Background information to a poster presented at the 16th International Acupuncture Research Symposium, King's College London, 29 March 2014

The fickleness of data: Estimating the effects of different aspects of acupuncture treatment on heart rate variability (HRV). Initial findings from three pilot studies.

© Tony Steffert (Open University) and David Mayor (University of Hertfordshire)

ABSTRACT

Background. Heart rate variability (HRV) is a measure of the interplay between sympathetic and parasympathetic influences on heart rate. Higher HRV is usually associated with relaxation and health benefits, lower HRV with stress/pathology. HRV is used increasingly in acupuncture research. Electroacupuncture (EA) and transcutaneous electrical acupoint stimulation (TEAS) are frequently used modalities, variants of manual acupuncture (MA). This is the fourth of a series of conference posters from a study investigating the effects of EA and TEAS on HRV and the EEG (electroencephalograph).

Objectives. To assess how treatment factors – particularly stimulation frequency (Hz) – contribute to changes in HRV.

Methods. Three small pilot studies were conducted. All intervention and monitoring 'segments' lasted for 5 minutes. In Pilot 1 (*N*=7, 12 visits in all), 5-minute electrocardiograph (ECG) monitoring followed each intervention segment. In Pilot 2 (*N*=12, 48 visits) & Pilot 3 (*N*=4, 16 visits), 5-minute monitoring and stimulation were concurrent; ECG and then photoplethysmography (PPG) were used, and HRV (or pulse rate variability, PRG) derived from raw interbeat interval data following standard procedures, including artefact processing. Stimulation was at different combinations of the acupoints LI4 and ST36, and at either 2.5 Hz or 10 Hz. Eight HRV/PRV measures were selected for analysis. For each factor, numbers of significant differences in these measures were counted (*N*), and normalised percentage differences calculated (Diff%). In addition, coefficient of variance (CV), Cohen's *d* (effect size) and correlation ratio *eta* (η) were computed for the differences in measures induced by the various factors.

Results. Several methods of assessing differences suggested a small, non-significant difference in HRV measures in favour of 2.5 Hz. However, most of these could be explained by intrinsic variation (CV) of the measures rather than as a specific effect of stimulation frequency.

Further analysis. There were highly significant correlations between *N*, Diff%, *d* and η for the treatment factor comparisons made (e.g. stimulation frequency, amplitude, location, visit, participant and baseline values of five main HRV measures). The sum of η^2 for all factors considered was 0.678, suggesting that >2/3 of factors responsible for variance in outcomes were identified. This variance was mostly dependent on differences among participants, and least on stimulation frequency.

Conclusions. The analytical methods employed are accessible even to those with little statistical expertise. They offer a simple way of assessing the contribution of different experimental factors to outcomes when statistical significance is elusive and sample size is small. They are thus be appropriate for application in acupuncture research, which tends to involve a number of independent variables in small-scale studies. However, a mixed models approach and multivariate analysis should also be used to analyse new and existing results, with Bootstrap to ensure a sufficiently large sample size.

In this study, the effects of stimulation frequency on HRV are likely to be masked by those of other treatment factors.

CONTENTS

3. Changes in HRV values

3. Changes in values over 5-minute segments (**Hz**)
3. Changes in HRV values from baseline to follow up (**Hz**)

4.Beneficial effect ratio (**BER**)

4. Beneficial effect ratio (**BER**) (**Hz**, **Mod**)
4. Beneficial effect ratio (**BER**) (**Visit**)
4. Beneficial effect ratio (**BER**) (**Loc**)
4. Beneficial effect ratio (**BER**) (**ID**)

5. Ratios of high and low measures relative to the group median

5. Ratios of high and low measures (**Hz**)
5. Ratios (**Visit**)
5. Ratios (**Loc**)
5. Ratios (**Mod**)
5. Ratios (**ID**)
5. Ratios (**ID**, **Mod**)

FURTHER ANALYSIS

Baseline comparisons: Test-retest reliability (**TTR**)

High or low baseline HRV measures as another factor in outcome

Coefficients of variance (normalised SD) and effect size

5. Ratios – Coefficients of variance (normalised SD)
5. Ratios – Effect size using modified Cohen's *d*
5. Ratios – Comparing CV and Cohen's *d*

Diff% for the various methods of analysis (Values, Correlations, BER and H/L ratios) used in this study

Partial correlations between Cohen's *d*, CV and Diff% for the various methods of analysis used

Example: Diff% and nSD (CV) for Pilot 2 (segments EA1 to EA4)

Correlation ratio *eta* (η) for factors in this study

Interaction between factors (independent variables) – the χ^2 test

BACKGROUND

Heart rate variability (HRV)

Heart rate variability (HRV) is a measure of the continuous interplay between sympathetic and parasympathetic influences on heart rate (HR), and is considered to communicate information about autonomic flexibility and the capacity for regulated emotional response [Appelhans & Luecken 2006].

Assessment of HRV from the electrocardiograph (ECG) is an established technique in medicine [Gevirtz 2011], standardised since 1996 [Malik 1996]. It is used increasingly in medical research, as is pulse rate variability (PRV) from the photoplethysmograph (PPG), although the two methods are not completely interchangeable, particularly in disordered breathing or mentally stressful conditions [Dehkordi et al. 2013; Khandoker et al. 2011; Schäfer et al. 2013; Wong et al. 2012].

A number of HRV measures are available, all derived from the R-to-R (RR) inter-beat interval of the ECG. Corresponding PRV measures are derived from pulse cycle intervals [Schäfer et al. 2013] (**Fig 1**).

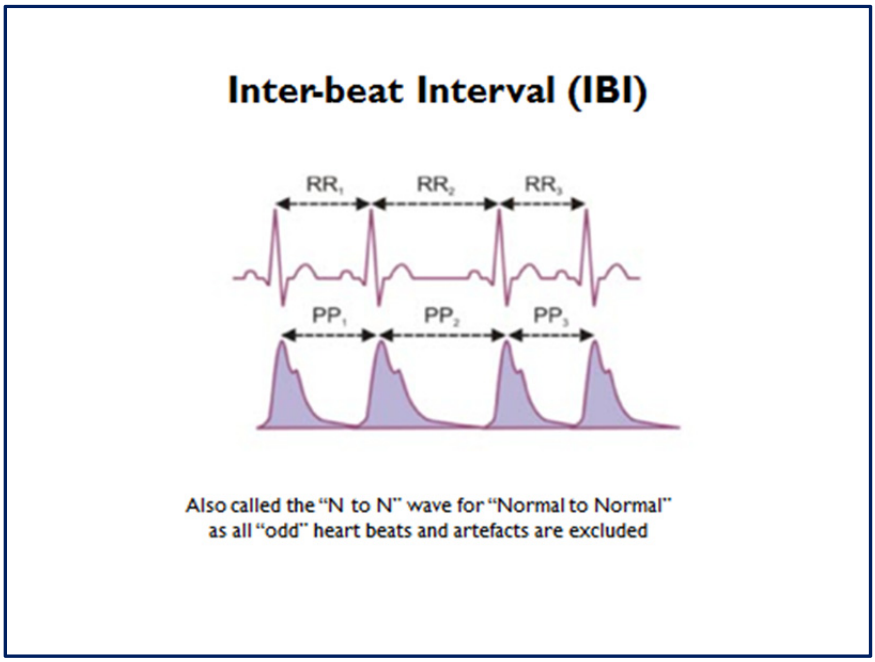


Fig 1. Inter-beat interval in HRV and PRV.

HRV is used increasingly in medical research, as can be seen from searching PubMed (**Fig 2**).

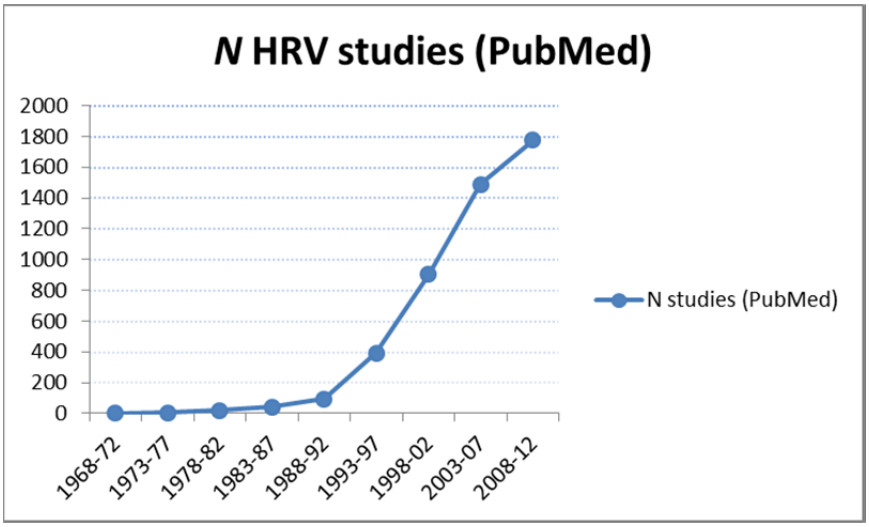


Fig 2. Numbers of studies in PubMed for consecutive 5-year periods when searching for 'HRV (NOT rhinovirus)' (15 Feb 2012).

For simplicity, here the term HRV will be used generically, although data was gathered using a PPG in the majority of Pilot 2 sessions and in all Pilot 3 sessions.

In general, reduced HRV is associated with ageing [Frewen et al. 2013; Fuller-Rowell et al. 2013; Nunan et al. 2010; Russoniello et al. 2013; Umetani et al. 1998] and increased risk of morbidity in many different conditions [Chang et al. 2014; Fagundes et al. 2011; Javorka et al. 2005; Kim et al. 2005; Kim et al. 2006; Lackschewitz et al. 2008; Lee et al. 2011; Malik 1996; Milovanovic et al. 2009; Zulli et al. 2005]. It is also found, for example, in heavy drinkers [Thayer et al. 2006] and stress-precipitated smoking [Ashare et al. 2012], as well as in chronic smokers (and even in the offspring of

the latter) [Dinas et al. 2013]. Of particular interest here is the association of reduced HRV with workplace stress [Rieger et al. 2014], anxiety [Cervantes Blásquez et al. 2009; Pittig et al. 2013], a potentiated startle reflex [Ruiz-Padial et al. 2003] and a tendency to panic [Friedman & Thayer 1998]. In contrast, those with higher HRV tend to perform better than those with low HRV on taxing cognitive tasks or in stressful situations (for example, the threat of electric shock) [Hansen et al. 2009]. Increased HRV is often considered an objective measure of improved subjective relaxation [Bothe et al. 2014; Lee et al. 2013; McFadden et al. 2012; Markil et al. 2012].

HRV measures – overview

HRV can be evaluated using different methods, usually grouped under the headings of 'time domain', 'frequency domain', 'rhythm pattern analysis' and 'nonlinear methods'. Rhythm pattern analysis was not used here.

Time domain methods are based on the normal-to-normal (NN) or RR interval (**Fig 1**). For short recordings (~5 minutes), frequency domain methods are more readily interpreted physiologically than time domain methods [Acharya et al. 2006; Faust et al. 2012; Malik 1996]. They focus specifically on frequency changes and power spectral density, estimated non-parametrically using the simple Fast Fourier Transform (FFT) algorithm (Welch's periodogram).

HRV reflects the stochastic autonomic input to the heart [Baillie et al. 2009]. Efferent vagal activity is accepted as a major contributor to the high frequency component of HRV (HF, 0.15-0.4 Hz), whereas the LF component (LF, 0.04-0.15 Hz) may include both sympathetic and vagal influences [Malik 1996]. Thus the LF/HF ratio is considered by some to mirror sympatho/vagal balance, and by others to reflect sympathetic modulation [*ib*.]. In any case, an increased LF/HF ratio is often associated with stress in some form or other [Cervantes Blásquez et al. 2009; Sauvet et al. 2009]. Thus, unlike HF absolute power (HFpwr) and the other HRV measures used in these Pilots, which *increase* with parasympathetic activation and reduced stress, LF/HF may *decrease*.

The normal RR series has been described as 'nonchaotic, nonlinear, and multifractal ' [Baillie et al. 2009], and so has frequently been subjected to nonlinear analysis. Reduced (fractal) complexity and stronger regularity may indicate deactivation of control loops within the cardiovascular system and a diminished adaptability of the cardiac pacemaker [Schubert et al. 2009]. Such an increase in more highly ordered dynamics has been associated with several pathologies, including Parkinson's disease (tremor), obstructive sleep apnoea, sudden cardiac death, epilepsy and foetal distress syndrome [Vaillancourt et al. 2002]. Many recent studies have used nonlinear methods to explore this field, although there is some lack of agreement on whether nonlinear methods are less [Malik 1996] or more [Schubert et al. 2009] sensitive than linear ones. In any case, nonlinear methods have been found particularly suited to short HRV records [Khandoker et al. 2009] and so to test-retest evaluations [Maestri et al. 2007].

Because it was suspected that application of a regularly repeating stimulus as with EA/TEAS might reduce HRV complexity, three nonlinear methods were selected in this pilot study: Correlation dimension (D_2) [Carvajal et al. 2005; Melillo et al. 2011], Approximate entropy (ApEn) [Carvajal et al. 2005; Richman & Moorman 2000] and Sample entropy (SampEn) [*ib*; Mohebbi & Ghassemian 2012]. All three measure the complexity or irregularity of the RR series, albeit in different ways [Richman & Moorman 2000; Yang et al. 2001]. Historically, D_2 was used initially in HRV studies, followed by ApEn,

and then SampEn. As for the linear measures, higher values of D_2 and ApEn indicate lower predictability [Nazeran et al. 2006], and reduced values may be predictive of morbidity, mortality [Pincus 2001] or stress [Mellilo et al. 2011]. The same is true for SampEn, a similar but less biased measure particularly suited to shortterm ECG recordings [Bornas et al. 2006; Henry et al. 2010; Khandoker et al. 2009; Lake et al. 2002; Richman & Moorman 2000; Vuksanović & Gal 2005]. Both D_2 and SampEn were found to increase significantly in one study of reflexology [Joseph et al. 2004]. However, increases in SampEn have not always been associated with beneficial findings [Ahamed et al. 2006; Akar et al. 2001; Mateo et al. 2012].

HRV reliability and changes over time

Short recordings of indices such as HF and total power repeated after several months show that their stability (test-retest reliability) is excellent (0.76-0.80 in one study of 70 healthy subjects), with that for HF the best [Alraek & Tan 2011]. However, many HRV measures can vary widely between individuals even within the same study, particularly HF [Nunan et al. 2010]. Relative reliability, the degree to which individuals maintain their position in a sample with repeated measurements, can be contrasted with absolute reliability, the degree to which repeated measurements vary for the individual, even if relative reliability is maintained [Anon n.d.]. In healthy subjects, there may be relatively large day-to-day random variations in HRV (i.e. low absolute reliability), which may make the detection of intervention effects using HRV difficult in individual participants [Sookan & McKune 2012].Similarly, HRV measurements in type 2 diabetics are characterised by poor absolute reliability but substantial to good relative reliability [Sacre et al. 2011]. In general terms, linear HRV indices show worse absolute reliability than nonlinear ones [Maestri et al. 2007; Sookan & McKune 2012].

Shortterm measures of HRV rapidly return to baseline after transient perturbations induced by mild exercise and other interventions, but take longer to do so following more powerful stimuli, such as maximum exercise [Malik 1996]. On this basis, it was thought suitable to make several repeat recordings of HRV during each participant visit.

HRV and acupuncture

Acupuncture research using both HRV and PRV has also become more frequent in recent years, so that currently in PubMed nearly 2% of all HRV and PRV studies are on acupuncture-related topics (whereas less than 0.1% of all studies indexed in PubMed are on acupuncture) [PubMed searches]. However, only recently have acupuncture-based HRV studies started to investigate the effect of using different acupuncture points [Kaneko et al. 2013; Litscher et al. 2013; Matsubara et al. 2011; Wu et al. 2009; Yang et al. 2013], and none appear to have explored using different frequencies of electroacupuncture (EA) or transcutaneous electrical acupoint stimulation (TEAS).

A literature review of HRV changes in response to acupuncture conducted in 2012 as a basis for the present study found the following:

HF power may decrease [EA: Chang et al. 2005] or increase [acupressure: Matsubara et al. 2011; MA: Haker et al. 2000; Hsu et al. 2007; Kurono et al. 2011; Li et al. 2003], or initially increase and then decrease [MA: Streitberger et al. 2008], or only increase after stimulation [Haker et al. 2000]. In conscious rats, EA increased HF [Imai et al. 2009].

LF power may decrease [MA: Agelink et al. 2003; Bäcker et al. 2008; Hsu et al. 2007] or increase [MA: Haker et al. 2000; Li et al. 2003; EA: Chang et al. 2005]. Whether LF power decreases or increases may depend on stimulation location, both with MA [Uchida et al. 2010] and EA [Imai et al. 2009].

LF/HF may decrease (Acupressure: Arai et al. 2011; EA: Imai et al. 2008; MA: Agelink et al. 2003; Chae et al. 2011, Hwang et al. 2011; Wang et al. 2011) or increase (MA: Streitberger et al. 2008 (shortterm); EA: Chang et al. 2005; Wu et al. 2009]. Again, whether LF/HF decreases or increases may depend on stimulation location [MA: Saito ; EA: Imai et al. 2009].

Thus:

- HR in general decreases (with MA or EA).
- HF power may decrease or increase, usually the latter (sometimes with MA it may increase and then decrease, or only increase following stimulation).
- LF power may decrease or increase (with MA or EA).
- LF/HF may decrease or increase (with MA or EA).
- Fatigue and location of stimulation may affect directions of change.

OBJECTIVES

To apply manual acupuncture (MA), electroacupuncture (EA) and transcutaneous electrical acupoint stimulation (TEAS) to healthy participants using a standard protocol, and assess changes in HRV due to the following factors:

- Stimulation frequency (**Hz**) [*primary objective*]
- Stimulation location (**Loc**)
- Stimulation duration (**Dur**)
- Stimulation amplitude (**Amp**)
- Stimulation modality (**Mod**)
- Participant (**ID**)
- Visit (**V**)
- Baseline HRV (**B**)

Ranges tested:

Hz	2.5 Hz, 10 Hz
Loc	B ('Bottom', ST36[2]), L ('Left' LI4 & ST36), R ('Right' LI4 & ST36), T ('Top', LI4[2]), Bilat* (L & R), LLSS* (B & T)
Dur	5, 10, 15 or 20 minutes
Amp	Pilot 1: 2.5-10.0 dial units; Pilot 2 and Pilot 3 (EA): 0.2-8.8 mA; Pilot 3 (TEAS): 0.2-24.7 mA
Mod	Pilots 1 and P3 (TEAS): TEAS; Pilots 2 and P3 (EA): MA and EA
ID	Pilot 1: 7 participants; Pilot 2: 12 participants; Pilot 3: 4 participants (all of whom also took part in Pilot 1)

[* In Pilot 1 only]

Visit Pilot 1: 2 visits; Pilots 2 and 3: 4 visits
Baseline Values of all 8 HRV measures at baseline (EO1)

METHODS

Participants

Ethics committee approval for this study was obtained from the University of Hertfordshire, provided that participants were professional acupuncturists or other complementary health practitioners with prior experience of acupuncture. Healthy volunteers were recruited from members of the Acupuncture Association of Chartered Physiotherapists and the British Acupuncture Council, and from local practitioners known to the lead researcher (DM).

