Month-to-Month and Year-to-Year Reproducibility of High Frequency QRS 
ECG signals

Niles J. Batdorf¹, B.A., Alan H. Feiveson², Ph.D., and Todd T. Schlegel², M.D.
¹National Space Biomedical Research Institute, Houston, TX 77030; ²Human 
Adaptation and Countermeasures Office, National Aeronautics and Space 
Administration, Lyndon B. Johnson Space Center, Houston, TX, 77058

Running head: Reproducibility of High Frequency QRS signals

Address Correspondence to:
Todd T. Schlegel, M.D.
NASA Johnson Space Center
Mail Code SK3
Houston, Texas 77058
Phone: 281-483-9643
FAX: 281-244-5734
E-mail: todd.t.schlegel@nasa.gov

Acknowledgement of Financial Support: Financial support came from a Presidential 
Early Career Award, and NASA grant number 199161157
ABSTRACT:

High frequency (HF) electrocardiography analyzing the entire QRS complex in the frequency range of 150 to 250 Hz may prove useful in the detection of coronary artery disease, yet the long-term stability of these waveforms has not been fully characterized. We therefore prospectively investigated the reproducibility of the root mean squared (RMS) voltage, kurtosis, and the presence versus absence of reduced amplitude zones (RAZs) in signal averaged 12-lead HF QRS recordings acquired in the supine position one month apart in 16 subjects and one year apart in 27 subjects. Reproducibility of RMS voltage and kurtosis was excellent over these time intervals in the limb leads, and acceptable in the precordial leads using both the V-lead and CR-lead derivations. The relative error of RMS voltage was 12% month-to-month and 16% year-to-year in the serial recordings when averaged over all 12 leads. RAZs were also reproducible at a rate of up to 87% and 81%, respectively, for the month-to-month and year-to-year recordings. We conclude that 12-lead HF QRS electrocardiograms are sufficiently reproducible for clinical use.

KEY WORDS: coronary artery disease, ischemia, reduced amplitude zone, root mean square voltage, kurtosis, CR leads
INTRODUCTION

High frequency (HF) analysis of the entire QRS complex of the electrocardiogram (ECG) between 150-250 Hz shows promise as a clinical tool for enhanced diagnosis of myocardial ischemia and coronary artery disease\textsuperscript{1-10}. However, few studies have investigated the reproducibility of this technique. Goldberger et al.\textsuperscript{11} examined the temporal stability of HF QRS root mean squared (RMS) voltage between the frequencies of 80-300 Hz in lead aVF over a 100-day period and reported a high degree of reproducibility between the two recordings with respect to both amplitude and morphology. Aversano et al.\textsuperscript{4} analyzed the HF QRS RMS amplitude between 150-250 Hz using the Frank lead configuration and found that the temporal variability in an average lead in multiple recordings over a three-hour period was just 0.3 microvolts and acceptably low for clinical purposes. Pettersson et al.\textsuperscript{12} more recently observed even higher reproducibility of RMS voltages in the standard 12 HF QRS leads through comparison of two consecutive recordings that were each 2.5 minutes in length. However, neither the Aversano et al. nor the Pettersson et al. studies involved the reapplication of electrodes, which in a clinical setting can detrimentally affect reproducibility. In the present investigation, we therefore tested both the intermediate and long-term reproducibility of 12-lead HF QRS recordings by analyzing the stability of their amplitude and morphology over intervals of approximately one month and one year, respectively.
MATERIALS AND METHODS

The present study was approved by the Johnson Space Center Committee for the Protection of Human Subjects. All subjects who volunteered for the study gave informed consent and were screened at Johnson Space Center's Human Test Subject Facility via a physical examination that included blood pressure measurements and an evaluation of serum lipids before ECG testing. Subjects were asked to avoid rigorous exercise, caffeine, nicotine and alcohol for at least 12 hours before testing. Two 12-lead HF QRS ECG recordings were acquired approximately one month (29.19 ± 3.10 days, mean ± SD) apart for the first part of the study, and approximately one year (340.59 ± 41.24 days) apart for the second part of the study.

