{-1})</td>
<td>12</td>
</tr>
<tr>
<td>2</td>
<td>1.11 cm(^{-1})</td>
<td>12</td>
</tr>
<tr>
<td>3</td>
<td>2.12 cm(^{-1})</td>
<td>12</td>
</tr>
<tr>
<td>4</td>
<td>2.45 cm(^{-1})</td>
<td>12</td>
</tr>
</tbody>
</table>

Note: For a valid submission, $0.27 \leq R \leq 3.3$. 
as shown in equation (1), and it is physiologically relevant for a 10-year life in the shoulder region. The electrical integrity of the coil was monitored continuously throughout the test. Failure was defined as an interruption of the electrical continuity or a rise in direct current resistance above 150% of the initial value. Tests that ended because of failures at the grips were treated as suspensions according to the advice given in Nelson.

Minimum and maximum curvatures of cyclic lead bending were measured at the apex position of test samples by imaging the samples optically or with x-ray at the operator’s discretion. Each operator developed software to fit a circle to the outer edge of the lead using the apex and the two points at ±20° to the apex. The centerline curvature was then found by subtracting 0.635 mm, which represents half of the nominal lead body diameter, and was used to calculate the curvature.

In order to compare results among operators more easily, the $k_a$ was normalized so that normalized alternating curvature, $s$, is $k_a / (2.831 \text{ cm}^{-1})$. Normalizing data used in complex statistical analysis has been adopted as a best practice as it minimizes numerical errors. (In preliminary testing, 2.831 cm$^{-1}$ was the maximum curvature amplitude practically achievable.)

Using maximum-likelihood techniques, the fatigue data were fit to the model,

$$\ln(\text{cycles}) = \beta_0 + \beta_1 s^{-\beta_2} + e,$$

(1)

where $\beta_n$ are coefficients, and $e$ represents random errors that are normally distributed with a mean of zero and standard deviation, $\sigma$ (see Meeker and Escobar). The estimate for $\sigma$ represents variability about the fatigue model for any particular value of the normalized curvature amplitude. Bayesian methods are used to provide posterior distributions for the four fit parameters, $\beta_n$, and $\sigma$. Leads that do not fracture (suspensions) are accounted for in the maximum-likelihood technique according to Meeker and Escobar. Tolerance limits for the model predictions of $\ln(\text{cycles})$ are also computed from the posterior distributions. (See Gelman for information regarding Bayesian data analysis.)

Calculation of the FOM is explained in detail in Haddad, Himes, and Campbell but will be described here in brief. The output of the Bayesian algorithm is the predicted survival of a particular lead segment. (The FOM, again, is the fifth percentile of 10-year survival based on the set of simulated lead survival curves.) The inputs into the algorithm are the fatigue data and the distribution parameters for the population age, activity level, and use conditions (fig. 2). The use-condition data are used to create posterior distributions of the mean and variance (see Gelman). The fatigue data, from benchtop testing of a particular section of the lead, are used to create a posterior distribution of fatigue models. Using these posterior distributions, a use condition and fatigue model are randomly selected for $M$ simulated patient cohorts (the outer loop in fig. 2). The survival curve for each cohort is calculated out to 10 years. To do that, the algorithm simulates a patient cohort by randomly selecting $N$ patients according to the age distribution (inner loop in fig. 2).
Then, an activity level is assigned randomly to each patient according to the activity-level distribution. The life of the particular lead segment can now be predicted for each simulated patient. The survival as a function of time is then predicted using the data from the entire cohort. Thus, $M$ survival curves are generated, and the FOM is calculated from the set of simulated survival curves. (Here, $M = 500$ and $N = 1,000$.)

For this round-robin, the FOM is calculated utilizing approximate use conditions that might be found in the shoulder region. The shoulder region has a relatively high applied cyclic curvature, and a lower number of fatigue cycles, than the intracardiac region. For the comparisons used in this research, Haddad, Himes, and Campbell\textsuperscript{7} was simplified as follows: (1) Any flexion of the shoulder beyond a random angle between $100^\circ$ and $160^\circ$ is considered a maximum cycle; there is no distribution in arm angle (the number of bins equals one). (2) Consequently, there is just one cyclic-curvature level constant throughout the implant duration, and the cumulative damage model in Haddad, Himes, and Campbell\textsuperscript{7} is not necessary. (3) The return to activity time is set to zero. (4) The fatigue model shown in equation (1) is used. (5) The load ratio is fixed ($R = 0.3$). (6) No competing risks were considered—all simulated patients survive longer than their implant. (7) There are no orientation effects due to the position of the lead. The arm activity distribution is taken from Coley et al.\textsuperscript{11} and Schurr and Ada,\textsuperscript{12} while the age distribution is from Arnsbo and Moller.\textsuperscript{13}