Exclusion criteria were past head injury, epilepsy, current cancer, wearing of an implanted electronic device, dependence on psychoactive medication and pregnancy.

ECG/PPG were used to gather data as a basis for assessing HRV/PRV in a study with the wider aims of investigating the effects of TEAS and EA on the electrical activity of the brain (using electroencephalography, EEG) and the heart.

Participants were asked to abstain from consuming caffeine, nicotine, alcohol or a heavy meal for at least two hours before attending for a session. They were also asked to avoid any strenuous activity to which they were not accustomed for two hours before or after attending a session.

On arrival, participants were seated in a comfortable chair with arms. An explanation of the experiment was provided, after which they had the opportunity to ask questions and then signed a consent form and completed some brief state questionnaires (they had already received detailed information about the study and completed several online background and trait questionnaires beforehand). Any wrist bangles or bracelets were removed, and ECG electrodes or a PPG were positioned. Subjects were asked to remain relaxed but awake. To avoid unduly affecting the HRV, they were instructed to 'breathe normally'. Talking during recording was discouraged, and in general the atmosphere in the room was one of calm concentration. TS took charge of the recording and timing, DM of the stimulation.

Protocols

Three small pilot studies were conducted. All intervention and ECG/PRV monitoring 'segments' lasted for 5 minutes.

In **Pilot 1**, 5-minute ECG monitoring *followed* each intervention segment. In **Pilots 2 & 3**, 5-minute monitoring and stimulation were *concurrent*.

Acupoint and electrical parameter factors tested are as listed above, under Objectives.

In **Pilot 1** (TEAS), all point combinations were used in every session, in balanced order. Five participants attended for two sessions (2.5 Hz or 10 Hz TEAS), two for only one session each.

In **Pilot 2**, one point combination was used per session, and 12 participants attended for four sessions. In **Pilot 3**, two combinations were used per session, and four participants from Pilot 1 (a year before) also attended for four sessions, each experiencing four of a possible eight interventions.

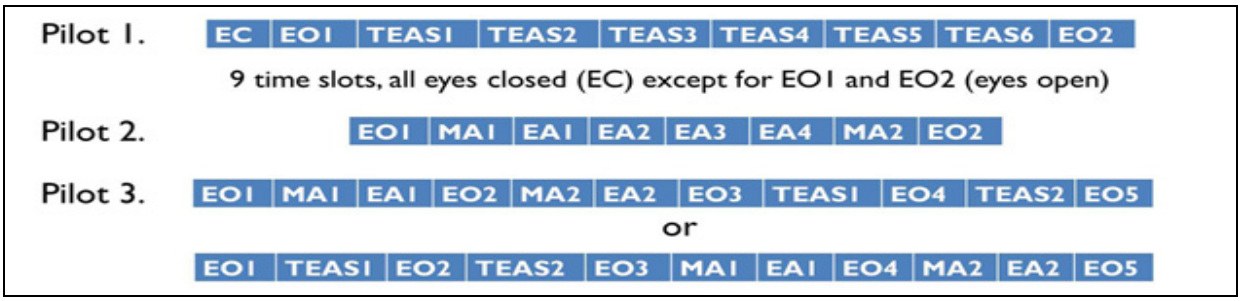


Fig 3. Order of 'segments' in each Pilot

TEAS in Pilot 1 was carried out using an Equinox stimulator (Equinox, Liverpool). In Pilots 2 and 3, both EA and TEAS were from a Classic4 stimulator (Harmony Medical, London). Acupuncture needles (Classic Plus, 25 mm x 0.22 [HMD Europe]) and self-adhering electrodes (Stimex, 32 mm diam. [schwa-medico, Ehreingshausen]) were also supplied by Harmony Medical.

Data collection

Procedures used for HRV data collection and analysis followed the accepted standards [Malik 1996]. In Pilot 1 and the first two sessions of Pilot 2, the EEG-202 [Mitsar, St Petersburg]) was used to gather ECG data from three 24 mm diameter disposable gel electrodes [ARBO, Henleys Medical Supplies, Welwyn Garden City], with passive and ground electrodes on one forearm (usually the right) and the active electrode on the other forearm. After that, a Nexus Blood Volume Pulse Sensor was used as the PPG, usually attached to the forefinger or middle finger of the right hand. A sampling rate of 1024 Hz was used. Data was sent from the NeXus-10 physiological amplifier via Bluetooth link to a laptop for processing using BioTrace software. The inter-beat interval was calculated in BioTrace, and then exported as a text file for further processing.

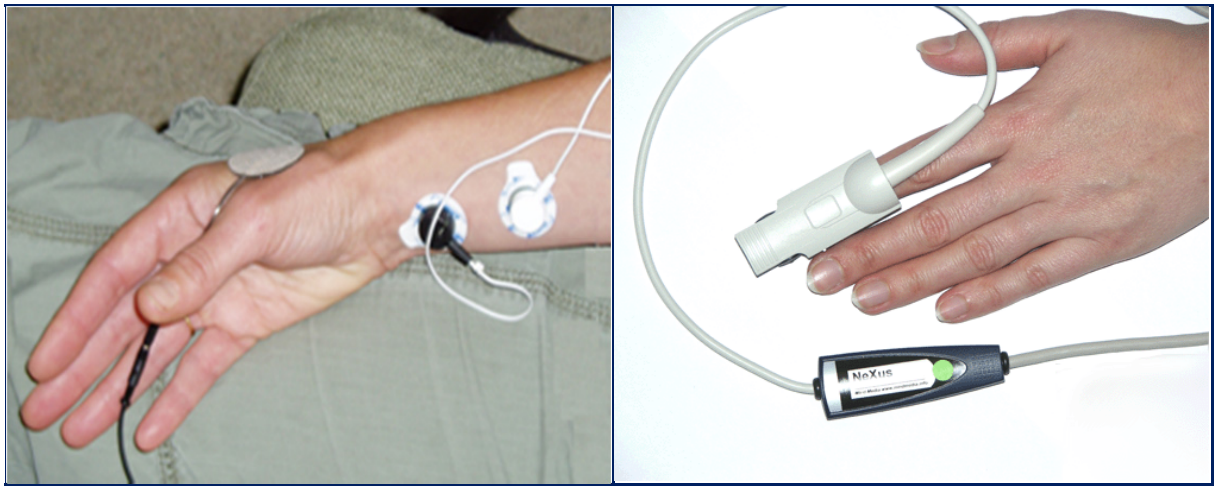


Fig 4. Sensors used. *Left*: Passive and ground ECG electrodes on right forearm, with TEAS stimulation electrode at LI4. *Right*: BVP sensor on left forefinger.

Data was processed using open-access Kubios HRV software v 2.0 (Biosignal Analysis and Medical Imaging Group, University of Eastern Finland: [Kubios HRV 2012], commonly used in HRV studies. Following visual inspection of raw records for ectopic beats, missing data and noise, artefact

preprocessing was conducted in Kubios using a 'medium' setting with 'smoothness priors' detrending to reduce the requirement that data for the nonlinear measures should be tested for nonlinearity and stationarity prior to HRV determination. The standard HRV frequency bands described above were used. Other Kubios default options adopted were 256 points/Hz for spectrum estimation, 256 second windowing with 50% overlap for FFT spectrum analysis, and embedding dimension *m* of 2 for ApEn and SampEn (tolerance 0.2 x SDNN), but 10 for D_2 (threshold 3.1623 x SDNN) [Tarvainen & Niskanen 2012].

HRV measures

Of the 47 possible HRV measures available as outputs from Kubios (9 time domain, 26 frequency domain, 12 nonlinear), a number were discounted because of only being suited to longterm (e.g. 24-hour) monitoring, others because of their known variation with respiration and emotional state, and others because of difficulties of interpretation. The remaining eight measures were considered appropriate for our purposes:

Time domain (3)

- **RR** Mean R-R interval (ms)
- **SDNN** R-R standard deviation (ms)
- **RMS SD** Root mean square of successive differences (ms)

Frequency domain (FFT spectrum using Welch's periodogram) (2)

- **HFpwr** HF power (mA^2)
- **LF/HF** LF/HF power ratio,

Nonlinear (3)

- **ApEn** Approximate entropy
- **SampEn** Sample entropy
- **D_2** Correlation dimension.

Analysis

Differences in these HRV measures for the above experimental factors were assessed using:

1. **Values** of the HRV measures over all segments during which stimulation was applied
2. **Correlations** between these values
3. **Changes** in HRV values between session baseline and follow-up segments
4. **BER** ('beneficial effect ratio') for a series of segments, defined as:

$$\frac{\Sigma\ (N \text{ increases in 7 measures}) + (N \text{ decreases for LF/HF})}{\Sigma\ (N \text{ decreases in 7 measures}) + (N \text{ increases for LF/HF}) + 1}$$

A BER > 0.8 indicates a beneficial effect, and one of <0.8 a non-beneficial effect.

5. **Ratios** of 'high' and 'low' measures relative to the group median, either for segments during which stimulation was applied, or comparing baseline and follow up.

In addition to the statistical significance of these assessments, 'normalised percentage differences', **Diff%**, were calculated. For example, value Diff% for Hz is defined as:

$$\frac{\text{(Value at 10 Hz) – (Value at 2.5 Hz)}}{\text{(value at 2.5 Hz)}} \text{ X 100}$$

For a comparison among > 2 factors, the mean Diff% for all comparisons was taken.

Statistical analysis was carried out using SPSS v20 (IBM 2011) and Excel v14 (Microsoft 2010).

RESULTS and initial analysis

(Data on which these results and analysis are based can be found in Appendices at the end of this document.)

1. Values

Table 1. Numbers of significant differences in the 8 HRV measures for the various factors over stimulation segments (*after* segments in Pilot 1; *during* segments in Pilots 2 and 3).

Pilot	Hz	Loc	Dur	Amp	Mod	ID	V	Baseline	**All**
Pilot 1	2 (3)	0 (0)	n/a	2 (6)	n/a	8 (8)	1 (4)	(6)[a]	13 (27)
Pilot 2	0 (1)	0 (1)	0 (0)	5 (5)	n/a	8 (8)	1 (0)	(5)[a]	14 (20)
Pilot 3 (EA)	1 (1)	0 (1)	n/a	5 (5)	0 (0)	5 (6)	0 (0)	(6)[a]	11 (19)
Pilot 3 (TEAS)	0 (0)	0 (0)	n/a	1 (1)	0 (0)	5 (6)	0 (0)	(5)[a]	6 (12)
All	3 (5)	0 (2)	0 (0)	13 (17)	0 (0)	26 (28)	2 (4)	(21)[a]	**46 (77)**

T-tests or 1-way ANOVA with Bootstrap were used except for Baseline, not yet computed (Mann-Whitney or Kruskal-Wallis test counts in parentheses). a. Averages over 5 initial measures, rounded to nearest whole number.

This allows a rough estimate of the contribution of each factor to the changes in HRV that result from stimulation. Although 77 out of 256 (8 x 8 x 4) possible comparisons (30%) were significant when using non-parametric tests, 66 of these, or more than 25% were attributable to the effects of stimulation amplitude, participant ID and values at baseline.

Table 2. Numbers of significant differences in each HRV measure for the various factors over stimulation segments (non-parametric results and all results for Baseline not yet entered).

	RR	SDNN	RMS SD	HFpwr	LF/HF	ApEn	SampEn	D_2	All

Hz	(0)	(0)	(1)	(1)	(0)	(2)	(1)	(0)	(5)
Loc	(0)	(0)	(0)	(0)	(0)	(2)	(0)	(0)	(2)
Dur	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)
Amp	2 (3)	2 (3)	1 (2)	0 (2)	1 (1)	1 (1)	3 (2)	3 (3)	13 (17)
Mod	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)
ID	(4)	(4)	(4)	(4)	(2)	(2)	(4)	(4)	(28)
V	(0)	(1)	(1)	(0)	(1)	(0)	(1)	(0)	(4)
All	(7)	(8)	(8)	(7)	(4)	(7)	(8)	(7)	**(56)**

Significant comparisons occurred for all HRV measures 7 or 8 times, except for LF/HF, for which significant differences only occurred 4 times.

1.Values (**Hz**)

In Pilot 1 (P1), only three HRV measures demonstrated significant differences for the two stimulation frequencies: RMS SD, HFpwr and SampEn. In Pilot 2 and Pilot 3 (EA segments), only ApEn showed significant differences.

Analysing actual differences in mean values rather than their significance was more informative. To make comparison across the different measures meaningful, normalised percentage differences (Diff%) were considered (as defined above):

Measures showing greatest absolute (un-signed) Diff% for the two frequencies in more than one Pilot were **SDNN** (2), **RMS SD** (2) and **HFpwr** (2); those showing least differences were RR (3) and SampEn (3), with similar results when HF/LF is used instead of LF/HF. Thus the first three of these measures might be more sensitive to differences in stimulation frequency than the last two.

However, it should be noted that SDNN, RMS SD and HF anyway showed greater intrinsic variation (coefficient of variance, CV, or normalised standard deviation, nSD) than RR and SampEn, regardless of stimulation frequency:

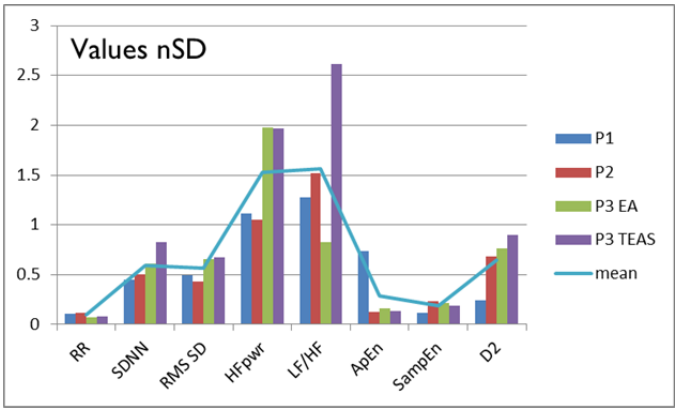


Fig 5. Normalised standard deviation (coefficient of variance) of HRV measures for *all* stimulation segments.

Unpacking nSD for the individual EA segments in Pilot 2 shows a very similar pattern:

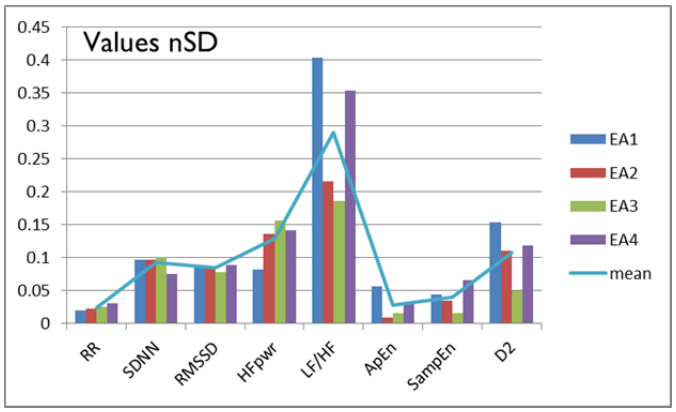


Fig 6. Normalised standard deviation (coefficient of variance) of HRV measures for *individual* stimulation segments in Pilot 2 (P2).

The same pattern is also recognisable in nSD at baseline:

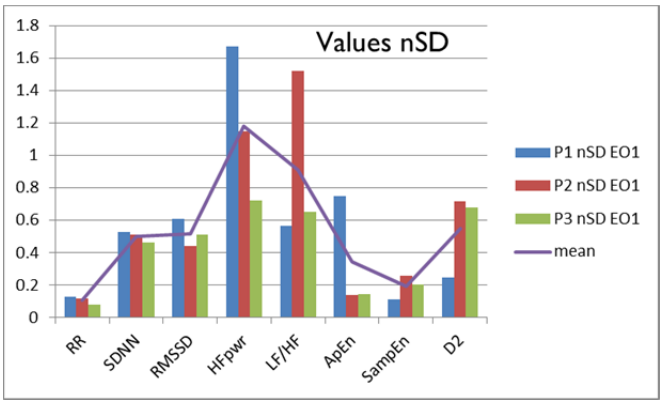


Fig 7. Normalised standard deviation (coefficient of variance) of HRV measures for *baseline* stimulation segments (EO1) in all Pilots.

This can also be visualised as in **Fig 8** (note that the CVs here were normalised separately, so that this comparison is one of overall pattern, not of numerical values).

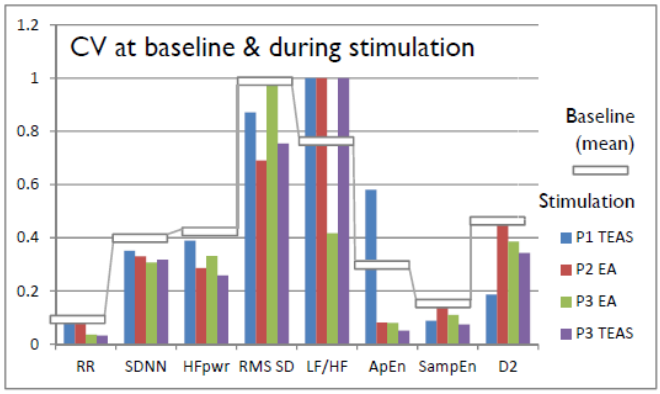


Fig 8. Comparing CV (nSD) patterns at baseline and for stimulation segments (values normalised separately).

Although LF/HF (and D_2 to a certain extent), like SDNN, RMS SD and HFpwr showed higher nSD than RR and SampEn, they do not appear to be particularly responsive to frequency.

As **Table 3** below shows, when the sign of Diff% is considered, in P1 (TEAS), P3 (EA) and P3 (TEAS), it was negative for more measures than it was positive. In other words, in these Pilots, especially P3 (EA), more HRV measures were **higher for 2.5 Hz than for 10 Hz**.