For the month-to-month reproducibility study, 16 asymptomatic individuals (9 males and 7 females, age 34.6 ± 15.3 years, mean ± SD, range 21-60 years) were recruited, fifteen having no previous history of cardiovascular disease and one having a prior inferior myocardial infarction. At the time of physical examination, 3 of the 16 subjects (19%) were considered hypertensive (SBP > 140 mm Hg, DPB > 90 mm Hg or taking antihypertensive medications) while 6 of 16 (38%) were considered hyperlipidemic (total cholesterol > 240 mg/dl or total cholesterol to HDL ratio > 5.0 or LDL cholesterol >160 mg/dl or on lipid-lowering medications). None of the subjects was diabetic.

For the year-to-year reproducibility study, 27 asymptomatic individuals (17 males and 10 females, age 40.4 ± 8.9 years, range 25-60 years) with no prior history of cardiovascular disease or diabetes were recruited. At the time of physical examination, 4
of the 27 subjects (15%) were considered hypertensive while 10 of the 27 (37%) were considered hyperlipidemic. Six individuals participated in both studies.

A Cardiax PC ECG system (IMED, Budapest, Hungary) with a frequency response range of 0.5 to 300 Hz and a sampling rate of 1000 samples/s was used to acquire the 12-lead surface ECG. All ECG data were obtained at approximately the same time of day (i.e., within 2.5 hours) during each of the serial recordings. The HF QRS signal was derived, filtered, and analyzed within software developed by the authors at NASA Johnson Space Center. The Mason-Likar configuration was used for the limb leads to reduce skeletal muscle noise. The QRS interval initial signal averaged templates for each channel were created by aligning the first ten 97% cross-correlated QRS waveforms using their respective R-wave fiducial synchronization points. Subsequent incoming beats were then added to the signal averages only when the cross correlation between the incoming beats and the existing templates in each channel was ≥ 97%. If in any channel the cross correlation between the template beat and the new incoming beat was less than 97%, then the beat was rejected for all channels. At each new accepted beat, the growing template was bandpass filtered in software to isolate those frequencies between 150 and 250 Hz, thus creating an HF QRS signal at each beat. The QRS interval was delineated from the conventional ECG signal, and the R-wave onset and S wave offset were then applied to the HF QRS signal. The final signal averaged HF QRS waveforms consisted of 200 accepted beats collected in the supine position.

Measures of HF QRS. The reproducibility of HF QRS amplitude from both month-to-month and year-to-year was assessed by serial measurements of the RMS voltage in each lead. The RMS voltage was determined by (1) squaring the voltage
amplitude of each sample point within the QRS interval; (2) calculating the mean of these squares; and (3) taking the square root of the calculated mean⁵.

To provide a measure of the reproducibility of HF QRS morphology, the presence versus the absence of three distinct types of reduced amplitude zones (RAZs) was also noted (Figure 1). As originally defined by Abboud³, a RAZ occurs in an HF QRS signal when at least two local maxima of the signal’s upper envelope or two local minima of its lower envelope are present (Figure 1B). A local maximum or minimum (darkened circles in the HF QRS signals in Fig. 1 A-D) is in turn defined as an HF envelope sample point (peak or trough) within the QRS interval wherein the absolute value of its voltage exceeds that of the three envelope sample peaks (or troughs) immediately preceding and following it. The RAZ (arrows in Fig. 1B-D) is thus the region lying between the two neighboring maxima or minima. The NASA software automatically searches for local maxima and minima of the HF QRS envelope not only according to the original criteria of Abboud, but also according to stricter criteria that have been developed in an effort to improve the usefulness of RAZ detection for clinical diagnoses. These empirical criteria identify, in addition to the “Abboud RAZ” described above (“A RAZ”), Figure 1B), stricter RAZs that we have termed the “Abboud Percent RAZ” (“AP RAZ”) and the “NASA RAZ” (“N RAZ”), respectively. An Abboud Percent RAZ (Figure 1C) is simply an Abboud RAZ that meets one additional criterion: namely, its secondary local maximum (or minimum) has a voltage that is at least X% of the voltage of the primary local maximum (or minimum) located on the same side of the HF QRS complex. (For the present study, X% was set at 33%). In turn, the NASA RAZ (Figure 1D), which is the strictest RAZ, has both a secondary local maximum and a secondary local minimum,
both having a voltage of at least X% of their respective primary local maximum and minimum. The RAZ nomenclature is such that when an N RAZ is present in any given HF QRS complex, an AP RAZ and an A RAZ must also be present in the same complex by definition. Similarly whenever an AP RAZ is present, an A RAZ must also be present by definition.

