



**HAL**  
open science

# Normative values, reliability and sample size estimates in heart rate variability

Gavin R Sandercock

► **To cite this version:**

Gavin R Sandercock. Normative values, reliability and sample size estimates in heart rate variability. Clinical Science, 2007, 113 (3), pp.129-130. 10.1042/CS20070137 . hal-00479380

**HAL Id: hal-00479380**

**<https://hal.science/hal-00479380>**

Submitted on 30 Apr 2010

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Heart rate variability measures: a fresh look at reliability

Pinna G, Maestri R, Torunski A, Danilowicz-Szymanowicz L, Szwoch D, La Rovere MT, and Raczak G, Heart Rate Variability Measures: A fresh look at reliability. *Clin Sci (Lond)*, 2007.

By Dr. Gavin Sandercock.

### Abstract.

Heart rate variability (HRV) is a non-invasive marker of cardiac autonomic modulation utilised in many hundreds of scientific studies each year. The reliability of heart rate variability has been frequently investigated yet remains poorly quantified. Assessing the reliability of a measure that assesses dynamic physiological processes and shows large between and within-subject variation is a complex task. In this issue of *Clinical Science* Pinna et al. [1] provide excellent insight into the test-retest reliability of commonly used HRV indices and put the values obtained into context by comparing them with levels of between-subject variation and by producing sample size estimates.

### Commentary.

Scientific publications utilising heart rate variability (HRV) as an index of cardiac autonomic modulation are increasing year on year [2]. It would therefore seem logical to conclude that the thousands of authors who publish scholarly articles using this measure are fully aware of its concurrent validity and retest reliability, two qualities that should be easily identifiable in any clinical measurement. Recent discussions [2] and reviews of the literature [3] demonstrate that an adequate consensus on either of these qualities has not been fully reached.

In this issue of *Clinical Science*, Pinna et al. [1] assess the reliability of time and frequency domain measures of HRV, in a well designed and statistically sound investigation. Their study [1] provides much needed information, enabling the reader to determine the value of HRV in terms of retest reliability.

Pinna et al. [1] promote an interesting discussion regarding absolute and relative reliability of HRV indices. Their paper evaluates the reliability of selected measures and suggests potential, theoretical sample size requirements needed during future empirical investigations. The authors conclude that reliability of HRV may appear low, or even unsatisfactory, insofar that even the better-performing indices often double or halve in value from test to retest. When this relatively poor absolute reliability is put into context, the authors are able to conclude that: ‘for most indexes...random variation represents a limited portion of the between-subject variability’.

Assessing the reliability of HRV measures in this way and the creation of sample size estimates is original, innovative and provides the reader with useable information regarding application and use of HRV. What this innovative approach demonstrates is that many of the existing studies which use HRV as a dependent variable lack statistical power. When we reviewed the literature examining reliability of short-term HRV analysis we found that high quality studies were rare [3]. In fact, nearly half the studies identified in our initial search were rejected due to poor design and in particular, poor choice of statistical analysis. One previous study of high methodological quality [4] provided findings highly consistent with those of Pinna et al.[1], both in terms of reliability and of absolute values for HRV indices obtained. Sinnreich et al. [4] produced, what was then, probably the best quality study in relation to HRV reliability and the present study [1] adds greatly to this body of literature.

The similarity in HRV index values between Sinnreich et al.[4] and Pinna et al. [1] leads to a second noteworthy point. Pinna et al. [1] raise an important point by using simulated data for sample size estimates because: ‘reference values as to what change in HRV parameters would be clinically relevant are lacking’. It has long been known that depressed levels of selected HRV indices are excellent predictors of mortality in some populations [5] but what represents a normal value for HRV on which to base sample size estimates proves elusive.

Measures of central tendency and dispersion show us that there appears to be a genuine and large inter-subject variation in many HRV indices, particularly spectral measures.

Similarly, some studies show very large changes (and therefore large effect sizes) in HRV indices due to a variety of interventions. The reliability of HRV must, therefore, be interpreted both in terms of expected values at baseline and predicted magnitude of change. Such interpretation is, however, problematic due in part to the heterogeneity in values reported in the literature.