Table 3. Normalised percentage differences (Diff%) between means (10 Hz – 2.5 Hz)

10 Hz – 2.5 Hz	RR	SDNN	RMS SD	HFpwr	LF/HF	ApEn	SampEn	D_2	+ -	mean
P1	1.427	-16.966	-23.221	-16.610	-7.816	-7.015	3.524	16.163	3 + 5 -	-6.314
P2	1.534	-1.134	4.752	11.098	-1.140	3.557	-0.407	-7.060	4 + 4 -	1.400
P3 EA	-0.441	-22.293	-22.427	-60.011	-21.185	-12.136	13.240	-5.668	1 + 7 -	-16.365
P3 TEAS	2.316	-33.652	-25.029	-57.361	-62.626	-3.129	0.528	4.200	3 + 5 -	-21.844
mean	1.209	-18.51	-16.481	-30.721	-23.192	-4.681	4.221	1.908	3 + 5 -	-10.781
+ -	3 + 1 -	0 + 4 -	1 + 3 -	1 + 3 -	0 + 4 -	1 + 3 -	3 + 1 -	2 + 2 -	**11 + 21 -**	**-10.780**

However, if HF/LF is considered rather than LF/HF, only in P3 was this the case (**Table 4**), although now the mean Diff% for all HRV measures is negative for all Pilots.

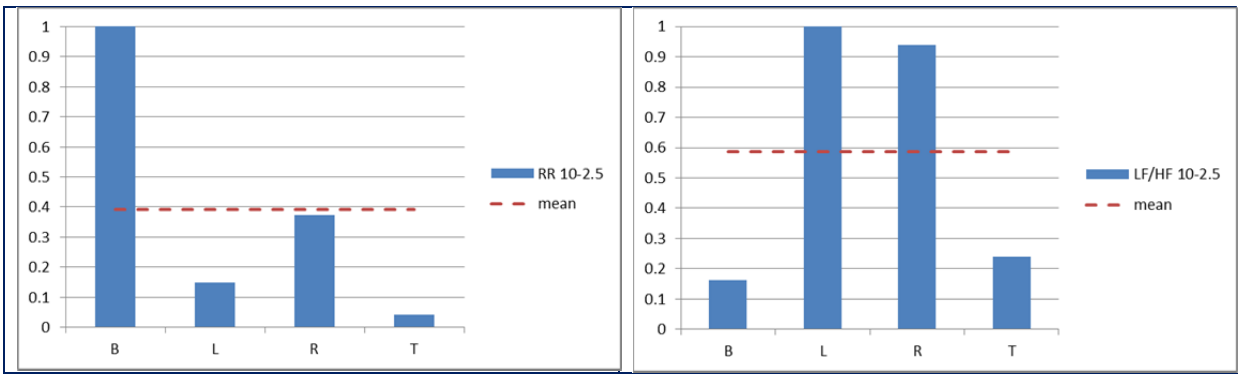
Table 4. Diff% between means when HF/LF is considered instead of LF/HF.

10 Hz – 2.5 Hz	**HF/LF**	+ -	mean
P1	**2.891**	4 + 4 -	-4.976
P2	**-17.715**	4 + 4 -	-0.672
P3 EA	**-2.203**	1 + 7 -	-13.992
P3 TEAS	**-15.311**	3 + 5 -	-15.930
mean	**-8.085**	3 + 5 -	-8.893
+ -	1 + 3 -	**12 + 20 -**	-8.892

Further details on differences in HRV measures for the two frequencies can be found in Appendix C.

1. Values (**Hz**, **Loc**)

In Pilot 2, averaged over the four EA segments, greatest difference between the two frequencies occurred for two of the measures at each location (B, L, R and T), as in the examples in **Fig 9**.



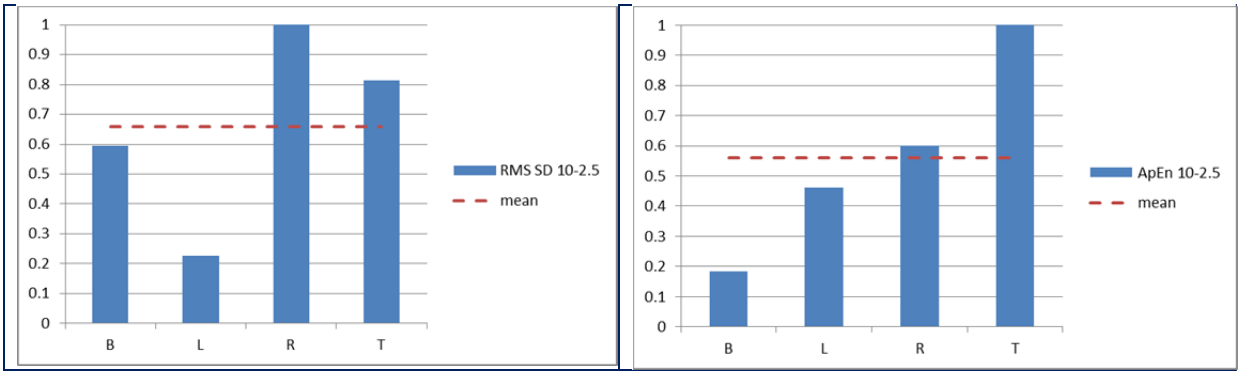


Fig 9. Mean absolute differences (10 Hz – 2.5 Hz) in HRV measures at different stimulation locations (Pilot 2, stimulation segments).

The number of measures greater than the mean for all Locations was as follows:

L 2, B and R 4, T 5 (excluding SDNN, for which the mean value at T was only 0.004, or 0.4%, less than the mean for all Locations). In Pilot 2, more measures showed significant differences for the two frequencies when stimulation was at R or T points, rather than L or B (see Appendix C for details).

Thus here **T appeared to be the Location where greatest differences might be found**.

1.Values (**Dur**)

If there is a general 'relaxation effect' over the course of a session, regardless of the intervention used, mean RR is likely to increase over time. The graphs in **Fig 10** compare results for the three Pilots.

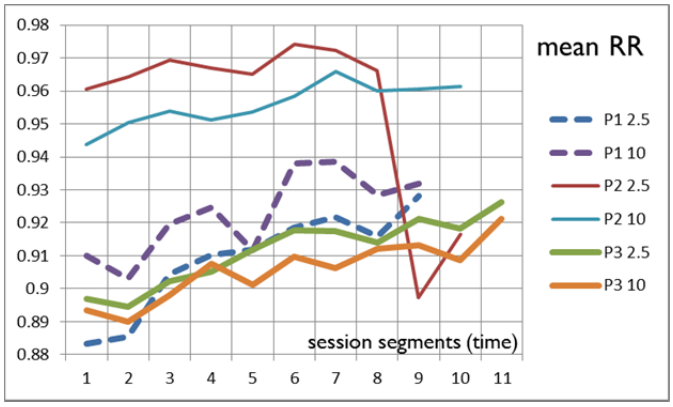
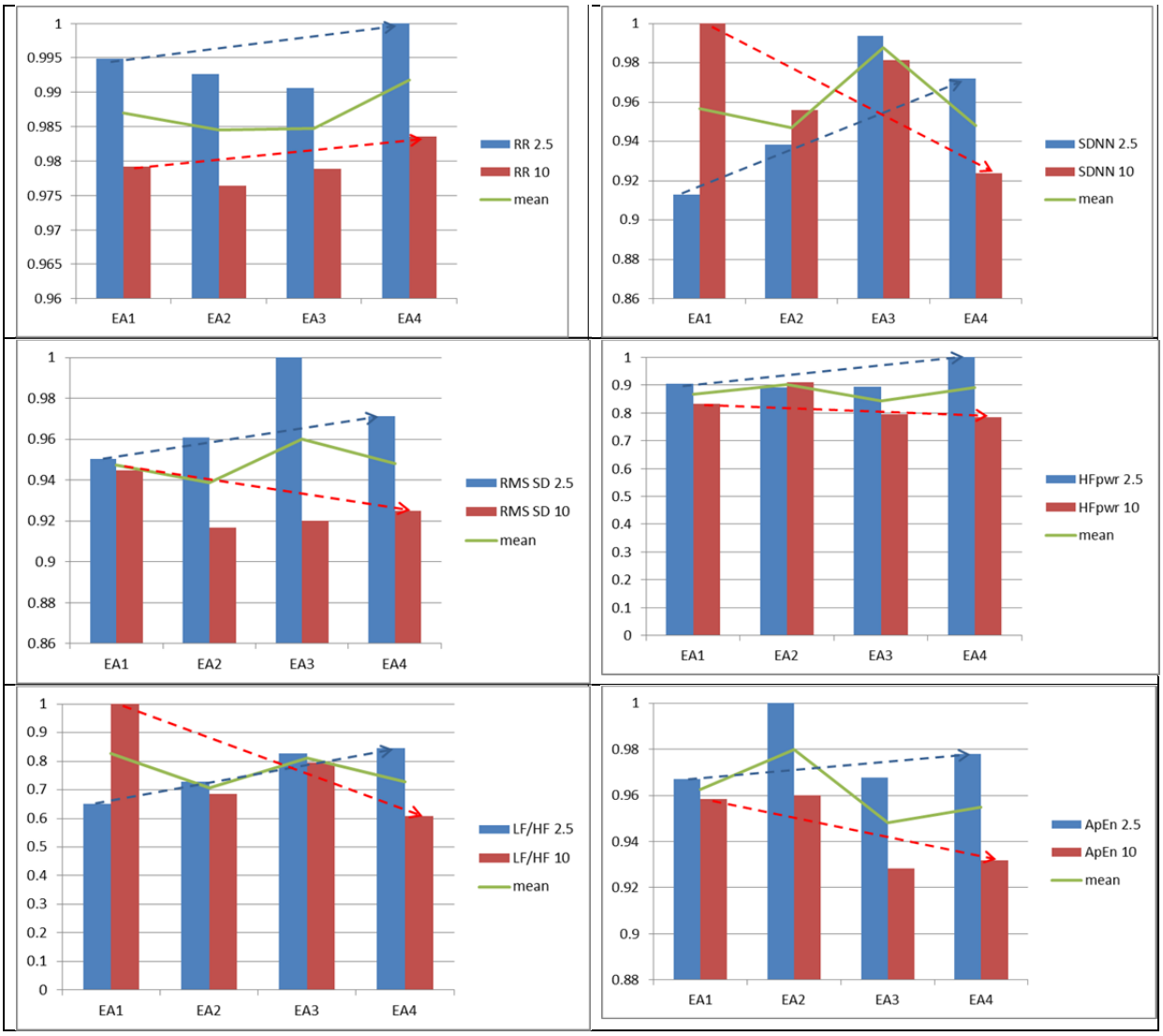


Fig 10. Changes in RR over time, suggesting a small 'relaxation effect' that may not be due to stimulation.

For both stimulation frequencies, RR can be seen to increase over time in Pilot 1 (9 segments) and Pilot 3 (11 segments). In Pilot 2, which in most sessions consisted of 8 segments, the same pattern holds, although there is a decrease at both frequencies between segment 7 (MA2) and segment 8 (EO2). At 2.5 Hz, this decrease started after the last segment of EA (segment 6), and RR decreased dramatically after segment 8. Segments 9 and 10 were added to assess the effect of extended monitoring; needles had been removed, and no further stimulation was provided (*N*=4 for 2.5 Hz, *N*=7 for 10 Hz).

The sudden decrease in mean RR in segment 9 appears to be mostly attributable to one session with participant (6899) who at baseline, even before stimulation, commented “I’m very tense, I don’t know why. My body feels tense. My jaw is feeling tense even though I know it’s not supposed to”. During segment 9, this participant reported that a pre-existing shoulder pain “[feels] in spasm now … more aware of it as the rest of me is more relaxed; I can feel it twitching”.

For segments EA1 to EA4 in P2, changes in the various measures over time (EA1-EA4) are more easily compared if units are normalised. In **Fig 11** this has been done by equating the maximum in the comparison with 1.



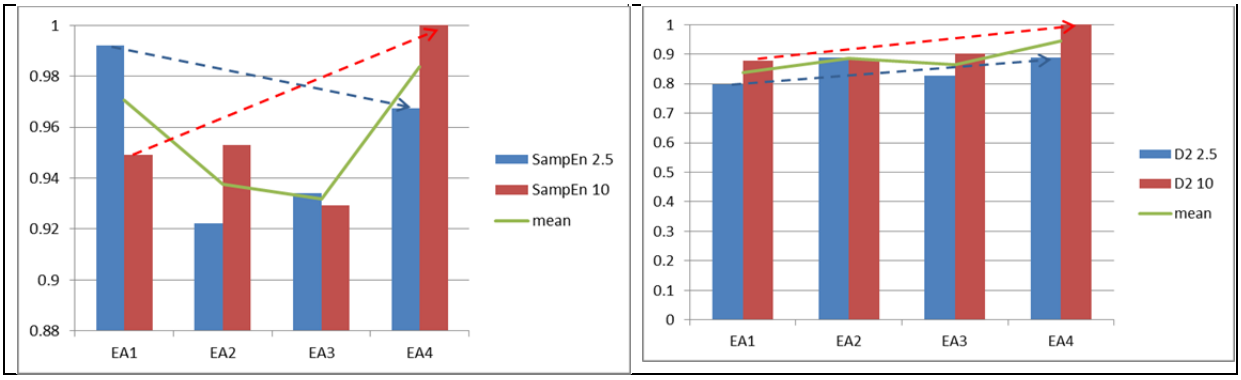


Fig 11. Changes during stimulation in Pilot 2 for each HRV measure taken separately, showing a possible effect of stimulation duration.

For six measures, changes with 2.5 Hz and 10 Hz over the four EA segments were in opposite directions, and for the remaining two in the same direction (**Table 5**).

Table 5. Changes during stimulation in Pilot 2: + increasing; – decreasing.

	2.5 Hz	10 Hz
RR	+	+
SDNN	+	–
RMSSD	+	–
HFpwr	+	–
LF/HF	+	–
ApEn	+	–
SampEn	–	+
D_2	+	+
All	7 +	5 –

Thus most measures **increased over time at 2.5 Hz**, but **more decreased than increased at 10 Hz**.

However, these differences may well be due to the inherent variability of the measures themselves (cf **Figs 6-8** above).

1. Values (**Amp**)

Transforming Amp into a 2-valued variable Amp-N, with Amp-N = 1 if Amp ≥ median amplitude for that Pilot and Amp-N = 0 if Amp < median amplitude for the Pilot, shows that it has a considerable effect on HRV.

Median values of Amp were:

P1 5.125 units on the Equinox device amplitude dial
P2 1.200 mA (as indicated by the Classic4 programming screen)
P3 EA 1.100 mA
P3 TEAS 4.700 mA.

The significance of the resulting differences in HRV for the two Amp-Ns during the stimulation segments are shown in **Table 6** below.

Table 6. Significant differences in HRV measures for stimulation segments when Amplitude is 'high' or 'low'.

Pilot	RR	SDNN	RMS SD	HFpwr	LF/HF	ApEn	SampEn	D_2	**All**
P1	0.000 (0.000)	ns (0.008)	ns (0.075)	ns (0.015)	ns (0.874)	ns (ns)	ns (ns)	0.000 (0.000)	2 (6)
P2	ns [0.005] (0.021)	0.002 (0.002)	ns (ns)	ns (ns)	0.001 [0.003] (ns)	0.047 [0.046] (0.016)	0.021 [0.033] (0.041)	0.011 [0.013] (0.011)	5 (5)
P3 EA	0.001 [0.002] (0.001)	0.046 [ns] (0.012)	0.027 [ns] (0.006)	ns [ns] (0.005)	ns (ns)	ns (ns)	0.029 [0.047] (ns)	0.022 [0.040] (0.040)	5 (5)
P3 TEAS	ns (ns)	ns (ns)	ns (ns)	ns (ns)	ns (ns)	ns (ns)	0.010 [0.010] (0.028)	ns (ns)	1 (1)
All	2 (3)	2 (3)	1 (2)	0 (2)	1 (1)	1 (1)	3 (2)	3 (3)	13 (17)

T-test; [Bootstrap; equal variances not assumed]; (Mann-Whitney U test; 2-tailed asymptotic significance)

In contrast, during the stimulation segments there were *no* significant differences in HRV measures for Loc B versus Loc T (generally the largest differences for the various Loc combinations) in P1, P2 or P3 (EA or TEAS) (See Appendix D). Thus the effect of Amp (Amp-N) on the HRV appears to be far stronger than that of Loc.

2. Correlations

2. Correlations between values (**Hz**)

Taking all Pilots together, there are more significant correlations overall **for 2.5 Hz than 10 Hz** (with more significant at the 0.01 level and fewer at the 0.05 level), but the proportion of these is not significantly different from that expected by chance (**Table 7**).

Table 7. Numbers of significant correlations between HRV measures for all Pilots.

2.5 Hz	54 (44** 10*)
10 Hz	49 (34** 18*)

Spearman's *rho.* ** 2-tailed significance at 0.01 level; * at 0.05 level.

Here Diff%: -9.26%.

2. Correlations between values (**Visit**)

There are more significant correlations between HRV measures in Visit 1 than subsequent visits.

Further data on correlations can be found in Appendix G.

3. Changes in HRV values

3. Changes in values over 5-minute segments (**Hz**)

The overall change in a HRV measure from beginning to end of a session (or part-session) of *n* segments of equal duration is given by $S_n - S_1$, where S_n is the value of the measure at the end of the

session or part-session, and S_1 its value at the beginning. The change during each segment *m* is given by $S_m - S_{m-1}$. Thus $S_n - S_1 = (S_n - S_{n-1}) + (S_{n-1} - S_{n-2}) + \ldots (S_3 - S_2) + (S_2 - S_1)$.

In other words, $S_n - S_1 = n$ x (average segment change).

As we are concerned here more with differences between groups than comparing changes during individual segments within a session, 5-minute segment changes will not be considered further at this juncture.

3. Changes in HRV values from baseline to follow up (**Hz**)

Changes between baseline (EO1) and final session segments (EO2 or EO5, depending on Pilot protocol) were assessed for frequency-dependent normalised differences (Diff):

Diff = [(value at 10 Hz) – (value at 2.5 Hz)]/(value at 2.5 Hz).

Table 8. Normalised differences (Diff) for changes between baseline and follow up.

HRV measure	P1 Diff	P2 Diff	P3 EA Diff	P3 TEAS Diff	10 Hz > 2.5 Hz [A]	10 Hz < 2.5 Hz [B]	**Ratio A/B**
RR	-0.993	2.715	-0.476	1.690	1,1,0,1	0,0,1,0	3:1
SDNN	-34.580	-0.266	2.591	13.273	0,0,1,1	1,1,0,0	2:2
RMS SD	0.268	-1.404	-7.882	0.378	0,0,0,1	1,1,1,0	1:3
HFpwr	-1.025	-0.745	-14.569	2.758	0,0,0,1	1,1,1,0	1:3
LF/HF	-0.379	-0.796	-0.439	-6.996	0,0,0,0	1,1,1,1	0:4
HF/LF	*0.358*	*-0.617*	*-0.525*	*-0.508*	*1,0,0,0*	*0,1,1,1*	*1:3*
ApEn	-3.011	-1.188	-4.578	0.488	0,0,0,1	1,1,1,0	1:3
SampEn	4.123	-3.482	-3.713	-2.158	0,0,0,0	1,1,1,1	0:4
D_2	4.892	-0.557	-1.470	2.066	0,0,0,1	1,1,1,0	1:3
***N* + & – Diffs**	3+, 5– [4+, 4–]	1+, 7– [0+, 8–]	1+, 7– [1+, 7–]	6+, 2– [6+, 2–]	9 [10]	23 [22]	p=0.020 [p=0.050]
mean Diff	-3.838 [-3.746]	-0.715 [-0.693]	-3.820 [-3.828]	1.440 [2.484]	2.866	-2.930 [-2.075]	-0.950 [-0.710]

Results when HF/LF is used rather than LF/HF are indicated in square brackets.
P-values are from the Binomial test for the ratio of positive to negative Diffs.