Because the presence of a RAZ often alters the "peakedness" of an HF QRS signal, we also calculated the kurtosis\(^{15}\) of the individual signals. Kurtosis is a measure used in statistical science to provide a quantification of the peakedness of a statistical distribution. To determine the kurtosis of an HF QRS signal, we first plot the absolute value in microvolts of the signal's envelope sample points against time in milliseconds. The resulting complex, when normalized to have an area of 1.0 can be thought of as a probability density function with central moments \(\mu_k = \int (t-\mu)^k f(t) dt\) \((k = 1, 2, ...)\), where \(f(t)\) is the normalized absolute voltage at sample point \(t\) and \(\mu = \int t f(t) dt\). The kurtosis (\(\gamma\)) is the normalized fourth moment \(\gamma = \frac{\mu_4}{\mu_2^2}\). Because a bimodal distribution has a relatively low kurtosis, and because the presence of a RAZ may cause the shape of an HF QRS envelope to resemble a bimodal distribution, HF QRS signals with RAZs typically have a low kurtosis.

Finally, in addition to determining the reproducibility of RMS voltages, RAZs and kurtosis using the standard precordial leads (V1-V6), we also determined the reproducibility of these same measures when using the right arm electrode as the reference point for the precordial leads (CR1-CR6 leads) rather than Wilson's central terminal\(^{16,17}\). This additional analysis was performed because the use of CR leads
maximizes QRS voltages\textsuperscript{16,17} and therefore might theoretically enhance the clinical utility of HF QRS electrocardiography.

Statistics. Statistical analysis was performed using Stata software (College Station, TX). The intra-subject reproducibilities of RMS voltage and kurtosis were evaluated in each lead by calculating the "relative error" (RE), the ratio of the absolute difference between two repeated measurements to the common mean of the measurements, expressed as a percent. Mathematically, RE is equal to 200*|A-B|/(A+B), where A and B are the two measurements\textsuperscript{18}. The reproducibility of the three types of RAZs was assessed by first noting the presence versus the absence of a RAZ in each lead in each recording, with a given RAZ being considered present in a given lead only when it occurred during two-thirds or more of the HF QRS recording (i.e., to avoid over-reliance upon any single HF QRS snapshot beat). A concordant result was then noted if a RAZ was present or absent in a given lead in both recordings, and a discordant result if the result changed between recordings.

Comparisons of reproducibility were made for HF QRS signals in the V leads versus those in the CR leads through a repeated-measures analysis of variance (ANOVA) with RMS voltage and kurtosis REs as dependent variables. However, a square-root transformation was applied to the actual REs to make their distribution closer to normal, which is assumed by the ANOVA. Asymmetric confidence limits for mean REs by lead (Figure 2) were obtained by fitting a general linear model to actual REs assuming a gamma distribution and reciprocal link function\textsuperscript{19}. Comparisons between lead types (limb or precordial) and the effect of time lapse (one month or one year between measurements) were made using the method of generalized estimating equations (GEE)
to account for repeated observations pertaining to each subject\textsuperscript{20} with either a normal or gamma distribution model. A standard analysis of variance model could not be used because there were some (but not all) subjects participating in both the month-to-month and the year-to-year tests. Because any quantitative diagnostic tool must be sensitive to actual change in addition to being reproducible, we also compared within-subject REs to between-subject REs. The between-subject RE is calculated exactly the same as the within-subject RE, except that the measurements being compared come from two different subjects. Thus, the between-subject RE is a relative measure of actual difference between subjects. Measurements with good diagnostic capability should have relatively low within-subject REs ("noise") compared with between-subject REs ("signal"). We used the Wilcoxon unpaired rank test statistic $Z$ to quantify this comparison. The rationale for doing so is that if all REs (within and between) are arranged in increasing order, a preponderance of the within REs should occur near the beginning of the sequence; i.e. they should have lower than random rank orders. Positive values of $Z$ indicate a higher than random incidence of within-subject REs with low rank orders, thus high $Z$ scores are indicative of an efficient diagnostic measurement.

