ELSEVIER

International Journal of Cardiology 103 (2005) 238–247

International Journal of Cardiology

www.elsevier.com/locate/ijcard

Review

The reliability of short-term measurements of heart rate variability

Gavin R.H. Sandercock*, Paul D. Bromley, David A. Brodie

Research Centre for Health Studies, Buckinghamshire Chilterns University College, Newlands Park, Gorelands Lane, Chalfont St.Giles, Buckinghamshire HP8 4AD, United Kingdom

Received 28 April 2004; received in revised form 10 August 2004; accepted 4 September 2004
Available online 23 February 2005

Abstract

Short-term assessment of heart rate variability (HRV) is a non-invasive technique to examine ANS function. Within the literature, HRV is commonly referred to as a reliable measurement technique. The aim of this review was to assess the accuracy of this description based upon a comprehensive review of the available data concerning reliability of short-term HRV measures.

Reviewing only studies using appropriate statistical analyses, it was determined that reliability coefficients for HRV measures were highly varied. Coefficients of variation ranged from <1% to >100%. Similar variation was found in studies using the intraclass correlation coefficient values, and limits of agreement.

Reliability coefficients reported displayed some distinct patterns. Firstly, where measurements were made during interventions such as tilt or pharmacological stimulation, reliability was poorer than when HRV was measured at rest. Secondly, clinical populations displayed poorer reliability than healthy subjects.

There was little effect of test–retest duration on reliability and although no single HRV measurement appeared less reliable than another, there was evidence that optimal data collection conditions for specific frequency domain measures exist.

Describing HRV as a reliable measurement technique appears to be a gross oversimplification, as results of reliability studies are heterogenous, and dependent on a number of factors. Further studies are required, particularly in clinical populations to assess HRV reliability. Authors should refer to coefficients from similar populations measured under similar conditions when making future sample size calculations.

Keywords: Heart rate variability; Autonomic control; Measurement; Reliability

1. Introduction

It is commonly perceived that a regular heartbeat is a sign of cardiac health. In truth however, the rhythm of a healthy heart is characterised by significant beat-to-beat variability [1]. This heart rate variability (HRV) has been used as a simple, non-invasive technique to examine autonomic nervous function [2,3]. Depressed levels of HRV have been shown to be present in a number of pathological conditions including heart disease [4,5], heart failure [6–9], diabetes [10–13], hypertension [14], asymptomatic left ventricular dysfunction [7], and following myocardial infarction [4,15].

Reduced HRV is a significant predictor of cardiac event including death in coronary artery disease [16], heart failure [17], stable angina pectoris [18,19], and following MI [15,20]. It is also a predictor of sudden cardiac event and death in the elderly [21] and the general population [22].

In addition to clinical applications, HRV measures have been utilised to describe differences in autonomic function in epidemiological studies. HRV measures are known to differ according to age [23,24] and sex [24,25]. Many studies have illustrated the effect of habitual levels of physical activity on HRV. For the most part, studies support the notion that HRV is elevated in active populations [26], although significant differences are not always evident [27].

* Corresponding author.
E-mail address: gsande01@bcuc.ac.uk (G.R.H. Sandercock).

0167-5273/$ - see front matter © 2005 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.ijcard.2004.09.013

There are also limited longitudinal data to suggest HRV can be modified by increased physical activity [28,29], although data are again inconsistent [30,31]. The findings of these studies have been summarised in recent reviews [32,33].

In these reviews, and commonly within the literature, HRV is described as a reliable measure. Several issues arise when such broad statements are made. Firstly, analysis of the variability characteristics of a single ECG can lead to numerous time and frequency domain measures. The Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology [3] gave recommendations on the use of 18 measures derived from the RR interval time series. In addition to this, there is ongoing development of new methods by which this time series may be analysed, such as, coarse grained spectral analysis [34], fractal scaling [35], and wavelet transformation [36].

Numerous reviews are available concerning the physiological meaning of the various HRV measures available [3,32,33] that also describe the various measurement techniques used in capturing ECG signals for HRV analysis. Briefly, there are two general methodologies by which ECG traces HRV analysis can be obtained. These are 24-h ambulatory recordings and short (usually 5–15 min) recordings. Ambulatory recordings can be used to generate all time and frequency domain measures of HRV and although time domain analysis can be carried out on short-term recordings, these are more suited to analysis via spectral methods. The possible measures achieved from short-term recordings are somewhat limited, as they are unable to measure accurately some of the very slow oscillations observed in longer recordings. However, short-term recordings do have several advantages of uncontrolled ambulatory ECG measurements. Firstly, they are quick to perform and to analyse. Five-minute recordings contain a number of individual RR intervals which can be edited manually in only a few minutes. Secondly, short-term recordings can be made under controlled conditions and supervised by a physician or researcher to ensure standardisation. Thirdly, they can be made under a variety of conditions such as before and after tilt or other postural, psychological or pharmacological interventions. Such comparative recordings allow researchers to observe autonomic responses to such stimuli.

1.1. Aims

The first aim of this article is, therefore, to systematically review the available literature concerning reliability of the various HRV measures from short-term ECG recordings. From this, we hope to determine whether HRV measures made at rest or during specific interventions are reliable in healthy subjects and clinical populations. The effect of time from test to retest will also be assessed as will the use of specific types of HRV measures (time domain and frequency domain). Finally we will make recommendations concerning the use of repeated HRV measurement and give direction for future possible research.

2. Study selection based on reliability analysis

For intra-individual change to be monitored, the method used must be reliable. Reliability has a number of definitions and for a full review of the concept the reader is directed to one of the excellent articles which exist in this area [37,38]. Concerning the application of a measurement such as heart rate variability (HRV) the researcher or clinician should be interested in how much a measurement, made on an individual or group, varies when it is repeated. This is akin to absolute reliability defined by Baumgarter [39] as the degree to which repeated measurements vary for individuals.

Inclusion criteria for the present review were as follows: The study was undertaken with the specific aim of assessing the reliability of HRV; studies were required to have used repeated measurements to examine reliability; the investigators were required to have given a methodological account of standardised procedures used to control for possible confounding variables that may have affected reliability.

Critical analysis of the statistical methods used to assess reliability was undertaken and only studies using appropriate statistical analysis, as defined by Atkinson and Nevill [37] were included. Four studies were rejected for the use of inappropriate statistical analyses (see Table 1) including the use of Pearson product moment correlation [40–42] or merely examining differences between groups [43]. Neither association between scores nor the absence of significant differences between test and retest quantify reliability of repeated measurements [44,45]. One further study was rejected due to the use of a unique repeatability coefficient which was non-comparable with the remaining data [46]. One further study used sampling periods which were too short (1 min) and had insufficient sample size to give any meaningful assessment of reliability [47].

Acceptable quantitative assessments of reliability included the intraclass correlation coefficient or ICC [48,49] where used appropriately [50] utilising current recommendations for result interpretation [51], the coefficient of variation (CV) and the limits of agreement (LoG) method [44,52]. Studies meeting these criteria were reviewed systematically, according to data collection method and subject population. For the purpose of this review, a short-term test is defined as any test that does not use 24-h ambulatory monitoring. Such tests can be subdivided into two further categories:

1. Stable, resting protocols, where conditions remain consistent throughout the ECG recording, and resting HR is measured.