No differences in any pre-to-post HRV values for the two frequencies were significant, but in each Pilot, except for P3 (TEAS), there was a greater average pre-to-post increase in HRV measures **at 2.5 Hz than at 10 Hz**. The Binomial test of the ratio of negative to positive Diffs showed significance when LF/HF is used (p=0.020), and near-significance when HF/LF is used instead (p=0.050).

Frequency for EA may have more of a differential effect than for TEAS.

Of the different HRV measures, only RR and SDNN tended not to show consistently higher values at 2.5 Hz than 10 Hz. When (absolute) Diffs were ranked, RR and SampEn were each ranked highest or next to highest twice in the four Pilots, and SDNN and D_2 (or HF/LF) were twice ranked lowest or next to lowest. Thus **RR** and **SampEn** might be more sensitive to stimulation frequency than SDNN and D_2. Adding rankings together suggests that, in addition to RR and SampEn, **RMS SD** and **HFpwr** might also vary considerably with stimulation frequency (**Fig 12)**.

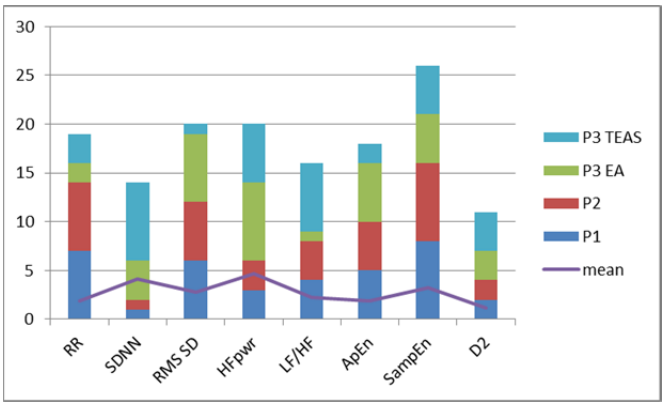


Fig 12. Ranked and summed pre-to-post absolute Diffs (10 Hz – 2.5 Hz) for HRV measures in the 3 Pilots.

However, coefficients of variance for the (absolute) Diffs indicate that changes of HFpwr and SampEn in response to stimulation frequency may in part be due to their intrinsic variability (**Fig 13**).

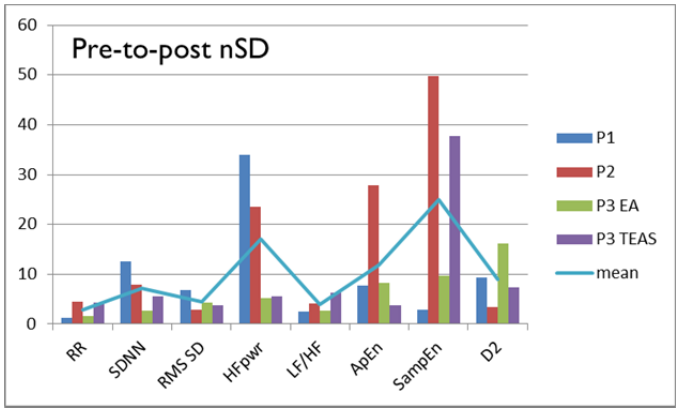


Fig 13. CV (nSD) of pre-to-post absolute Diffs (10 Hz – 2.5 Hz) for HRV measures in the 3 Pilots.

4.Beneficial effect ratio (BER)

4. Beneficial effect ratio (**BER**) (**Hz**, **Mod**)

Table 9 shows the BER for the three Pilots.

Table 9. BER for the three Pilots, data separated for stimulation frequency (Hz) and modality (Mod).

	P1 TEAS	P2 EA	P3 EA	P3 TEAS	mean (all)	mean (EA)	mean (TEAS)
2.5	1.742	0.418	1.167	1.061	1.097	0.793	1.402
10	1.375	0.446	1.504	0.908	1.058	0.975	1.142
Diff%	-21.1	6.7	28.9	-14.4	-3.5	23.0	-18.6
10>2.5	n	y	y	n	n	y	n

Here the **TEAS** Pilots suggest greater BER for **2.5 Hz**, but the **EA** Pilots for **10 Hz**. Mean BER is **greater for 2.5 Hz than for 10 Hz**, and at both frequencies is greater for TEAS than for EA (**Fig 14**). However, none of these findings are statistically significant.

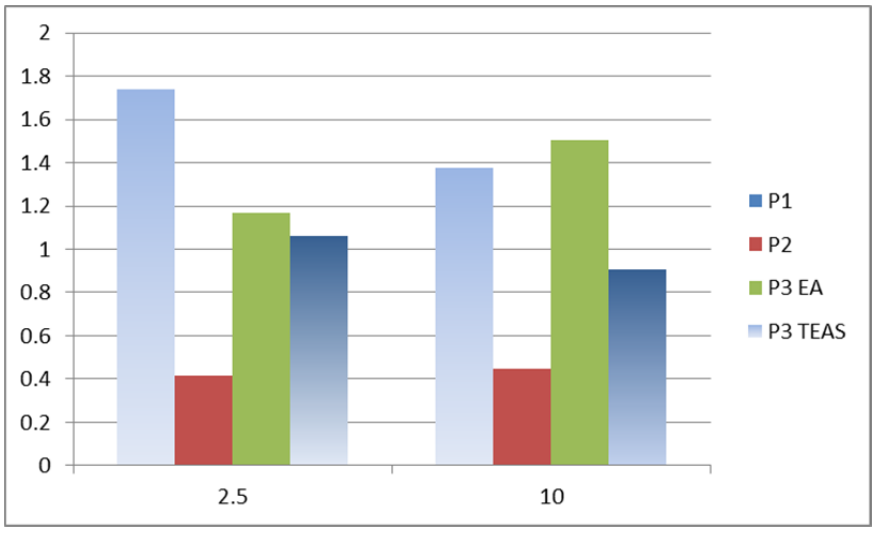


Fig 14. BER for the three Pilots, showing differences between EA and TEAS for the two stimulation frequencies.

4. Beneficial effect ratio (**BER**) (**Visit**)

Table 10 shows the BER for the three Pilots.

Table 10. BER for the three Pilots, data separated for Visit (V) and modality (Mod).

Visit	P1	P2	P3 EA	P3 TEAS
1	1.430	1.164	1.093	1.256
2	1.518	1.084	1.299	0.757
3		1.263	1.023	0.843
4		1.304	1.021	1.083
significance	ns[a]	ns[b]	ns[b]	ns[b]

a. Mann-Whitney U test; b. Kruskal-Wallis test

There is no real consistency of change over visits for the different interventions (**Fig 15**).

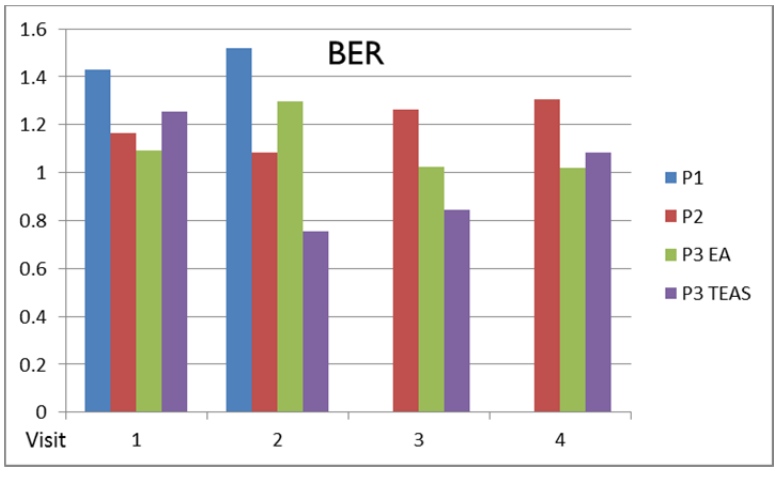


Fig 15. BER for the three Pilots, data separated for Visit (V)

Table 11 shows the same data, separated out by stimulation frequency.

Table 11. BER for the three Pilots, data separated for Visit (V) and Frequency (Hz).

Pilot	Visit	2.5 Hz	10 Hz	diff%	p value
P1	1	1.952	0.845	-56.7	0.001
	2	1.324	1.905	43.9	0.003
P2	1	1.939	1.925	-0.7	ns
	2	1.230	0.939	-23.7	ns
	3	1.218	1.308	7.4	ns
	4	1.082	1.525	40.9	ns
P3 EA	1	1.39	0.80	-42.4	ns
	2	1.38	1.22	-11.6	ns
	3	0.29	1.75	503.4	ns
	4	1.10	0.94	-14.5	ns
P3 TEAS	1	1.43	1.08	-24.5	ns
	2	0.93	0.58	-37.6	ns
	3	0.75	0.94	25.3	ns
	4	1.23	0.93	-24.4	ns

Although there appears to be little consistency of pattern for the normalised difference between $BER_{2.5}$ and BER_{10} across the different interventions, curiously in P2 and P3, $BER_{10} > BER_{2.5}$ in visit 3, but otherwise $BER_{10} < BER_{2.5}$. In P1, $BER_{10} > BER_{2.5}$ in visit 2.

4. Beneficial effect ratio (**BER**) (**Loc**)

Table 12 shows BER for the three Pilots, data separated for Location (Loc) and Frequency (Hz).

Table 12. BER for the three Pilots, data separated for Location (Loc) and Frequency (Hz).

Pilot	Location	2.5 Hz	10 Hz	Diff%
P1	**B (St36[2])**	1.564	0.939	-40.0
	L (Left)	0.818	2.389	192.1
	R (Right)	0.752	0.650	-13.6
	T (LI4[2])	1.038	2.389	130.2
	L & R (Bilat)	2.113	1.175	-44.4
	B & T (LLSS)	2.350	0.357	-84.8
P2	**B (St36[2])**	1.273	1.327	4.2
	L (Left)	1.059	1.251	18.1
	R (Right)	1.333	1.192	-10.6
	T (LI4[2])	1.040	1.155	11.1

(no differences significant)

Differences were in the same direction for L, R and T in the two Pilots, but not for B.

Note that in P3, each BER measure was calculated for two locations, so could not be analysed here.

4. Beneficial effect ratio (**BER**) (**ID**)

BER for individual participants in Pilots 1 and 3 was analysed (**Table 13**).

Table 13. Comparison of BER in Pilots 1 and 3 for individual participants.

ID	P1	P3 EA	normalised difference % (P3 – P1)	P3 TEAS	normalised difference % (P3 – P1)	normalised difference % (P3 TEAS - EA)
8311	0.832					
8875	1.124					
8680	1.127	1.101	-2.3%	1.022	-9.3%	-7.2%
7032	1.148					
2185	1.284	0.964	-24.9%	0.932	-27.4%	-3.3%
8954	1.377	0.937	-32.0%	0.827	-39.9%	-11.7%
5611	2.896	1.433	-50.5%	1.159	-60.0%	-19.1%

BER was less in P3 than in P1 for all repeating participants, whether EA or TEAS is considered, suggesting a possible 'novelty effect' in Pilot 1.

BER was less for TEAS than for EA in Pilot 3, suggesting either a novelty effect (participants were already familiar with TEAS in the experimental setting from P1), or that TEAS is in fact less effective than EA.

Participant 5611 appears to be a 'strong responder' for all three interventions, whereas for 8680 there was least (normalised) difference between BER across the interventions; 8680 could therefore be considered as exhibiting less variation in responsiveness.

Fig 16 shows the variation in participant BER responsiveness in Pilots and 3. Note the difference between 8680 and 5611, for example, in the two Pilots.

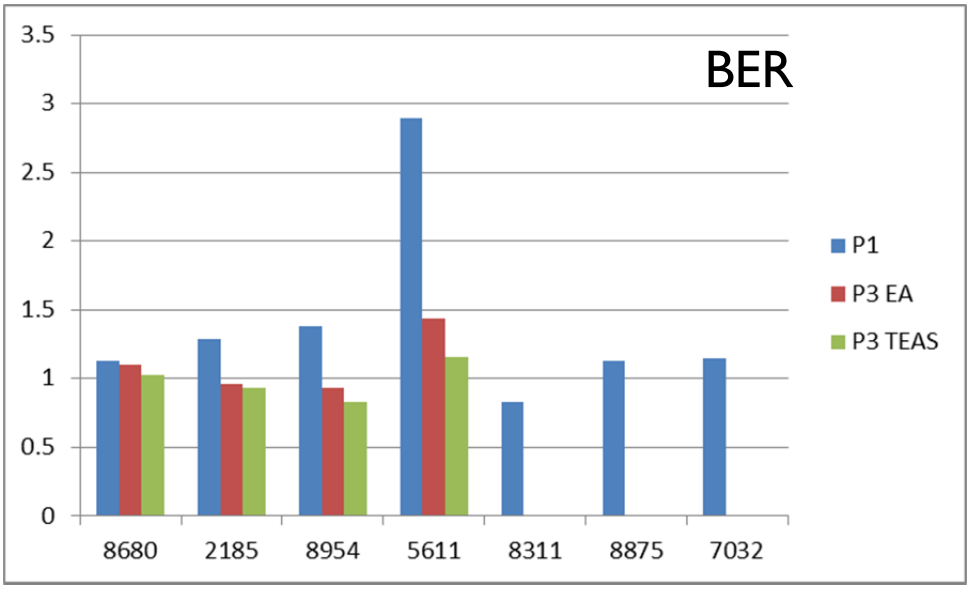


Fig 16. BER for the participants in Pilots 1 and 3.

Table 14 shows the same data, separated out by stimulation frequency.

Table 14. BER in Pilots 1 and 3 for individual participants, separated out by stimulation frequency (BERs in **bold** are those larger for that stimulation frequency than the other).

	P1			P3 EA			P3 TEAS		
ID	2.5	10	diff%	2.5	10	diff%	2.5	10	diff%
2185	**1.272**	1.109	-12.8	**1.220**	0.709	-41.9	0.856	**1.008**	17.8
5611	**4.720**	2.332	-50.6	1.297	**1.570**	21.0	**1.1486**	0.833	-27.5
7032	**1.262**		-34.5						
8311		0.826							
8680	**1.377**	0.601	-56.4	**1.202**	1.001	16.7	**1.083**	0.961	-11.3
8875	0.908	**1.484**	63.4						
8954	0.916	**1.900**	107.4	**0.951**	0.922	-3.0	0.821	**0.832**	1.3

8680 is the only participant for whom the difference between 2.5 and 10 Hz is consistent across all three interventions (with BER greater for 2.5 Hz than 10 Hz). This echoes the result above, where 8680 showed least variation in responsiveness.

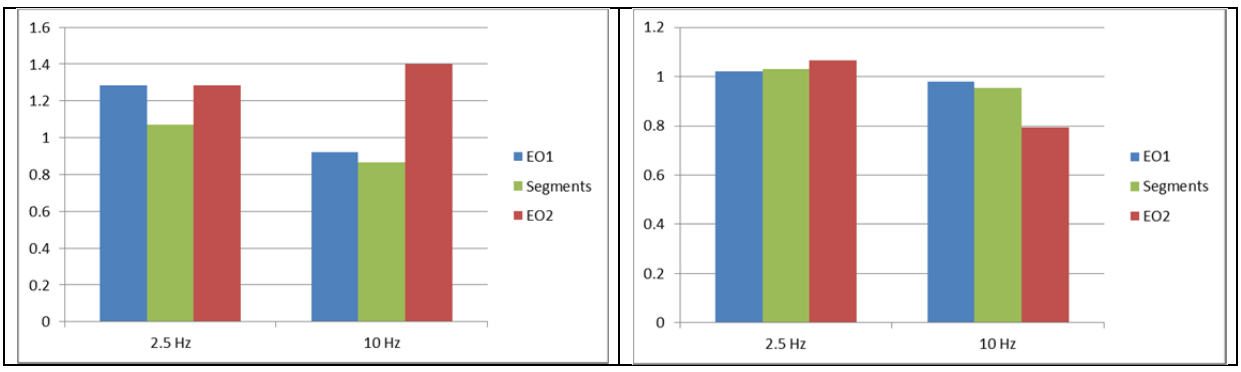
5. Ratios of high and low measures relative to the group median

5. Ratios of high and low measures (**Hz**)

Taking all stimulation segments together in each Pilot, and considering all HRV measures, ratios of number of high to number of low measures relative to the median for each Pilot were as shown in **Table 15** and **Fig 17** (with 'high' and 'low' reversed for LF/HF).

Table 15. Ratios of number of high to number of low measures relative to the median for each Pilot at start (EO1), end (final EO) and in stimulation segments (with 'high' and 'low' reversed for LF/HF).

Pilot	Segments	2.5 Hz	10 Hz	start	2.5 Hz	10 Hz	end	2.5 Hz	10 Hz
P1 TEAS	3 to 8	1.071	0.867	EO1	1.286	0.920	EO2	1.286	1.400
P2 EA	3 to 6	1.032	0.954	EO1	1.021	0.979	EO2	1.065	0.794
P3 EA	EA1, EA2	1.000	2.354	EO1 or EO3	1.000	0.882	EO3 or EO5	1.065	0.829
P3 TEAS	TEAS1, TEAS2	0.969	0.969	EO1 or EO3	1.286	1.000	EO3 or EO5	0.778	0.730
means		1.018	1.286		1.148	0.945		1.049	0.938



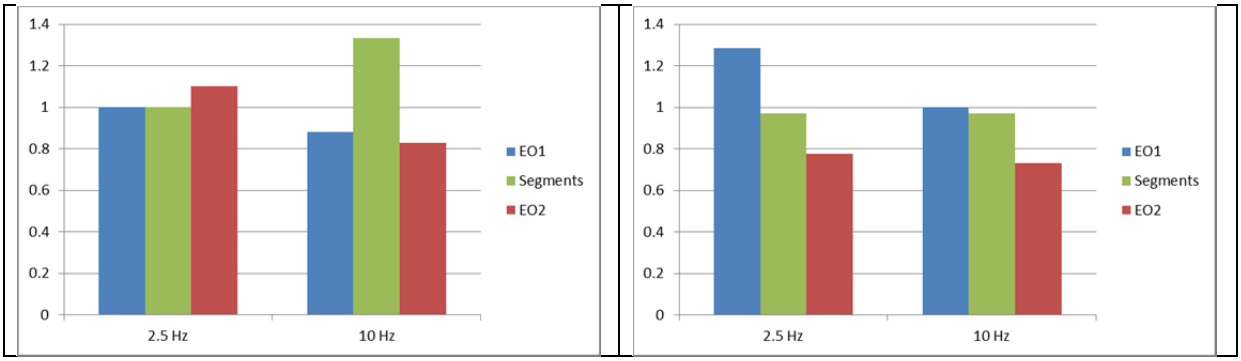


Fig 17. *Top Left*: Pilot 1; *Top Right*: Pilot 2; *Bottom Left*: Pilot 3 EA; *Bottom Right*: Pilot 3 TEAS.