A bivariate probit model\textsuperscript{21} was used to compare the rates of discordance of the A, AP, and N RAZs in the V leads versus the CR leads after month and year-long intervals. This model is especially suited for the comparison of dependent binary response variables. In this case, the dependence arises because the rates of concordance and discordance observed in each of the V leads and the CR leads were obtained from the same ECG recording. In addition, standard errors of model parameter estimates were adjusted to allow for additional possible dependence between different leads for the same
subject. For comparison of the reproducibility of the three different RAZ types, we used GEE on RAZ concordances assuming a binomial distribution with logit link. Statistical significance for the ANOVA and bivariate probit model tests was defined as $P < 0.05$.

RESULTS

Reproducibility of RMS and kurtosis. Figure 2 shows the lead-specific REs (as derived from the limb and V precordial leads) for RMS voltage and kurtosis from month-to-month and year-to-year recordings, respectively. In general, both RMS and kurtosis were more reproducible (lower mean RE) in the limb leads than in the precordial leads ($P = 0.001$ and $P = 0.007$ for RMS and kurtosis, respectively). As can be seen from Figure 2, some differences were found between individual leads within limb or within precordial leads, but no consistent pattern was evident.

Table 1 shows sample means, standard deviations and sample sizes of REs over all observations (subjects and leads) for RMS voltage and kurtosis for each lead derivation (limb leads, V precordial leads and CR precordial leads) and observation interval (month-to-month or year-to-year). However, because REs are by definition non-negative and can occasionally take on relatively large values, the actual skewed distribution of REs is better portrayed by histograms, as shown in Figure 3. For both RMS and kurtosis, regardless of the interval between recordings, repeated-measures ANOVA on square-root transformed REs showed no statistical difference in average reproducibility when using the V versus the CR precordial leads.

Reproducibility of RAZs. Table 2 shows the month-to-month and year-to-year reproducibility of the three types of RAZs. Data are expressed as the number of leads,
out of the maximum possible 12 leads, having a concordant result after comparison of the serial recordings. The N RAZ was without exception the most reproducible RAZ, with for example a concordance of 10.44 out of 12 leads (87.0%) from month-to-month when using the standard leads. The reproducibility of the AP RAZ was slightly less, but still better than that of the A RAZ. Using GEE to compare overall rates of concordance in the standard leads, we found that the N RAZ (overall concordance = 9.83 out of 12 leads) and AP RAZ (overall concordance = 9.23 out of 12 leads) had significantly greater concordance than the A RAZ (8.54 out of 12 leads), ($P = 0.0048$, $P = 0.0054$, respectively), although the concordances of the N and AP RAZs were not significantly different from one another ($P = 0.128$). Using the bivariate probit model, no significant difference was found between rates of concordant versus discordant results for the V leads versus the CR leads for any of the three types of RAZs in either month-to-month or year-to-year recordings. In addition, no clear pattern emerged with respect to the any given individual lead(s) having a clearly superior RAZ reproducibility.

**Effect of observation interval.** Readings taken 1 month apart appeared more reproducible (smaller RES) than those taken one year apart; however these differences were at most borderline statistically significant depending on the analysis model used. For example, using GEE with a normal distribution on square-root transformed RESs, P-values for the effect of interval (one month or one year) were 0.061 and 0.076 for RMS and kurtosis, respectively. Conversely, using the original RESs with a gamma distribution gave P-values of 0.113 and 0.052 for RMS and kurtosis respectively. As can be seen from Table I, no substantive interaction was found between lead type (i.e., limb or precordial) and observation interval for either RMS or kurtosis reproducibility.
Additionally, when using GEE with a binomial distribution and logit link, no statistically significant difference was found between month-to-month and year-to-year concordance rates for either the A or AP RAZ (Table 2). However, the N RAZ was statistically more reproducible in month-to-month than in year-to-year recordings ($P = 0.0051$).