A nonparametric, bootstrap resampling technique from Efron and Tibshirani\textsuperscript{14} was used to generate distributions of the FOM for each operator. The FOM was computed for each of 2,000 bootstrapped samples that were randomly selected.
from the four effective alternating curvature levels. While operators were asked to apply alternating curvature at the four effective alternating curvature levels, achieving the target exactly was not possible. Thus, the effective alternating curvature levels can be thought of as bins that contain similar alternating curvature measurements. The actual alternating curvature measurements for each bin (level) were used in the analysis.

Specifically, samples of size 10, 10, 5, and 5 were randomly selected (with replacement) from the fatigue data for the four alternating curvature levels, respectively, and used to calculate the FOM. (The proposed standard calls for a minimum of 30 fatigue samples to be tested across the range of use conditions.) One FOM was produced for each sampled fatigue data set. The resultant survival curves (see Haddad, Himes, and Campbell7) used to obtain the FOM were based on the simulated use condition of a lognormal distribution, $Y = \ln(X)$, where units of $X$ are in $\text{cm}^{-1}$, of alternating curvature with location $= 0.21$ and scale $= 0.53$. This use condition is a reasonable approximation given proprietary preliminary data for a well-performing lead. The use condition should only be utilized to compare the reproducibility of the FOM among the four operators.

Results

The actual alternating curvature levels achieved by the operators were lower than those asked for in table 1. In preliminary testing, Operator 2 was the only operator able to obtain data at the 0.78 cm$^{-1}$ alternating curvature level; however, all the values were suspended (did not fracture). The alternating curvature levels returned by the operators are given in table 2.

Some leads did not fracture in the test section near the apex but at or near a clamp. These data are included in the analyses and are considered suspended because the lead itself did not fracture in the test section. In general, the four operators’ data showed agreement with the usual scatter associated with fatigue (fig. 3). (Figure 3 raw data are available from NIST upon request.)

The model parameters estimated for each operator, defined as the medians of the posterior distributions, are shown in table 3. Medians were used instead of means because of skewness in the posterior distributions.

<table>
<thead>
<tr>
<th>Nominal Level (1/cm)</th>
<th>OP1</th>
<th>OP2</th>
<th>OP3</th>
<th>OP4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (1/cm)</td>
<td>S.D. (1/cm)</td>
<td>Mean (1/cm)</td>
<td>S.D. (1/cm)</td>
</tr>
<tr>
<td>1</td>
<td>1.098</td>
<td>0.077</td>
<td>1.125</td>
<td>0.061</td>
</tr>
<tr>
<td>2</td>
<td>1.400</td>
<td>0.059</td>
<td>1.277</td>
<td>0.068</td>
</tr>
<tr>
<td>3</td>
<td>2.244</td>
<td>0.168</td>
<td>2.160</td>
<td>0.061</td>
</tr>
<tr>
<td>4</td>
<td>2.431</td>
<td>0.156</td>
<td>2.305</td>
<td>0.143</td>
</tr>
</tbody>
</table>
Figure 3 displays S-N curves (labeled “predicted”) for each operator based on the model parameters shown in Table 3. Also shown in Figure 3 are tolerance limits with 90% coverage and 95% confidence for the value of \( \ln(\text{cycles}) \) predicted by the models. Figure 5 is a comparison of the fitted models across all operators.

Table 4 lists summary statistics for the distribution of the FOM for each operator based on the bootstrapped samples. Figure 6 shows box plots of the bootstrapped values of the FOM for each operator.

A bootstrap technique was used to obtain estimates of the repeatability and reproducibility standard deviations. The standard deviations shown in Table 4 are very similar, so we pooled the values to obtain an estimate of the repeatability standard deviation, \( s_r = 0.0016 \). The standard deviation of the mean FOM values, \( s_L = 0.0005 \), serves as an estimate of the between-laboratory standard deviation. Thus, the reproducibility standard deviation is \( \sqrt{s_r^2 + s_L^2} = 0.0017 \). (We use the definition of reproducibility standard deviation from equation X1.2 in ASTM E691-18, Standard Practice for Conducting an Interlaboratory Study to Determine the Precision of a Test Method.\(^{15} \))
FIG. 4  Fatigue data and curves for each operator. The dashed lines represent the predicted fatigue curves based on the parameters given in table 4 for each operator. The plots also show curves corresponding to tolerance intervals containing 90% of predicted ln(cycles) with 95% confidence. Open symbols denote suspensions.