At first sight, it appears that in P1 and P2, but not P3, there are more 'high' (H) than 'low' (L) values of HRV measures (H/L ratio > 1) for 2.5 Hz than 10 Hz stimulation. However, this may be due to a **difference at baseline**: in all three Pilots, ratios for 2.5/10 Hz >1 prior to stimulation ('start' column), and this inequality was maintained post-stimulation ('end' column) except in P1.

Thus no conclusions on the effects of stimulation frequency can be drawn from these figures.

5. Ratios (**Visit**)

Note, however, that comparing high to low ratios for **Visit** rather than **Hz**, all three ratios (during Segments, and at start and end) were greater in Visit 1 than in Visit 2 (except for Segments in P2), and more often >1 in Visit 1 (11 instances) than in Visit 2 (3 instances) (**Table 16**).

Table 16a. High/Low (H/L) ratios (stimulation segments only): means for each Visit.

Pilot	Segments	V1	V2	V3	V4
P1 TEAS	3 to 8	1.105	0.796		
P2 EA	3 to 6	1.065	1.122	0.753	1.076
P3 EA	EA1, EA2	1.000	0.939	0.600	2.368
P3 TEAS	TEAS1, TEAS2	1.207	0.524	0.778	1.783
means					

Table 16b. High/Low (H/L) ratios (start and end segments only): means for each Visit.

Pilot	start	V1	V2	V3	V4	end	V1	V2	V3	V4
P1 TEAS	EO1	1.333	0.818			EO2	1.435	1.222		
P2 EA	EO1	1.133	0.959	1.000	0.920	EO2	0.920	0.846	0.778	1.182
P3 EA	EO1/3	1.000	0.524	1.286	1.133	EO3/5	1.000	0.882	0.684	1.286
P3 TEAS	EO1/3	1.286	1.000	0.882	1.462	EO3/5	1.133	0.391	0.684	1.000
means		1.188	0.825	1.056	1.172		1.122	0.835	0.715	1.156

This suggests that in general there was a **greater responsiveness in Visit 1**, when participants might not have been sure what to expect, compared with subsequent visits, when they were more familiar with the setting and protocol. This is shown graphically in **Fig 18**.

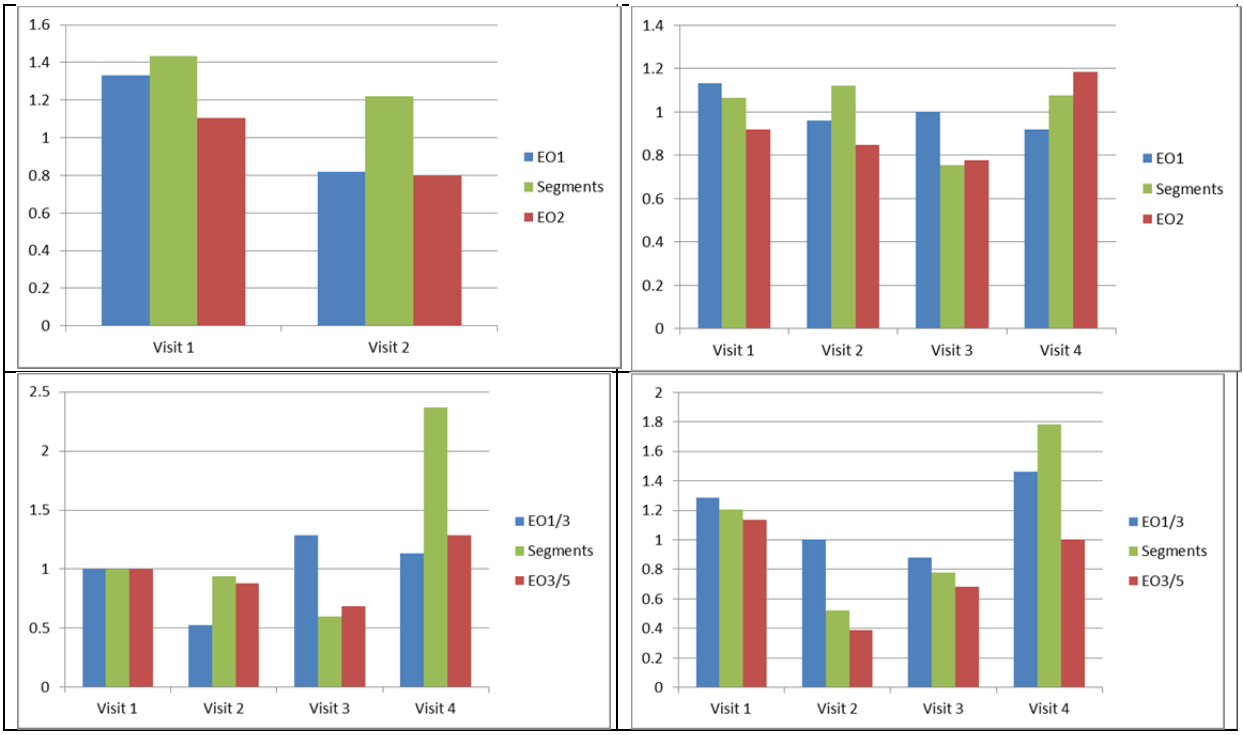


Fig 18. H/L ratios for the different visits in each Pilot, in the start, stimulation and end segments. *Top Left*: Pilot 1; *Top Right*: Pilot 2; *Bottom Left*: Pilot3 EA; *Bottom right*: R Piolt 3 TEAS.

In Pilot 1, Visit 1 showed more high HRV values than V2. In Pilot 2, there was no clear pattern of increase or decrease in H/L ratios. In Pilot 3, there were no clear parallel trends for EA and TEAS, although the last visit appeared to result in highest H/L ratios during stimulation than the other visits. This suggests a possible cumulative effect of treatment (although this was not evident in Pilot 2), but may also be partly a baseline effect (higher for TEAS in Visit 4, but not for EA).

5. Ratios (**Loc**)

Table 17. H/L ratios compared by Location
(note that only in Pilot 2 is it possible to assess EO1 and EO2 for Loc).

Pilot	Segments	B	L	R	T	Bilat	LLSS
P1 TEAS	3 to 8	1.043	1.043	0.778	1.087	1.133	0.778
P2 EA	EO1	1.233	1.042	0.846	0.920		
P2 EA	3 to 6	1.098	0.864	1.110	0.920		
P2 EA	EO2	1.182	0.745	1.087	0.745		
P3 EA	EA1, EA2	1.560	0.939	0.778	1.133		
P3 TEAS	TEAS1, TEAS2	1.065	1.065	1.065	0.730		
means		1.197	0.950	0.944	0.923		

The overall means suggest that higher HRVs (ratio >1) are found at B, then at L, R and lower at T (ratio<1). However, ratios are not greatly different from 1 in any Pilot, except for Pilot 3 (EA), where B shows most high values.

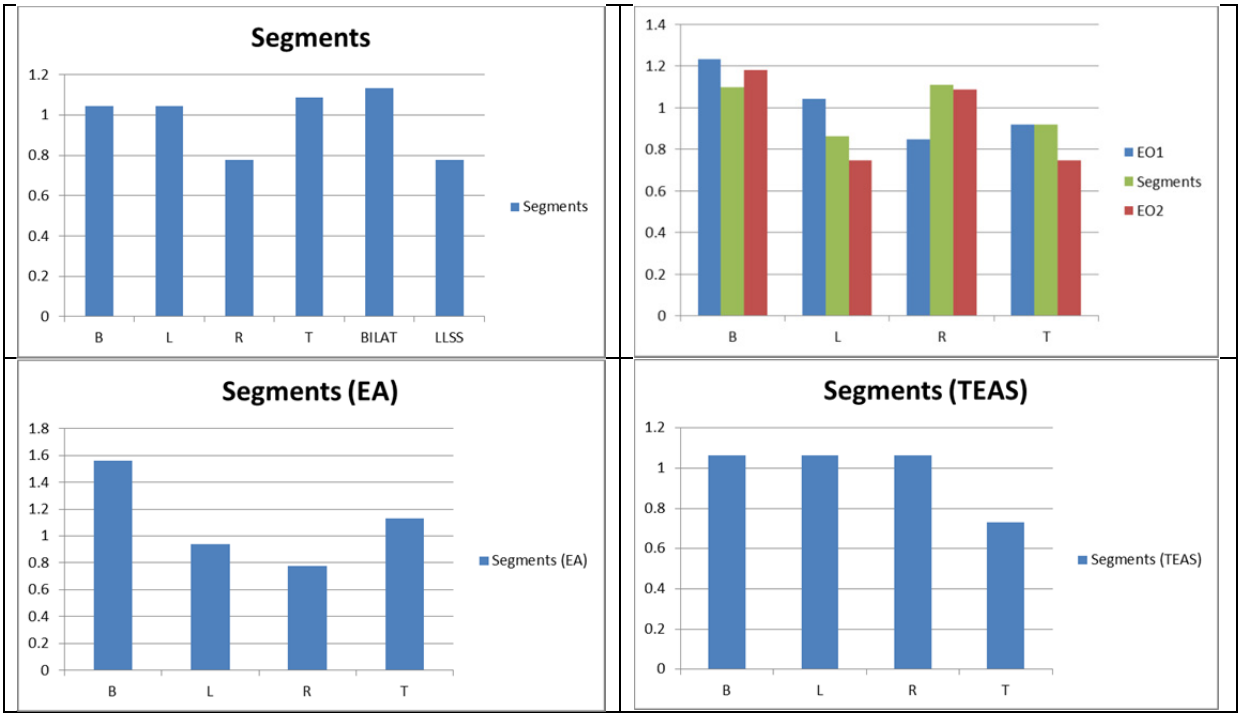


Fig 19. Graphical representation of H/L ratios for the three Pilots. *Top Left*: Pilot 1; *Top Right*: Pilot2; *Bottom Left*: Pilot 3 EA; *Bottom Right*: Pilot 3 TEAS.

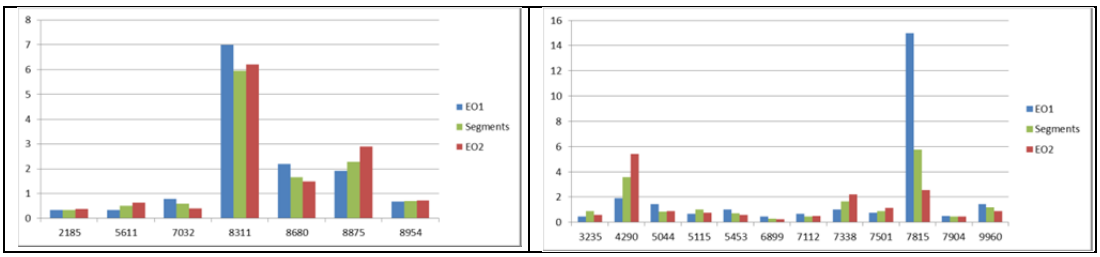
5. Ratios (**Mod**)

Note that for Segments in Table 24 (p. 34), CV is consistently higher for EA than for TEAS in Pilot 3. This is not the case when comparing Pilot 1 (TEAS) and Pilot 2 (EA).

5. Ratios (**ID**)

Some participants scored consistently higher than others in Pilot 1, with differences in EO1 carrying through to EO2. Four showed an overall increase in counts over the course of sessions, 2 a decrease, and 1 a decrease during stimulation compared to before and after.

In Pilot 2, again there were high and low scorers, with 3 showing increasing numbers of 'high' values over the course of a session, 5 decreasing numbers, 2 showing an increase during stimulation compared to before and after, and 2 a decrease during stimulation compared to before and after (**Fig 20)**.



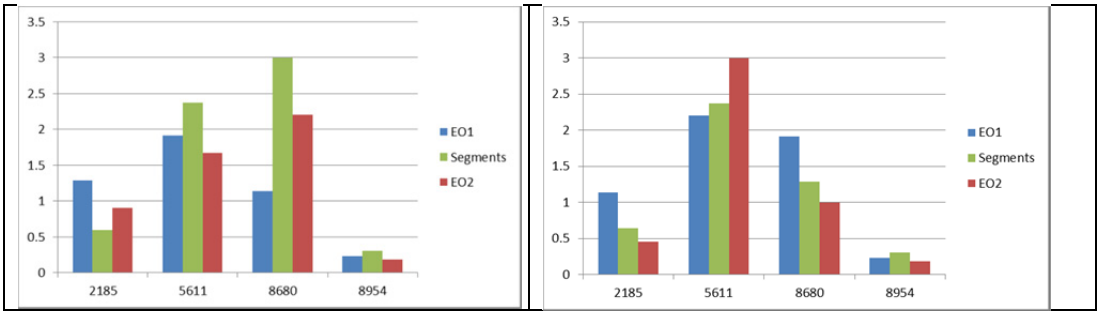


Fig 20. H/L ratios, showing different patterns of change over the course of sessions in each Pilot. *Top Left*: Pilot 1; *Top Right*: Pilot 2; *Bottom Left*: Pilot 3 EA; *Bottom Right*: Pilot 3 TEAS.

Although 8954 showed relatively low H/L ratios in both Pilot 1 and Pilot 3 and 8680 showed middling ratios in both Pilot 1 and Pilot 3, ratios were low for 2815 and 5611 in Pilot 1, but higher in Pilot 3. This suggests that whereas some participants may show similar characteristics at different times, for others these may change. Thus it would make sense to **compare changes within sessions rather than between sessions**, and for **each participant separately**, rather than grouping them together, or at least separating out 'strong reactors' and 'weak reactors'.

5. Ratios (**ID**, **Mod**)

In Pilot 3, ratios were of a similar order for EA and TEAS each participant: for 5601 and 8680, ratios were high, for 8954 low, and for 2185 somewhere between. However, apart from 8954, patterns of changing ratios from EO1→segments→EO2 were dissimilar for the two interventions (EA and TEAS).

FURTHER ANALYSIS

Baseline comparisons: Test-retest reliability (TTR)

Significant correlations between values of HRV measures for the initial segments of all Visits were counted (SPSS bivariate correlations, with default SPSS Bootstrap settings, and Spearman's *rho* as a confirmatory non-parametric coefficient).

In Pilot 1, 4 measures showed significant TTR in the initial 'eyes-closed' segment (EC1), and 2 in the initial 'eyes-open' segment (EO1).

In Pilot 2, significant TTR for all 6 pairwise Visit combinations was found for four HRV measures, with least TTR for ApEn (2 comparisons significant) and SampEn (1 comparison significant), indicating that these nonlinear measures may be very sensitive to noise (in contrast, D_2 showed good TTR for 4 comparisons).

In Pilot 3, only 4 measures showed significant TTR, and then only for one comparison each. Only D_2 showed significant TTR when EO1 results for P1 and P3 were compared for those who participated in both Pilots.

Table 18 summarises these findings.

Table 18. Counts of significant test-retest reliability (TTR) for values of HRV measures in segments EO1 (Pilots 1-3) and EC1 (P1) – shown as P1(C).

HRV	V1-V2	V1-V3	V1-V4	V2-V3	V2-V4	V3-V4	**All**
RR	P2	P2	P2	P2 P3	P2	P2	7
SDNN	P1(C) P1 P2 P3	P2	P2	P2	P2	P2	9
RMSSD	P2	P2	P2	P2	P2	P2	6
HFpwr	P1(C) P2	P2	P2	P2	P2	P2	7
LF/HF	P1(C) P2		P2		P2		4
ApEn					P2	P2 P3	3
SampEn	P2						1
D_2	P1(C) P1 P2			P2	P2	P2 P3	7
All	14 (4 C 10 O)	4	5	6	7	8	**44**

This is illustrated in graphic form in **Fig 21**, showing total number of significant visit-to-visit correlations for each HRV measure (Pilots 1 and 3). Greater test-retest reliability (more correlations) suggests greater stability of the measure.

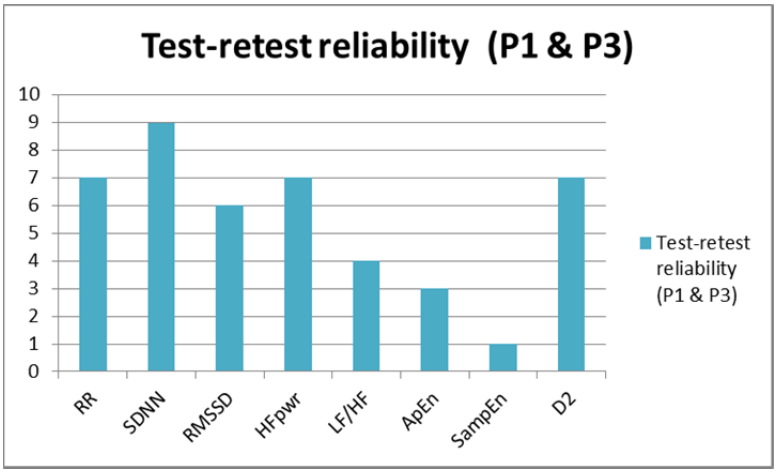


Fig 21. Total number of significant visit-to-visit correlations for each HRV measure in Pilots 1 and 3.

It is instructive to compare **Fig 21** with the nSD charts above (**Figs 6-8, 13**).

High or low baseline HRV measures as another factor in outcome

HRV measures at baseline (EO1) were compared to the median for the whole Pilot sample (all participants, in all segments), transformed into binary numbers (1 if > median, 0 if ≤ median), and relabelled as RR-ini, SDNN-ini, etc.. Median values for Pilot 1 were taken from the Pilot 1 data alone, but for each of Pilot 2 and 3 from their combined data. In addition to the usual 8 HRV measures, peak LF frequency was coded into 1s and 0s in the same way (as 'LFpk-ini'), and also relative to 0.1 (see Appendix A), as 'LFpk-0.1'.

In each Pilot, significant differences in HRV values during stimulation segments were found depending on initial state, using the Mann-Whitney U test (2-tailed asymptotic significance), and counted.