**DISCUSSION**

The results of the present study suggest that RMS voltage and kurtosis of 150-250 Hz HF QRS waveforms have excellent stability in the limb leads over both month and year-long intervals. The stability of these same measures in the precordial leads is less, but probably still acceptable for clinical purposes. (For example, the precordial lead RMS reproducibility in the present study compares favorably to that of RMS voltage vector magnitudes (40-250 Hz) used clinically for analysis of late potentials)$^{18,22}$. Moreover, identification of RAZs, which shows promise as a diagnostic tool for the detection of coronary artery disease in 150-250 Hz HF QRS waveforms$^{2,3,6}$, revealed an acceptable degree of reproducibility over both month and year-long intervals, especially for the NASA RAZ.

Whereas previous studies have investigated the normal variation in RMS voltage during short periods of monitoring that did not involve a change in electrodes$^{4,12}$, the present study examined the reproducibility of several HF QRS parameters in all 12 leads after both month-long and year-long intervals that necessitated a change of electrodes. If HF QRS is to be used clinically for the diagnosis of coronary artery disease and related conditions, it is necessary for the clinician to know the extent to which HF QRS-related
parameters normally fluctuate with successive visits during which variability in electrode placement may occur.

In the present study, kurtosis and RMS voltages were equally reproducible in the CR leads and V leads, as were the three types of RAZs. In a recent evaluation of resting 12-lead HF QRS ECG in a clinical setting, our preliminary data indicate that along with the limb leads, the use of the CR precordial leads (which maximizes QRS voltages)\textsuperscript{16,17} rather than the V precordial leads may enable the best distinction between individuals with and without coronary artery disease\textsuperscript{13}. The statistically equivalent reproducibility of HF QRS complexes derived from the V and CR precordial leads therefore lends support to the acceptability of the use of the CR leads in the clinical setting.

Not surprisingly, the reproducibility of the month-to-month HF QRS ECG recordings was greater in our study than that of the year-to-year recordings. This might be explained purely on a physiologic basis given that any changes within the coronary arteries and cardiac conduction system would be more likely to occur with increasing time intervals. In addition, the superior reproducibility of the N RAZ as compared to the other types of RAZs might be explained by the fact that N RAZs are less susceptible to being formed on the basis of changes in the noise levels at the margins of the HF QRS signals. Moreover, kurtosis, a relatively new measure of HF QRS morphology, appeared to be amongst the most reproducible of the parameters we investigated (Table 1).

The results of the present study will hopefully provide a good reference for expected normal variations in RMS voltage, kurtosis, and RAZs in 150-250 Hz HF QRS ECG recordings over month- and year-long intervals.
REFERENCES


17. De Gasperi RN, Ezzudin SH, Bauerlein EJ, Sequeira R, Lemberg L, Duncan RC: Digitized electrocardiograms recorded with bipolar right-infraclavicular leads compared to electrocardiograms recorded with unipolar chest V leads and bipolar lead II. J Electrocardiol 35:125-34, 2002


Figure Legends:

Figure 1. Screen snapshots showing HF QRS complexes below their respective signal-averaged conventional QRS complexes. Darkened circles connote local maxima and minima, whereas arrows, when present, connote a reduced amplitude zone (RAZ). A. An HF QRS complex from lead I that contains no RAZs. This complex is “normal”. B. An HF QRS complex from lead aVF that contains an “Abboud RAZ” (RAZ A) on both the right upper and lower portion of the signal. The amplitudes of both the secondary local maximum and minimum are nonetheless insufficient to define an “Abboud Percent” RAZ (RAZ AP), because they are less than 33% (empirically defined) of the amplitudes of their corresponding primary local maximum and minimum. C. An HF QRS complex from lead aVR that has a RAZ AP. The RAZ AP is present because the absolute amplitude of the secondary local maximum (left side of signal) that forms the RAZ is ≥33% of that of the corresponding primary local minimum. D. An HF QRS signal that has a “NASA RAZ” (RAZ N). A RAZ N is present because both a secondary local maximum and a secondary local minimum are present and both have amplitudes exceeding 33% of the amplitudes of their respective primary local maximum/minimum.