FIG. 5  The fatigue curves for each operator based on the model parameters given in table 3.
Discussion

Using the results of the FOM bootstrap analysis for the four operators, the reproducibility standard deviation is just 0.17 % of the grand mean. The proposed standard does not restrict how the lead may be clamped; it only requires that a specific curvature be applied. The four operators used a wide variety of clamps (fig. 7) including a grooved flat (fig. 7A), an inclined groove (fig. 7B), and an “S” buckle configuration (fig. 7C). The low variation of the FOM across operators indicates that the fatigue measurement is not sensitive to clamp design.

The ability of an operator to test at a specific alternating curvature depends on the testing apparatus that is being used. The typical apparatus is testing multiple specimens at once (i.e., one machine stroke setting is generating bending in multiple specimens), each in its own set of grips. In addition, the test specimens are assemblies of multifilar coil inside tubing; thus, a set of such specimens will have

<table>
<thead>
<tr>
<th>Operator</th>
<th>Minimum</th>
<th>First Quartile</th>
<th>Median</th>
<th>Mean</th>
<th>Third Quartile</th>
<th>Maximum</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>OP1</td>
<td>0.9747</td>
<td>0.9795</td>
<td>0.9805</td>
<td>0.9804</td>
<td>0.9816</td>
<td>0.9850</td>
<td>0.0016</td>
</tr>
<tr>
<td>OP2</td>
<td>0.9738</td>
<td>0.9794</td>
<td>0.9806</td>
<td>0.9804</td>
<td>0.9816</td>
<td>0.9858</td>
<td>0.0016</td>
</tr>
<tr>
<td>OP3</td>
<td>0.9754</td>
<td>0.9798</td>
<td>0.9809</td>
<td>0.9808</td>
<td>0.9818</td>
<td>0.9855</td>
<td>0.0015</td>
</tr>
<tr>
<td>OP4</td>
<td>0.9760</td>
<td>0.9805</td>
<td>0.9816</td>
<td>0.9816</td>
<td>0.9827</td>
<td>0.9860</td>
<td>0.0016</td>
</tr>
</tbody>
</table>

FIG. 6 Box plots of the bootstrapped values of the FOM for each operator. Points are jittered to improve visibility.
some dimensional variation. Given the variation in specimen dimensions, and the limitation of one test machine motion that bends multiple specimens at one time, the system is too constrained to achieve precisely the same curvature in each specimen. The curvature in each specimen is measured independently, however. The 0.17 % reproducibility standard deviation in the FOM suggests that the analysis approach is robust to differences in cyclic curvature among test specimens. In fact, the proposed standard calls for testing specimens across the range of cyclic curvature to ensure that the failure curve is well defined. It automatically penalizes a

**FIG. 7** A wide variety of clamping methods were used by the operators: (A) a flat grooved plate, (B) an inclined grooved plate, and (C) an “S” buckle grip. The images on the left are at minimum stroke, and the images on the right are at maximum stroke.
sparsity of data because this will increase the confidence bound on the failure curve, which in turn decreases the FOM and therefore makes it more likely that a lead design will be rejected.

It is important to note that the operators chose not to test at very low levels of curvature (around 0.78 cm⁻¹). There are two reasons for this. First, these low levels of curvature are difficult to achieve with \( R = 0.3 \) because the stroke for the non-“S” buckled setups would have to be quite large (perhaps greater than 25 mm). Secondly, in preliminary testing, none of the samples failed before five million cycles at these alternating curvature levels.

**Conclusion**

The Bayesian framework provides a method to predict the FOM (the fifth percentile of 10-year survival of the lead) based on benchtop fatigue data, cardiac lead use conditions, and patient age. The four operators participating in a round-robin study were able to produce similar FOM values even though fatigue test fixture designs were quite different. Using bootstrap resampling of the fatigue data, in conjunction with the Bayesian framework, we found that the FOM described here is very reproducible (reproducibility standard deviation = 0.17 %).

**References**

1. *Active Implantable Medical Devices—Part 1: General Requirements for Safety, Marking and Information to be Provided by the Manufacturer*, EN 45502-1, (Nordhavn: CENELEC, Danish Standards Foundation, approved April 20, 2015).