The initial values which differentiated between values in stimulation segments of over half the HRV measures were then tabulated (**Table 19**).

In addition, the correlation ratio *eta* (*η*) (see below, p. 40) were calculated. Those HRV measures for which *eta* (*η*) was highest or lowest were also tabulated.

Following analysis of Pilot 1, LFpk-0.1 was excluded from further analysis as having the lowest mean *eta* (*η*) across all measures (see Appendix F).

Table 19. Summarising significant differences in HRV values during stimulation segments resulting from high or low initial values.

Pilot	initial HRV with most effect	HRV most affected	HRV least affected	initial *eta* (*η*) largest	initial *eta* (*η*) least	stim *eta* (*η*) largest	stim *eta* (*η*) least
P1	RR, SDNN, RMS SD, HFpwr, D_2, LFpk	RR D_2,	ApEn SampEn	SDNN RR	ApEn LF/HF	RR SDNN	ApEn SampEn
P2	SDNN, RMS SD, HFpwr, ApEn, SampEn, D_2, LFpk	RMS SD HFpwr	RR LF/HF	SDNN RMS SD	RR LF/HF	RMS SD SDNN	ApEn RR
P3 EA	RR, SDNN, RMS SD, HFpwr, SampEn, D_2	RMS SD HFpwr	LF/HF ApEn	SampEn HFpwr	ApEn LFpk	D_2 RR	ApEn LF/HF
P3 TEAS	RR, HFpwr, D_2	SDNN RMS SD HFpwr	LF/HF ApEn	D_2 HFpwr	LFpk ApEn	D_2 RR	ApEn LF/HF

Those initial measures with most effect over all Pilots were **HFpwr** and **D_2** (4 occurrences), with **RR**, **SDNN** and **RMS SD** not far behind (3 occurrences). Those measures most affected by baseline values during stimulation segments were RMS SD and HFpwr (3 occurrences), those least affected being LF/HF and ApEn (3 occurrences).

No initial measure showed consistent largest *eta* more than twice in the above Table, but ApEn occurred 3 times as the initial measure with lowest *eta*. Of the measures during stimulation segments showing highest *eta*, only RR occurred 3 times in **Table 19**, and of those showing lowest *eta,* ApEn appeared 4 times.

On the basis of these results, the 5 measures in bold above (3 time domain, 1 frequency domain, 1 nonlinear) were selected for calculation of mean CV (normalised SD) and *eta* (*η*) (**Table 20**, and see below, pp. 41, 42).

Table 20. Number of significant differences in HRV measures (in stimulation segments) induced by the five most influential HRV measures at baseline, with corresponding correlation ratios *eta* (η).

HRV	*N* Significant differences					Correlation ratio *eta* (η)				
	P1	P2	P3 EA	P3 TEAS	**mean**	P1	P2	P3 EA	P3 TEAS	**mean**
RR	6	2	6	6	5	0.381	0.185	0.276	0.372	0.304
SDNN	6	5	5	4	5	0.417	0.463	0.379	0.269	0.382
RMS SD	6	5	5	4	5	0.354	0.435	0.339	0.290	0.355
HFpwr	5	6	6	6	5.75	0.280	0.396	0.407	0.385	0.367
D_2	5	6	6	6	5.75	0.292	0.365	0.388	0.412	0.364
mean	5.6	4.8	5.6	5.2	5.3	0.345	0.369	0.358	0.346	0.354

Table 21. Mean *eta* (η) for each HRV measure in stimulation segments, for the five selected baseline measures taken together.

Pilot	RR	SDNN	RMS SD	HFpwr	LF/HF	ApEn	SampEn	D_2	**mean**
Pilot 1	0.531	0.424	0.420	0.351	0.159	0.106	0.197	0.570	0.345
Pilot 2	0.208	0.558	0.563	0.439	0.190	0.105	0.312	0.574	0.369
Pilot 3 EA	0.566	0.383	0.357	0.249	0.091	0.105	0.444	0.662	0.358
Pilot 3 TEAS	0.406	0.375	0.415	0.260	0.190	0.132	0.367	0.621	0.346
mean	0.428	0.435	0.439	0.325	0.1575	0.112	0.330	0.607	0.355
CV of η	0.378	0.195	0.200	0.274	0.296	0.119	0.315	0.072	0.032

*Association of ID and Visit with baseline values (**B**)*

A *Chi*-square (χ^2) test for Values during stimulation segments was conducted to assess the association between initial state and ID or Visit (**Table 22**).

Table 22. Results of the *Chi*-square (χ^2) test for Values during stimulation segments, showing mean Pearson's χ^2) in each Pilot, and whether the test results were significant.

Pilots 1-3	ID			Visit		
	P1	P2	P3	P1	P2	P3
RR-ini	0.001	<0.001	<0.001	ns	ns	0.008
SDNN-ini	<0.001	<0.001	<0.001	<0.001	0.001	<0.001
RMSSD-ini	<0.001	<0.001	<0.001	<0.001	0.004	<0.001
HFpwr-ini	<0.001	<0.001	<0.001	ns	ns	0.008
LF/HF-ini	<0.001	<0.001	<0.001	ns	<0.001	<0.001
ApEn-ini	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
SampEn-ini	<0.001	<0.001	<0.001	0.043	0.049	<0.001
D_2-ini	<0.001	<0.001	<0.001	ns	<0.001	<0.001
LFpk-ini	<0.001	<0.001	<0.001	0.013	ns	0.004
mean	49.417	235.114	74.364	8.846	19.448	33.631

Note that the value of χ^2 depends on the degrees of freedom, so is bound to be lower for Visit (df 1 or 3) than ID (df 6, 11 or 3).

In Pilots 1 and 2, initial values are closely associated with ID, but much less so with Visit.

In Pilot 3, unlike P1 and P2, not only are initial values closely associated with ID, but also with Visit (although not to the same extent).

The association for ID was further explored for the individual participants. Obviously some participants showed more high or more low initial HRV measures; sometimes the proportion of these was significant. However, proportions (and significance) were not consistent across visits, as shown in **Table 23**.

Table 23. Participants in Pilots 1 to 3 showing significant differences in numbers of baseline 'high' and low' values for the 9 HRV measures tabulated in Table 22.

Pilot	ID	High/low initial HRV	With LF/HF reversed	Binomial significance
Pilot 1	2185	3/15	3/15	0.008
	8875	15/3	13/5	0.008 (ns)
Pilot 2	4290	25/11	27/9	0.029 (0.004)
	5044	11/25	15/21	0.029 (ns)
	5611	12/24	10/26	ns (0.011)
	6899	23/13	25/11	ns (0.029)
	7815	8/28	4/32	0.001 (<0.001)
	7904	24/12	28/8	ns (0.001)
Pilot 3	8680	13/23	9/27	ns (0.004)
	8954	28/8	24/12	0.001 (ns)

The high/low proportions were also not consistent over longer periods: although 2185, for instance, showed significantly more high than low initial measures in Pilot 1, in Pilot 3 this was no longer the case.

When LFpk-0.1 (rather than LFpk-ini) was considered in isolation, then participant 8680 showed high initial LFpk (>0.1) in one session of Pilot 1, and in all four sessions in Pilot 3. Two participants in Pilot 2 (4290 and 6899)showed high initial LFpk-0.1 in 3 out of 4 sessions. Others only showed high initial LFpk-0.1 in 2 sessions at the most. 11 out of the total 19 participants in the three Pilots demonstrated high initial LFpk-0.1 in one or more sessions (25 out of a total of 76 sessions, or approximately 1/3).

Further data on differences in HRV measures with baseline state can be found in Appendix F.

Coefficients of variance (normalised SD) and effect size

5. Ratios – Coefficients of variance (normalised SD)

Table 24. Coefficients of variance (normalised SD) for H/L ratios.

Comparison	Pilot	EO1	Segments	EO2	mean CV (SD)
Hz	P1 P2 P3 EA P3 TEAS	0.235 0.030 0.089 0.177	0.149 0.055 0.571 0	0.060 0.206 0.176 0.045	0.149 (0.153)
Visit	P1 P2 P3 EA P3 TEAS	0.339 0.092 0.334 0.229	0.113 0.168 0.637 0.513	0.230 0.190 0.261 0.414	0.293 (0.163)
ID	P1 P2 P3 EA P3 TEAS	1.254 1.928 0.608 0.645	1.159 1.081 0.841 0.789	1.172 1.074 0.711 1.097	1.030 (0.360)
Loc	P1 P2 P3 EA P3 TEAS	n/a 0.167 n/a n/a	0.143 (0.161[a]) 0.125 0.306 0.171	n/a 0.243 n/a n/a	0.193 (0.069) [0.196 (0.067)[a]]
Amplitude	P1 P2 P3 EA P3 TEAS	n/a n/a n/a n/a	1.095 1.098 0.540 0.569	n/a n/a n/a n/a	0.826 (0.313)
Dur	P2 only	n/a	1.165 (0.037)	n/a	1.165 (0.037)
Baseline	P1 P2 P3 EA P3 TEAS	n/a n/a n/a n/a		not calculated	

a. Including Bilat and LLSS

Time did not permit calculation of CV for Baseline state.

As expected, CV is highest for ID, followed by Visit and then Loc. CV is least for Hz. Note the strong effect of stimulation amplitude.

5. Ratios – Effect size using modified Cohen's *d*

One measure of effect size, most commonly used for independent samples (as when comparing a treatment and a non-treatment group), is Cohen's *d* [Anon (Wikipedia); Cohen 1992; Taş-Cebe & Cummings 2013; Thalheimer & Cook 2002]. A *d* of <0.015 is considered negligible, around 0.20 (0.15-0.4, or 0.2 to 0.3) small, around 0.5 (0.4-0.75) medium, and around 0.80 large (or >075, sometimes subdivided into large, 0.75-1.1, and very large, 1.1-1.45). Cohen's *d* is most meaningful when calculated *after* rejecting the null hypothesis in a statistical test [Anon 2010-2012]. However, it is still a useful indicator of the magnitude of mean differences where the truth or otherwise of a null hypothesis cannot be established. Where such quantification is problematic, "Cohen's effect size criteria may serve as a last resort" (Ellis 2010).

The equation for Cohen's *d* for two groups (1 and 2, with means m_1 and m_2, standard deviations sd_1 and sd_2, and numbers n_1 and n_2) is:

$$d = \frac{(m_1 - m_2)}{(\text{pooled } sd)}$$

Where pooled $sd = \frac{\sqrt{(n_1-1)sd_1^2 + (n_2-1)sd_2^2}}{\sqrt{(n_1 + n_2 - 2)}}$

In some discussions of Cohen's *d* the '– 2' is included in the denominator (Hartung et al. 2008); in others it is omitted (Thalheimer & Cook 2002). It is included here to give a more stringent assessment of effect size. Whether it is included or not was found to have a negligible effect on whether *d* is classified as small, medium or large.

Cohen's *d* is used in this analysis on the assumption that the groups compared are in effect independent (because of the number of independent variables considered). However, given that each Pilot is structured as a complex cross-over, with the same participants usually included at least twice in each comparison, *n* was determined from the number of 'cases' in the SPSS Descriptives output, not from the number of participants.

When comparisons were between more than two groups (e.g., comparing the effects of participant ID or stimulation Loc on outcome), the mean of Cohen's *d* for the various inter-group comparisons is presented.

When calculating Cohen's *d* for summed high/low ratios, some individual ratios will have zero as the denominator (if all HRV measures in the comparison concerned are 'high'), giving infinite values. However, in these Pilots, there are only 13 such ratios, three of which do not contribute to Cohen's *d* for stimulation segments (**Table 25**).

Table 25. 8/0 ratios [those highlighted do not contribute to calculated Cohen's *d*]

P1	P2	P3 EA	P3 TEAS
V1 T 10 8311	V2 T 2.5 EA1 4290	n/a	V1 L 2.5 TEAS1 8680
V1 EO2 10 8311	V2 T 2.5 EA4 4290		
V2 L 10 8875	V2 T 2.5 MA2 4290		
V2 Bilat 10 8875	V2 B 2.5 EO1 7815		
	V3 L 10 MA2 7815		
	V4 T 2.5 EO1 7815		
	V4 T 2.5 EA4 7815		
	V4 T 2.5 MA2 7815		

By replacing 8/0 with 7/1 for the remaining 10 ratios, approximate values for Cohen's *d* were then calculated (**Table 26**).

Table 26. Cohen's *d* (calculated using Descriptives exported from SPSS into Excel). Where more than one inter-group comparison is possible, max and min *d* were calculated, and then their mean (median).

Comparison	Pilot	EO1	Segments	EO2	mean *d* (SD)
Hz	P1 P2 P3 EA P3 TEAS	0.272 0.056 0.414 0.0475	0.074 0.179 0.069 0.349	0.125 0.157 0.348 0.508	0.217 (0.157)
Visit (max, min)	P1 P2 P3 EA P3 TEAS	0.358 0.129 (0.217, 0.040) 0.600 (1.054, 0.145) 0.461 (0.799, 0.122)	0.164 0.061 (0.115, 0.006) 0.520 (1.002, 0.037) 0.823 (1.253, 0.393)	0.149 0.284 (0.490, 0.077) 0.455 (0.766, 0.144) 0.508 (0.910, 0.105)	0.376 (0.227)
ID (max, min)	P1 P2 P3 EA P3 TEAS	25.103 (50.205, 0) 0.332 (0.612, 0.051) 0.658 (1.259, 0.056) 1.095 (1.536, 0.653)	1.332 (2.334, 0.329) 1.712 (3.249, 0.175) 1.122 (2.034, 0.209) 1.356 (2.323, 0.389)	13.218[b] ('infinity', 0) 2.034 (4.068, 0) 0.439 (0.694, 0.183) 1.466 (2.116, 0.815)	4.156 (7.470)
Loc (max, min)	P1 P2 P3 EA P3 TEAS	n/a 0.152 (0.247, 0.056) n/a n/a	0.313 (0.465, 0.160) 0.096 (0.180, 0.011) 0.401 (0.742, 0.059) 0.233 (0.384, 0.082)	n/a 0.153 (0.268, 0.037) n/a n/a	0.225 (0.115)
Amplitude	P1 P2 P3 EA P3 TEAS	n/a n/a n/a n/a	1.006 0.085 0.999 0.158	n/a n/a n/a n/a	0.562 (0.510)
Dur	P2 only	n/a	113 (0.118)	n/a	113 (0.118)

b. Interpolated as mean of EO1 and Segment entries.

As for CV, Cohen's *d* is highest for ID, followed by Amp, Visit and then Loc. It is least for Hz. Time did not permit calculation of Cohen's *d* for Baseline state.

5. Ratios – Comparing CV and Cohen's *d*

Both CV and Cohen's *d* are normally distributed, and with acceptable skewness and kurtosis (albeit only just acceptable skewness for Cohen's *d*). Comparison between them shows good correlation for the factors Hz, V, ID and Loc (**Fig 22**).

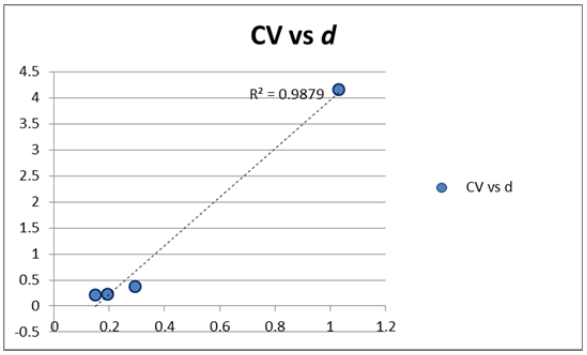


Fig 22. Correlation between CV and *d* for the factors Hz, V, ID and Loc.

However, correlation is less good if Amp is included as an additional factor (**Fig 23** *Left*), and poor if the results for Dur from Pilot 2 are also included (**Fig 23** *Right*).

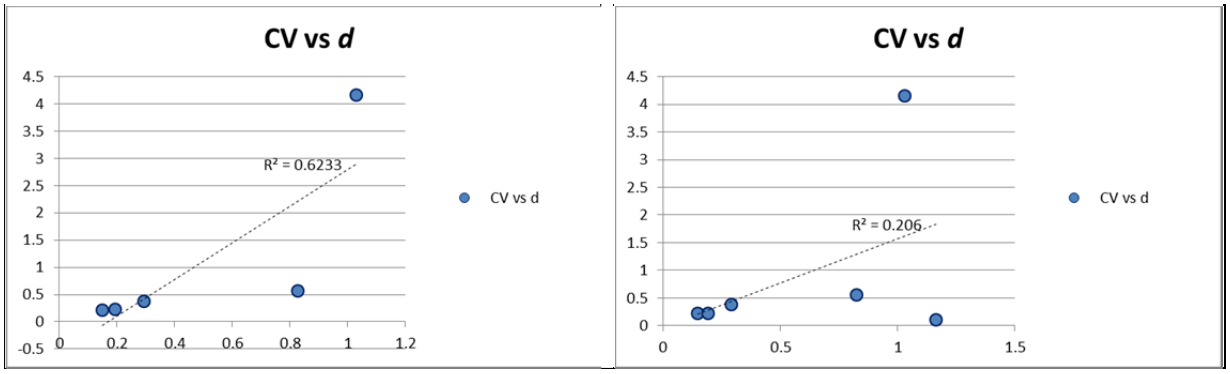


Fig 23. *Left*: Correlation between CV and *d* for the factors Hz, V, ID, Loc and Amp.
Right: Correlation for the same factors, and also Dur.

Diff% for the various methods of analysis (Values, Correlations, BER and H/L ratios) used in this study

In **Table 27**, Mean [and Max] Diff%s are entered for all comparisons when these are between more than two factors (e.g., for Visit in Pilot 2 or Pilot 3, or for ID or Loc in all Pilots). Other than for comparisons by frequency and pre-to-post values, data was sorted to ensure that all these Diff%'s were positive.