Figure 2. Asymmetric 95% confidence limits for mean month-to-month and year-to-year RMS and kurtosis REs by lead, obtained by fitting a general linear model to actual REs assuming a gamma distribution and reciprocal link function. n=16 (month-to-month), n=27 (year-to-year)
Figure 3. Histograms illustrating the skewed distribution of the REs for RMS and kurtosis. n=192 (month-to-month), n=324 (year-to-year)
Table 1

Reproducibility of HF QRS RMS Voltage and Kurtosis

<table>
<thead>
<tr>
<th>Interval</th>
<th>n</th>
<th>Lead Type</th>
<th>RMS (mean RE ± SD)</th>
<th>Kurtosis (mean RE ± SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Month</td>
<td>192</td>
<td>limb</td>
<td>8.18 ± 7.71%</td>
<td>8.51 ± 6.84%</td>
</tr>
<tr>
<td>Month</td>
<td>192</td>
<td>V precordial</td>
<td>15.91 ± 14.50%</td>
<td>11.57 ± 10.30%</td>
</tr>
<tr>
<td>Month</td>
<td>192</td>
<td>CR precordial</td>
<td>14.41 ± 11.81%</td>
<td>11.65 ± 10.57%</td>
</tr>
<tr>
<td>Year</td>
<td>324</td>
<td>limb</td>
<td>12.77 ± 11.44%</td>
<td>11.53 ± 9.58%</td>
</tr>
<tr>
<td>Year</td>
<td>324</td>
<td>V precordial</td>
<td>19.14 ± 15.96%</td>
<td>14.62 ± 11.09%</td>
</tr>
<tr>
<td>Year</td>
<td>324</td>
<td>CR precordial</td>
<td>18.47 ± 16.38%</td>
<td>13.01 ± 9.44%</td>
</tr>
</tbody>
</table>

n = number of observations; RMS = root mean squared voltage; RE = relative error; SD = standard deviation
Table 2

Average Concordance of RAZs in all Leads out of a Maximum of 12
Mean ± SD

<table>
<thead>
<tr>
<th>Period</th>
<th>n</th>
<th>Lead Derivation</th>
<th>A RAZ</th>
<th>AP RAZ</th>
<th>N RAZ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Month</td>
<td>16</td>
<td>V + limb leads</td>
<td>8.69 ± 1.89</td>
<td>9.44 ± 2.25</td>
<td>10.44 ± 1.50</td>
</tr>
<tr>
<td>Month</td>
<td>16</td>
<td>CR + limb leads</td>
<td>8.75 ± 2.05</td>
<td>9.19 ± 2.46</td>
<td>9.94 ± 1.39</td>
</tr>
<tr>
<td>Year</td>
<td>27</td>
<td>V + limb leads</td>
<td>8.37 ± 2.04</td>
<td>9.04 ± 1.79</td>
<td>9.41 ± 2.0</td>
</tr>
<tr>
<td>Year</td>
<td>27</td>
<td>CR + limb leads</td>
<td>8.04 ± 2.24</td>
<td>9.19 ± 1.86</td>
<td>9.70 ± 1.88</td>
</tr>
</tbody>
</table>

n = number of subjects; SD = standard deviation. See text for statistical results using the method of generalized estimating equations.
Figures 1A-1D

No RAZ ("normal")

Abboud (A) RAZ

Abboud Percent (AP) RAZ

NASA (N) RAZ
Figure 2

RMS, month-to-month

RMS, year-to-year

Kurtosis, month-to-month

Kurtosis, year-to-year
Figure 3

RMS, month-to-month

RMS, year-to-year

Kurtosis, month-to-month

Kurtosis, year-to-year