Time domain values show reasonable homogeneity but spectral measures appear particularly heterogeneous. Although some of the variation will be due to individual differences and methodological inconsistencies, the range of values appears unsatisfactory. To illustrate this point using the high frequency spectral power (HF), I cite values from two papers examining HRV in sportive populations, which used broadly similar data collection methods. In the first, HF was reported as (mean  $\pm$  S.D.)  $5839 \pm 1839 \text{ ms}^2$  [6] whilst in the second, HF was reported as  $3.4 \pm 1.8 \text{ ms}^2$  [7]. Both these values are somewhat disparate from the 1996 Taskforce report [8] which provides normal values for HF as  $975 \pm 203 \text{ ms}^2$ . The problem here lies not in which value is 'correct', this is merely an illustration, but in the fact that such values can be published without reference to the fact that they vary so widely from 'expected' values.

Large-scale, sex and age-determined normative data for short-term HRV do exist [4, 9, 10] but reference to such values rarely is made. Authors and reviewers alike should be cautious when values for spectral measures are radically different from those considered normal and authors in particular should seek to provide reasoning for such discrepancies. Authors should be encouraged to report spectral measures as raw ( $\text{ms}^2$ ) units and provide appropriate measures of central tendency and dispersion. Logarithmic transformation is often necessary prior to parametric analysis, but reporting only the mean ( $\pm$  S.D) of the natural logarithm does not adequately describe the data. The simple addition of appropriate descriptive statistics would greatly enhance the potential for the reader to evaluate the values for spectral HRV measures reported.

Perhaps when a recognised range of values for spectral HRV exists, research into expected magnitudes of change can begin in earnest. Finally, when such values are established, HRV studies with *a priori* power calculations will be more commonplace and HRV may be able to occupy a role as a clinical tool in risk stratification and therapeutic target for intervention.

To conclude, Pinna et al.[1] present an excellent example of an investigation into a very important field of research for HRV application, and provide well reasoned 'ball park' figures for required sample sizes in future intervention studies in which HRV indices are independent variables. Accuracy of future sample size calculations would, however, be much greater if reference values for HRV indices are agreed and cited. With over 500 peer-reviewed publications annually it surely cannot be unreasonable to assume that some consensus toward a normative range for HRV indices can be reached. Perhaps then, sample size calculations in HRV will be as common place as they are in the assessment of other physiological measurements.

#### References.

- [1]. Pinna G, Maestri R, Torunski A, Danilowicz-Szymanowicz L, Szwoch D, La Rovere MT, and Raczak G, Heart Rate Variability Measures: A fresh look at reliability. *Clin Sci (Lond)*, 2007.
- [2]. Parati G, Mancia G, Di Rienzo M, and Castiglioni P, Point: cardiovascular variability is/is not an index of autonomic control of circulation. *J Appl Physiol*, 2006. 101(2): 676-678; discussion 681-672.
- [3]. Sandercock GR, Bromley PD, and Brodie DA, The reliability of short-term measurements of heart rate variability. *Int J Cardiol*, 2005. 103(3): 238-247.
- [4]. Sinnreich R, Kark JD, Friedlander Y, Sapoznikov D, and Luria MH, Five minute recordings of heart rate variability for population studies: repeatability and age-sex characteristics. *Heart*, 1998. 80(2): 156-162.
- [5]. Kleiger RE, Miller JP, Bigger JT, Jr., and Moss AJ, Decreased heart rate variability and its association with increased mortality after acute myocardial infarction. *American Journal of Cardiology*, 1987. 59(4): 256-262.
- [6]. Sacknoff DM, Gleim GW, Stachenfeld N, and Coplan NL, Effect of athletic training on heart rate variability. *American Heart Journal*, 1994. 127(5): 1275-1278.
- [7]. Scott AS, Eberhard A, Ofir D, Benchetrit G, Dinh TP, Calabrese P, Lesiuk V, and Perrault H, Enhanced cardiac vagal efferent activity does not explain training-induced bradycardia. *Auton Neurosci*, 2004. 112(1-2): 60-68.
- [8]. Taskforce, Heart rate variability: standards of measurement, physiological interpretation and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. *Circulation*, 1996. 93(5): 1043-1065.
- [9]. Kuo TB, Lin T, Yang CC, Li CL, Chen CF, and Chou P, Effect of aging on gender differences in neural control of heart rate. *American Journal of Physiology*, 1999. 277(6 Pt 2): H2233-2239.
- [10]. Hemingway H, Shipley M, Brunner E, Britton A, Malik M, and Marmot M, Does autonomic function link social position to coronary risk? The Whitehall II study. *Circulation*, 2005. 111(23): 3071-3077.