Table 27. Diff% for the various methods of analysis (Values, Correlations, BER and H/L ratios) used in this study.

	value	**correls**	**pre-to-post value**	**BER**	**H/L ratio**	**H/L ratio**	**H/L ratio**
Hz					segments	start	end
P1	-6.314	-23.53	-72	-21.1	-33.249	-28.460	8.865
P2	1.400	18.75	-72	6.7	-9.484	-11.876	-34.154
P3 EA	-16.365	0.00	-382	28.9	33.300	-11.800	-24.636
P3 TEAS	-21.844	57.14	144	-14.4	0.000	-22.240	-6.170
Visit					segments	start	end
P1	3.513	26.67	5603.196	6.154	38.806	96.000	22.358
P2	17.047 [30.489]	21.429 [45.455]	1602.068 [3795.295]	11.328 [20.295]	23.932 [48.860]	11.981 [23.188]	25.980 [51.948]
P3 EA	12.136 [24.350]	170.833 [466.667]	621.701 [1246.671]	14.524 [27.228]	118.898 [294.737]	68.013 [145.455]	41.774 [87.912]
P3 TEAS	15.901 [31.467]	39.206 [80.000]	17.110 [28.275]	35.630 [65.918]	108.546 [240.316]	35.515 [65.641]	90.861 [189.630]
ID					segments	start	end
P1	50.758 [134.723]	135.331 [566.667]	705.997 [10948.01]	61.683	355.559 [1320.000]	443.636 [2933.333]	1078.942 [4800.000]
P2	58.426 [263.909]	101.760 [600.000]	-480.847 [3187.908]	26.617 [74.884]	246.731 [1774.040]	388.496 [3200.000]	259.786 [2240.000]
P3 EA	24.703 [48.452]	85.000 [175.000]	529.694 [1672.184]	53.050 [27.770]	395.181 [880.000]	284.318 [727.273]	422.657 [1088.000]
P3 TEAS	46.979 [101.666]	66.667 [133.333]	36.244 [57.375]	20.659 [40.200]	259.557 [673.684]	358.254 [853.333]	497.576 [1520.000]

Loc					segments	start	end
P1	3.927 [9.266]	6.250 [12.500] {15.913 [42.857]}[a]	n/a	10.406 [21.626]	6.107 [13.674]	n/a	n/a
P2	10.149 [19.156]	11.197 [23.077]	72.304 [-5495.08]	10.626 [18.505]	17.185 [28.448]	16.829 [28.858]	36.237 [58.537]
P3 EA	6.599 [11.136]	66.667 [150.000]	n/a	n/a	48.570 [100.571]	n/a	n/a
P3 TEAS	21.076 [40.470]]	92.857 [175.000]	n/a	n/a	22.939 [45.878]	n/a	n/a

a. In {}s: including Bilat and LLSS.

Taking the mean (rather than max) absolute values, averaging over all columns, Diff%s are ranked as follows:

P1 Loc << Hz << ID < V
P2 Hz < Loc << ID < V
P3 EA Loc < Hz << V < ID
P3 TEAS Hz < Loc < V << ID.

Out of a possible 28 individual columns, in 22 ID ranks highest, and in 18 Hz ranks lowest. Thus, as with CV and Cohen's *d*, of the four factors included in this Table, ID appears to have the most effect on HRV outcomes, and Hz the least.

Partial correlations between Cohen's *d*, CV and Diff% for the various methods of analysis used

Partialling out the effects of Pilot and the various factor Comparisons, **Table 28** shows the correlations which remain significant (for Segments only) when calculated using Bootstrap (with SPSS default settings).

Table 28. Significant partial correlations between Cohen's *d*, CV and Diff% for the various methods of analysis used.

all	Cohen's *d*	Diff% values (1)	Diff% correlations (2)	Diff% BER (4)	Diff% H/L ratios (5)
CV	**	**	*	**	**
Cohen's *d*		**	*	**	**
Diff% values (1)					**
Diff% correls (2)					*
Diff% BER (4)					**

** p<0.01; * p<0.05.

R^2 values for the correlations with Diff% are shown in **Table 29**. For the count of significant differences, see Table 1; for *eta* (η), see next section, p. 40.

Table 29. Correlation coefficients for some Diff% results with other methods of assessing effect.

all	Diff% values (1)	Diff% correlations (2)	Diff% BER (4)	Diff% H/L ratios (5)	**mean** (excl BER)	**nSD**
CV	0.984	0.841	0.001	0.927	0.917	0.078
Cohen's *d*	0.999	0.755	0.017	0.944	0.899	0.142
eta (η)	0.960	0.878	0.001	0.871	0.903	0.055
N signif diffs	0.994	0.690	0.021	0.977	0.887	0.193
mean	0.984	0.791	0.010	0.930	**0.902**	
nSD	0.018	0.107	1.052	0.048	**0.014**	

These findings support the use of three of the methods of analysis used in this study, in particular the Values and H/L ratios. However, BER is clearly assessing something rather different. Note that dispersion of R^2 was low for CV and *eta* (η), and higher for Cohen's *d* and the count of significant differences.

Example: Diff% and nSD (CV) for Pilot 2 (segments EA1 to EA4)

If the absolute (non-signed) differences in value of the various measures are normalised and taken as percentages (Diff%), they correlate very closely with coefficient of variance (CV), i.e. the normalised standard deviation nSD, of the measures themselves for the same sample (Pilot 2, EA1 to EA2), as shown in **Fig 24**.

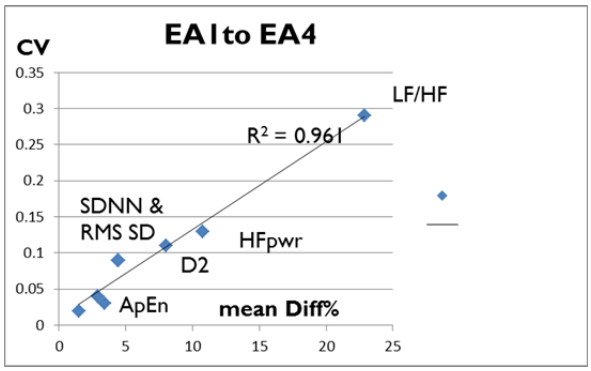


Fig 24. Scatter plot of CV vs mean Diff% for the 8 HRV measures in Pilot 2, segments EA1 to EA4, showing how they are closely correlated.

Correlation ratio *eta* (η) for factors in this study

To confirm that the effect of Amp (Amp-N) on the HRV appears to be far stronger than that of Loc (above, Values (**Amp**), p. 19), a comparison was made of the correlation ratio *eta* (η) for the various factors (**Table 30**).

Table 30. Correlation ratio *eta* (η) for stimulation amplitude with some nominal independent variables (significant values of η in **bold**).

Pilot	Hz	Visit	Time	Loc	B & T only	ID
P1	0.265	0.006	0.164	0.395	0.526	**0.789**
P2	0.099	0.179	0.213	0.396	0.403	**0.771**
P3 EA	0.252	0.296	0.164	0.366	0.520	0.615
P3 TEAS	0.005	0.296	0.167	0.363	0.418	0.583

In Pilot 1, for baseline values of the different HRV measures, *eta* (η) for stimulation amplitude was >0.4 only for HFpwr-ini (0.410), RR=ini (0.426) and D_2-ini (0.697). In Pilots 2 and 3, *eta* (η) did not reach 0.4 for any HRV measure. Thus there does not appear to be some 'HRV type' of participant with a particularly low or high tolerance for electrical stimulation.

This suggests that Amp is *not* a major confounding factor for any results other than ID.

Table 31 shows the mean *eta* (η) for the different factors (during stimulation segments only), with correlations for Amp values with HRV measures.

Table 31. Mean *eta* (η) for the different factors (during stimulation segments only), with correlations for Amp values with HRV measures (p values shown as ** (*) for Pearson <0.01 (Spearman <0.05)). *Eta* (η) values in **bold** if > 0.75.

Pilot	**Factor**	RR	SDNN	RMS SD	HFpwr	LF/HF	ApEn	SampEn	D_2	**mean**
P1	Hz	0.068	0.208	0.267	0.082	0.032	0.155	0.316	0.049	0.147
	Loc	0.033	0.077	0.150	0.217	0.128	0.123	0.234	0.094	0.132
	Amp	** (**)	n (**)	n (*)	n (**)	n (n)	n (n)	n (n)	** (**)	2 (5)
	Amp-N	0.479	0.192	0.081	0.139	0.069	0.067	0.678	0.102	0.226
	ID	0.459	0.707	0.594	0.535	0.496	0.433	0.303	**0.830**	0.545
	V	0.132	0.209	0.196	0.030	0.071	0.164	0.405	0.057	0.158
	Time	0.424	0.414	0.297	0.264	0.338	0.258	0.277	0.586	0.357
	Segment	0.076	0.191	0.202	0.205	0.186	0.272	0.220	0.077	0.179
	Baseline	0.531	0.424	0.420	0.351	0.159	0.106	0.197	0.570	0.345
P2	Hz	0.064	0.011	0.054	0.050	0.004	0.141	0.009	0.053	0.048
	Loc	0.177	0.157	0.166	0.077	0.124	0.080	0.097	0.123	0.125
	Amp	** (*)	n (**)	* (*)	** (n)	** (n)	n (**)	n (n)	n (*)	4 (4)
	Amp-N	0.127	0.227	0.100	0.079	0.257	0.144	0.168	0.182	0.161
	ID	**0.890**	**0.871**	**0.863**	**0.876**	**0.766**	0.536	0.738	**0.842**	0.798
	V	0.184	0.180	0.155	0.118	0.212	0.077	0.095	0.146	0.146
	Time	0.408	0.400	0.565	0.710	0.283	0.124	0.181	0.494	0.396
	Dur/Segm	0.025	0.034	0.018	0.024	0.044	0.100	0.099	0.065	0.051
	Baseline	0.208	0.558	0.563	0.439	0.190	0.105	0.312	0.574	0.369
P3 EA	Hz	0.032	0.210	0.196	0.220	0.146	0.413	0.289	0.039	0.193
	Loc	0.215	0.306	0.303	0.325	0.480	0.336	0.200	0.195	0.295
	Amp	* (**)	n (**)	* (*)	n (**)	* (**)	n (n)	n (n)	** (**)	4 (6)
	Amp-N	0.548	0.329	0.363	0.262	0.172	0.217	0.373	0.406	0.334

	ID	**0.790**	0.498	0.551	0.393	0.307	0.328	**0.818**	0.726	0.551
	V	0.421	0.383	0.378	0.322	0.342	0.391	0.178	0.368	0.348
	Time	0.360	0.375	0.318	0.213	0.075	0.066	0.406	0.608	0.303
	Segment	0.096	0.223	0.197	0.256	0.178	0.304	0.225	0.052	0.191
	Baseline	0.566	0.383	0.357	0.249	0.091	0.105	0.444	0.662	0.358
P3 TEAS	Hz	0.137	0.248	0.215	0.207	0.177	0.122	0.014	0.023	0.143
	Loc	0.296	0.218	0.154	0.268	0.323	0.308	0.396	0.190	0.269
	Amp	* (*)	n (n)	n (n)	n (n)	n (n)	n (n)	* (n)	n (n)	2 (1)
	Amp-N	0.330	0.264	0.213	0.254	0.150	0.156	0.444	0.102	0.239
	ID	**0.823**	0.561	0.616	0.481	0.302	0.254	**0.758**	0.672	0.558
	V	0.291	0.382	0.355	0.409	0.312	0.242	0.043	0.349	0.298
	Time	0.225	0.318	0.331	0.231	0.113	0.146	0.415	0.499	0.285
	Segment	0.124	0.359	0.388	0.323	0.277	0.303	0.231	0.282	0.286
	Baseline	0.406	0.375	0.415	0.260	0.190	0.132	0.367	0.621	0.346
P3	Mod	0.069	0.083	0.022	0.074	0.101	0.084	0.101	0.060	0.074
	Segment	0.128	0.271	0.274	0.260	0.290	0.302	0.247	0.214	0.248
mean		0.298	0.313	0.304	0.271	0.217	0.209	0.302	0.321	
CV		0.790	0.581	0.628	0.683	0.707	0.585	0.681	0.824	

As in the other analyses conducted, ID is the factor showing greatest effect.

Measures with the highest mean *eta* (η) are SDNN and SampEn; those with the lowest mean *eta* (η) are LF/HF and ApEn.

Table 32 shows the mean *eta* (η) for the various comparisons, combined for the different Pilots.

Table 32. Mean *eta* (η) for the various comparisons, combined for the different Pilots.

Factor	Mean (SD)	CV of η
Hz	0.133 (0.061)	0.458
Loc	0.205 (0.089)	0.435
Amp-N	0.240 (0.071)	0.298
ID	0.613 (0.123)	0.201
V	0.237 (0.101)	0.425
Time	0.335 (0.051)	0.152
Segment	0.177 (0.096)	0.546
Baseline	0.355 (0.011)	0.032

Interaction between factors (independent variables) – the χ^2 test

To further explore the role of Amp-N (Amp transformed into binary numbers), *Chi*-square (χ^2) tests for various categorical variables were performed (**Table 33**).

Table 33. Results of the *Chi*-square (χ^2) test for Values during stimulation (EA or TEAS) segments.

χ^2	Amp-N vs ID	Time vs ID	Amp-N vs Loc	Amp-N vs Hz
P1	37.455 (p<0.001, df 5)	148.800 (p<0.001, df 15)	4.303 (ns, df 5)	5.719 (p=0.017, df 1)
P2	66.510 (p<0.001, df 11)	365.929 (p<0.001, df 22)	41.244 (p<0.001, df 3)	14.139 (p<0.001, df 1)
P3 EA	20.825 (p<0.001, df 3)	32.000 (p<0.001, df 3)	0.508 (ns, df 3)	2.032 (ns, df 1)
P3 TEAS	12.698 (p=0.005, df 3)	32.000 (p<0.001, df 3)	4.571 (ns, df 3)	0.000 (ns, df 1)

Thus, as originally planned when allocating visit times to participants, Time and ID are closely associated. Also closely associated are ID and stimulation amplitude – some participants prefer, or can tolerate, stronger stimulation.

Interestingly, the association between Amplitude and Loc is only significant in Pilot 2. Nonetheless, in all Pilots mean amplitude was greatest at B ($ST36^2$) and least at T ($LI4^2$), ranked order in Pilots 2 and 3 being: B>L>R>T. In Pilot 1, order was B>R>L>T (although, when results were split by Hz, at 10 Hz Amp was higher in Pilot 3 TEAS at R than B). In Pilots 1 and 2, the difference between Amp for B and the next greatest Amp was more for 2.5 Hz than 10 Hz (this was not the case in Pilot 3). In this context it is important to remember that calculation of *eta* (η) shows that Amp is only a significant confounder for ID, not for Loc, Time, Visit or Hz (even when only the extremes of T and B are considered for Loc).

The interaction of Amp and Hz should also be noted as a possible confounding factor. In Pilot 1, mean amplitude for 10 Hz was greater than for 2.5 Hz at all locations. In contrast, in Pilot 2 mean amplitude was greater for 2.5 Hz than 10 Hz at both $LI4^2$ (T) and $ST36^2$ (B), although these differences were not significant.

In no Pilot did the *Chi*-square (χ^2) test indicate a significant association between Hz and Loc.

CONCLUSIONS

Few differences are significant for any of the comparisons undertaken, and few are consistent across all Pilots. In particular, it is not possible from these data to state unequivocally that the effect of one frequency is greater than that of the other, or that one is more likely than the other to benefit health.

However, comparing CV and Cohen's *d* for H/L ratios with the *eta* (η) values obtained above, and bearing in mind the numbers of significant differences for the various comparisons (**Table 1**), as well as the correlations with Diff% (**Tables 28, 29**, pp. 39, 40), it becomes evident that these five methods of assessing how much an experimental factor in these Pilots affects outcome are broadly in agreement (**Table 34**).

Table 34. Five methods of assessing how much an experimental factor in these Pilots affects outcome.

Factor	*N* significant differences	mean CV (SD)	Effect size *d* (SD)	mean *eta* (η) (SD)	mean Diff% Values (SD)
Hz	5	0.149 (0.153)	0.217 (0.157)	0.133 (0.061)	11.481 (9.302)
Loc	2	0.193 (0.069)	0.225 (0.115)	0.205 (0.089)	10.438 (7.536)
Visit	4	0.293 (0.163)	0.376 (0.227)	0.237 (0.101)	12.150 (6.128)
Amp	17	0.826 (0.313)	0.562 (0.510)	0.240 (0.071)	n/a
Dur	0	1.165 (0.037)	0.113 (0.118)	0.051 (0.033)	n/a
Baseline	21	n/a	n/a	0.355 (0.011)	n/a
ID	28	1.030 (0.360)	4.156 (7.470)	0.613 (0.123)	45.217 (14.481)
mean nSD	0.990	0.450	0.931	0.356	0.589

n/a: not yet calculated.

For example, there is excellent correlation between CV and Cohen's *d* (R^2=0.988) when factors Amp and Dur are excluded. This decreases to 0.623 if Amp is included, and 0.206 if Duration is included as well (**Table 35**). R^2 for other comparisons are shown below (for those with Diff%, see above, **Table 29**).

Table 35. Correlations between CV, *d* and *eta* (η).

Comparison	CV vs *eta*	CV vs *d*	*d* vs *eta*
R^2	0.991 (0.667; 0.071)	0.988 (0.623; 0.206)	0.970 (0.968; 0.909)

Note that, although in all three comparisons, R^2 decreases if Amp and Dur are included, this decrease is only marginal (to 0.909) for the correlation between Cohen's *d* and η. Given that η^2 is, like Cohen's *d*, also a measure of effect size, this is not surprising.

Table 36. Correlations of *N* (number of significant differences) with CV, *d* and *eta* (η).

Comparison	*N* vs CV	*N* vs *d*	*N* vs *eta*
R^2	0.965 (0.940; 0.202)	0.989 (0.781; 0.780)	0.848 (0.740; 0.792)

Correlations between *N*, which is quite a gross estimate of effect, with the more formal effect size estimates (Cohen's *d* and *eta*) are surprisingly good (**Table 36**). (*N* vs *eta* is 0.792, for example, when Baseline is included, and 0.803 with Baseline excluded.)

There is less dispersion of CV and Cohen's *d* than the other approach used to assess the impact of protocol factors on HRV measures, but they all indicate that there is considerable variation with participant (**ID**) and baseline (**B**), and little with stimulation frequency (**Hz**). Diff%, for example, is much less for Loc and Hz than for ID and Visit in all Pilots, with the rank order (for all Pilots considered together) being ID, Visit, Hz, Loc (although this varies in individual Pilots).

Thus the effects of the ID, Baseline, Amp, V, Loc and possibly Dur may mask those of Hz.

In theory, in a formal multivariate analysis where all factors are identified (and there is no additional error factor), the sum of η^2 for these is 1 (whether within or between subjects). Here it is 0.678,

suggesting that over 2/3 of the factors responsible for variance in outcomes have been accounted for.

Limitations

It is difficult to draw convincing conclusions from these three Pilot studies, for various reasons.

1. In Pilot 1, monitoring occurred *after* each 5-minute segment of stimulation, whereas in Pilots 2 and 3, it was carried out *during* stimulation.

2. In Pilot 1 and the first two sessions in Pilot 2, ECG/HRV was used, but thereafter PPG/PRV.

3. Too many variables were involved, particularly in Pilots 1 and 3, where sample size was small.

4. Furthermore, varying the order of interventions only added unnecessary complexity, especially in Pilot 3, where there might be an order effect for the three interventions used (MA, EA and TEAS). Here RR tended to increase marginally more during the whole session with 2.5 Hz than 10 Hz, and this was more marked when EA preceded TEAS (suggesting perhaps that ending a session with TEAS would be more relaxing than ending it with EA). In contrast, at 10 Hz, TEAS first could be interpreted as more relaxing.

5. Other interactions between the various factors compounded this problem – for example, between Amp and Hz, or Amp and Loc, or Amp and ID.

6. Differences at baseline were sometimes considerable, with individual participant tendencies at baseline tending to be repeated throughout subsequent segments, regardless of stimulation.

More specifically, the value of all measures at baseline tended to be maintained throughout subsequent segments (in 40 out of 64 segments, or 62.5%, in Pilot 1; in 42 segments, or 75%, in Pilot 2, and in 67 segments, or 83.75%, in Pilot 3) (cf. **Fig 8**, above). This was least often the case for LF/HF (occurring in only 10 out of a possible 24 segments) and D_2 (13 segments), most often for HFpwr (difference maintained in all 24 segments in the three Pilots) and RR (22 segments). The only measure which, in all three Pilots, was higher at baseline for 10 Hz than for 2.5 Hz, was LF/HF.

7. in the context of this study, Cohen's *d* needs to be interpreted with caution [Clark n.d.]. Results should not be taken to indicate that the factors considered contribute to the overall variance of HRV outcomes in a completely precise way. They are only a guide to ranking their effect. More formal multivariate analysis under expert guidance from a statistician would be required to take this further.

Future directions

Any further investigations should be designed carefully in order to obviate or compensate for these problems.

For example, small-scale Pilots should focus on individual participants, within individual sessions, and with stimulation at a single location within each session, rather than attempting to compare the effects of several variables at once. Careful attention should be paid to the effects of stimulation Amp, as well as Hz. One factor that has not been addressed so far is whether monitoring was

conducted during or after stimulation. Thus a possible next step would be to design a Pilot in which different frequencies of stimulation are used within each session, at the same locations, monitoring for HRV both during and after each stimulation segment. Another factor which has not been investigated is whether prior experience of EA or TEAS affects outcome. The data from Pilots 1 and 2 should be explored for this possibility before proceeding to recruit for further investigations.

The basic analytical methods employed so far are accessible even to those with little statistical expertise. They offer a simple way of assessing the contribution of different experimental factors to outcomes when statistical significance is elusive [Taş-Cebe & Cummings 2013] and sample size is small. They would thus be very appropriate in acupuncture research, which tends to involves quite a number of independent variables in small-scale studies. In addition, they provide useful indicators for further analysis. However, in principle – and using Bootstrap or with given a sufficiently large sample size – this should make use of more advanced methods such as mixed models and multivariate methodology in order to properly assess the interactions of the experimental factors (e.g. Hz and Loc) and their relative contribution to changes in HRV.

ACKNOWLEDGEMENTS

To Robert Kozarski and Tim Watson of the University of Hertfordshire for advice on statistics and study design, to Terje Alraek and Alan Watkins for sharing their considerable experience of HRV, to Harmony Medical for providing equipment (the Classic4 stimulator, needles and self-adhering electrodes), and of course to the acupuncturists and other health practitioners who gave up considerable amounts of their time to take part in our EEG/HRV study. Also to ourselves, for remaining patient with each other over a long period of collaboration. Any errors in this report remain, of course, our own responsibility.

REFERENCES

Acharya UR, Joseph KP, Kannathal N, Lim CM, Suri JS. Heart rate variability: a review. Medical and Biological Engineering and Computing. 2006; 44(12): 1031–51

Agelink MW, Sanner D, Eich H, Pach J, Bertling R, Lemmer W, Klieser E, Lehmann E. [Does acupuncture influence the cardiac autonomic nervous system in patients with minor depression or anxiety disorders?]. Fortschritte der Neurologie-Psychiatrie. 2003; 71(3): 141-9

Ahamed VI, Karthick NG, Joseph PK. Effect of mobile phone radiation on heart rate variability. Computers in Biology and Medicine. 2008; 38(6): 709-12

Akar JG, Everett TH, Kok LC, Moorman JR, Haines DE. Loss of spatiotemporal organization with the transition from acute to chronic atrial fibrillation in vivo. Journal of the American College of Cardiology. 2001; 37(2): A87-A141 (1269-132)

Alraek T, Tan CO. Acupuncture and heart rate variability. Acupuncture in Medicine. 2011; 29(1): 7-8

Anon. n.d. Basic concepts: reliability. Medical Research Council, UK. http://dapa-toolkit.mrc.ac.uk/basic-concepts/reliability.html [accessed 25.03.14]

Anon (Wikipedia). n.d. Effect size. http://en.wikipedia.org/wiki/Effect_size [accessed 26.03.14]

Anon. 2010-2012. Effect Size for Dependent Samples t-Test. http://www.statisticslectures.com/topics/effectsizedependentsamplest/ [accessed 26.03.14]

Appelhans BM, Luecken LJ. Heart rate variability as an index of regulated emotional responding. Review of General Psychology. 2006; 10(3): 229-40

Arai YC, Ushida T, Matsubara T, Shimo K, Ito H, Sato Y, Wakao Y, Komatsu T. The influence of acupressure at Extra 1 acupuncture point on the spectral entropy of the EEG and the LF/HF ratio of heart rate variability. Evidence-Based Complementary and Alternative Medicine. 2011; Article ID 503698. doi:10.1093/ecam/nen061

Ashare RL, Sinha R, Lampert R, Weinberger AH, Anderson GM, Lavery ME, Yanagisawa K, McKee SA. Blunted vagal reactivity predicts stress-precipitated tobacco smoking. Psychopharmacology. 2012; 220(2): 259-68

Bäcker M, Grossman P, Schneider J, Michalsen A, Knoblauch N, Tan L, Niggemeyer C, Linde K, Melchart D, Dobos GJ. Acupuncture in migraine: investigation of autonomic effects. Clinical Journal of Pain. 2008; 24(2): 106-15

Baillie RT, Cecen AA, Erkal C. Normal heartbeat series are nonchaotic, nonlinear, and multifractal: new evidence from semiparametric and parametric tests. Chaos. 2009; 19(2): 028503. doi: 10.1063/1.3152006

Bornas X, Llabrés J, Noguera M, López AM, Gelabert JM, Vila I. Fear induced complexity loss in the electrocardiogram of flight phobics: a multiscale entropy analysis. Biological Psychology. 2006; 73(3): 272-9

Bothe DA, Grignon JB, Olness KN. The effects of a stress management intervention in elementary school children. Journal of Developmental and Behavioral Pediatrics. 2014; 35(1): 62-7

Carvajal R, Wessel N, Vallverdú M, Caminal P, Voss A. Correlation dimension analysis of heart rate variability in patients with dilated cardiomyopathy. Computer Methods and Programs in Biomedicine. 2005; 78(2): 133-40

Cervantes Blásquez JC, Rodas Font G, Capdevila Ortís L. Heart-rate variability and precompetitive anxiety in swimmers. Psicothema. 2009 Nov;21(4):531-6

Chae Y, Park HJ, Kang OS, Lee HJ, Kim SY, Yin CS, Lee H. Acupuncture attenuates autonomic responses to smoking-related visual cues. Complementary Therapies in Medicine. 2011; 19(Suppl 1): S1-7

Chang CH, Huang JL, Ting CT, Chang CS, Chen GH. Atropine-induced HRV alteration is not amended by electroacupuncture on Zusanli. American Journal of Chinese Medicine. 2005; 33(2): 307-14

Chang HA, Chang CC, Tzeng NS, Kuo TB, Lu RB, Huang SY. Heart rate variability in unmedicated patients with bipolar disorder in the manic phase. Psychiatry and Clinical Neurosciences. 2014 Feb 26. doi: 10.1111/pcn.12178

Clark M. n.d. Effect size. More to life than statistical significance – reporting effect size. University of North Texas. www.unt.edu/rss/class/mike/5700/ES5700.ppt [accessed 22.03.14]

Cohen J. A power primer. Psychological Bulletin. 1992; 112(1): 155-9

Dehkordi P, Garde A, Karlen W, Wensley D, Ansermino JM, Dumont GA. Pulse rate variability compared with Heart Rate Variability in children with and without sleep disordered breathing. Conference Proceedings. Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 2013; 2013: 6563-6

Dinas PC, Koutedakis Y, Flouris AD. Effects of active and passive tobacco cigarette smoking on heart rate variability. International Journal of Cardiology. 2013; 163(2): 109-15

Ellis PD. The Essential Guide to Effect Sizes: An introduction to statistical power, meta-analysis and the interpretation of research results. Cambridge, UK: Cambridge University Press, 2010

Fagundes CP, Murray DM, Hwang BS, Gouin JP, Thayer JF, Sollers JJ 3rd, Shapiro CL, Malarkey WB, Kiecolt-Glaser JK. Sympathetic and parasympathetic activity in cancer-related fatigue: more evidence for a physiological substrate in cancer survivors. Psychoneuroendocrinology. 2011; 36(8): 1137-47

Faust O, Acharya UR, Molinari F, Chattopadhyay S, Tamura T. Linear and non-linear analysis of cardiac health in diabetic subjects. Biomedical Signal Processing and Control. 2012; 7(3): 295-302

Frewen J, Finucane C, Savva GM, Boyle G, Coen RF, Kenny RA. Cognitive function is associated with impaired heart rate variability in ageing adults: the Irish longitudinal study on ageing wave one results. Clinical Autonomic Research. 2013; 23(6): 313-23

Friedman BH, Thayer JF. Anxiety and autonomic flexibility: a cardiovascular approach. Biological Psychology. 1998; 49(3): 303-23

Fuller-Rowell TE, Williams DR, Love GD, McKinley PS, Sloan RP, Ryff CD. Race differences in age-trends of autonomic nervous system functioning. Journal of Aging and Health. 2013;25(5): 839-62

Gevirtz R. (2011). Autonomic nervous system markers for psychophysiological, anxiety, and physical disorders. In: Gordon E, Koslow SH. (Eds). Integrative Neuroscience and Personalized Medicine. Oxford : Oxford University Press, 164-80

Haker E, Egekvist H, Bjerring P. Effect of sensory stimulation (acupuncture) on sympathetic and parasympathetic activities in healthy subjects. Journal of the Autonomic Nervous System. 2000; 79(1): 52-9

Hansen AL, Johnsen BH, Thayer JF. Relationship between heart rate variability and cognitive function during threat of shock. Anxiety, Stress, and Coping. 2009; 22(1): 77-89

Hartung J, Knapp G, Sinha BK. (2008). Statistical Meta-Analysis with Application. Hoboken, New Jersey: Wiley

Henry BL, Minassian A, Paulus MP, Geyer MA, Perry W. Heart rate variability in bipolar mania and schizophrenia. Journal of Psychiatric Research. 2010; 44(3): 168-76. doi: 10.1016/j.jpsychires.2009.07.011

Hsu CC, Weng CS, Sun MF, Shyu LY, Hu WC, Chang YH. Evaluation of scalp and auricular acupuncture on EEG, HRV, and PRV. American Journal of Chinese Medicine. 2007; 35(2): 219-30

Hwang DS, Kim HK, Seo JC, Shin IH, Kim DH, Kim YS. Sympathomodulatory effects of Saam acupuncture on heart rate variability in night-shift-working nurses. Complementary Therapies in Medicine. 2011; 19(Suppl 1): S33-40

Imai K, Ariga H, Chen C, Mantyh C, Pappas TN, Takahashi T. Effects of electroacupuncture on gastric motility and heart rate variability in conscious rats. Autonomic Neuroscience. 2008; 138(1-2): 91-8

Imai K, Ariga H, Takahashi T. Electroacupuncture improves imbalance of autonomic function under restraint stress in conscious rats. American Journal of Chinese Medicine. 2009; 37(1): 45-55

Javorka M, Javorková J, Tonhajzerová I, Calkovska A, Javorka K. Heart rate variability in young patients with diabetes mellitus and healthy subjects explored by Poincaré and sequence plots. Clinical Physiology and Functional Imaging. 2005; 25(2): 119-27

Joseph P. Acharya UR, Poo CK, Chee J, Min LC, Iyengar SS, Wei H. Effect of reflexological stimulation on heart rate variability. ITBM-RBM. 2004; 25: 40–5

Kaneko S, Watanabe M, Takayama S, Numata T, Seki T, Tanaka J, Kanemura S, Kagaya Y, Ishii T, Kimura Y, Yaegashi N. Heart rate variability and hemodynamic change in the superior mesenteric artery by acupuncture stimulation of lower limb points: a randomized crossover trial. Evidence-Based Complementary and Alternative Medicine. 2013; 2013: 315982. doi: 10.1155/2013/315982

Khandoker AH, Jelinek HF, Palaniswami M. Identifying diabetic patients with cardiac autonomic neuropathy by heart rate complexity analysis. Biomedical Engineering Online. 2009; 8:3. doi: 10.1186/1475-925X-8-3

Khandoker AH, Karmakar CK, Palaniswami M. Comparison of pulse rate variability with heart rate variability during obstructive sleep apnea. Medical Engineering and Physics. 2011; 33(2): 204-9

Kim CK, McGorray SP, Bartholomew BA, Marsh M, Dicken T, Wassertheil-Smoller S, Curb JD, Oberman A, Hsia J, Gardin J, Wong ND, Barton B, McMahon RP, Sheps DS. Depressive symptoms and heart rate variability in postmenopausal women. Archives of Internal Medicine. 2005; 165(11): 1239–44

Kim DH, Lipsitz LA, Ferrucci L, Varadhan R, Guralnik JM, Carlson MC, Fleisher LA, Fried LP, Chaves PH. Association between reduced heart rate variability and cognitive impairment in older disabled women in the community: Women's Health and Aging Study I. Journal of the American Geriatrics Society. 2006; 54(11): 1751–57

Kubios HRV – Heart Rate Variability Analysis Software. Biosignal Analysis and Medical Imaging Group, University of Eastern Finland. http://kubios.uef.fi/ [accessed 26.03.14]

Kurono Y, Minagawa M, Ishigami T, Yamada A, Kakamu T, Hayano J. Acupuncture to Danzhong but not to Zhongting increases the cardiac vagal component of heart rate variability. Autonomic Neuroscience. 2011; 161(1-2): 116-20

Lackschewitz H, Hüther G, Kröner-Herwig B. Physiological and psychological stress responses in adults with attention-deficit/hyperactivity disorder (ADHD). Psychoneuroendocrinology. 2008; 33(5): 612-24

Lake DE, Richman JS, Griffin MP, Moorman JR. Sample entropy analysis of neonatal heart rate variability. American Journal of Physiology. Regulatory, Integrative and Comparative Physiology. 2002; 283(3): R789-97

Lee JO, Kang SG, Kim SH, Park SJ, Song SW. The Relationship between Menopausal Symptoms and Heart Rate Variability in Middle Aged Women. Korean Journal of Family Medicine. 2011; 32(5): 299-305

Lee MS, Park BJ, Lee J, Park KT, Ku JH, Lee JW, Oh KO, Miyazaki Y. Physiological relaxation induced by horticultural activity: transplanting work using flowering plants. Journal of Physiological Anthropology. 2013; 32: 15. doi: 10.1186/1880-6805-32-15

Li Z, Jiao K, Chen M, Wang C. Effect of magnitopuncture on sympathetic and parasympathetic nerve activities in healthy drivers--assessment by power spectrum analysis of heart rate variability. European Journal of Applied Physiology. 2003; 88(4-5): 404-10

Litscher G, Wang L, Wang X, Gaischek I. Laser acupuncture: two acupoints (baihui, neiguan) and two modalities of laser (658 nm, 405 nm) induce different effects in neurovegetative parameters. Evidence-Based Complementary and Alternative Medicine. 2013; 2013: 432764. doi: 10.1155/2013/432764

McFadden KL, Healy KM, Hoversten KP, Ito TA, Hernández TD. Efficacy of acupressure for non-pharmacological stress reduction in college students. Complementary Therapies in Medicine. 2012; 20(4): 175-82

Maestri R, Pinna GD, Porta A, Balocchi R, Sassi R, Signorini MG, Dudziak M, Raczak G. Assessing nonlinear properties of heart rate variability from short-term recordings: are these measurements reliable? Physiological Measuement. 2007; 28(9): 1067-7

Malik M (Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology). 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. European Heart Journal. 1996; 17(3): 354-81

Markil N, Whitehurst M, Jacobs PL, Zoeller RF. Yoga Nidra relaxation increases heart rate variability and is unaffected by a prior bout of Hatha yoga. Journal of Alternative and Complementary Medicine. 2012; 18(10): 953-8

Mateo M, Blasco-Lafarga C, Martínez-Navarro I, Guzmán JF, Zabala M. Heart rate variability and pre-competitive anxiety in BMX discipline. European Journal of Applied Physiology. 2012; 112(1): 113-23

Matsubara T, Arai YC, Shiro Y, Shimo K, Nishihara M, Sato J, Ushida T. Comparative effects of acupressure at local and distal acupuncture points on pain conditions and autonomic function in females with chronic neck pain. Evidence-Based Complementary and Alternative Medicine. 2011; 2011. pii: 543291. doi: 10.1155/2011/543291

Melillo P, Bracale M, Pecchia L. Nonlinear Heart Rate Variability features for real-life stress detection. Case study: students under stress due to university examination. Biomedical Engineering Online. 2011; 10:96. doi: 10.1186/1475-925X-10-96

Milovanovic B, Milinic N, Trifunovic D, Krotin M, Filipovic B, Bisenic V, Djuric D. Autonomic dysfunction in alcoholic cirrhosis and its relation to sudden cardiac death risk predictors. General Physiology and Biophysics. 2009; 28(Special No): 251-61

Mohebbi M, Ghassemian H. Prediction of paroxysmal atrial fibrillation based on non-linear analysis and spectrum and bispectrum features of the heart rate variability signal. Computer Methods and Programs in Biomedicine. 2012; 105(1): 40-9

Nazeran H, Krishnam R, Chatlapalli S, Pamula Y, Haltiwanger E, Cabrera S. Nonlinear dynamics analysis of heart rate variability signals to detect sleep disordered breathing in children. Conference Proceedings. Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 2006; 1: 3873-8

Nunan D, Sandercock GR, Brodie DA. A quantitative systematic review of normal values for short-term heart rate variability in healthy adults. Pacing and Clinical Electrophysiology. 2010; 33(11): 1407-17

Pincus SM. Assessing serial irregularity and its implications for health. Annals of the New York Academy of Sciences. 2001; 954: 245-67

Pittig A, Arch JJ, Lam CW, Craske MG. Heart rate and heart rate variability in panic, social anxiety, obsessive-compulsive, and generalized anxiety disorders at baseline and in response to relaxation and hyperventilation. International Journal of Psychophysiology. 2013; 87(1): 19-27

Richman JS, Moorman JR. Physiological time-series analysis using approximate entropy and sample entropy. American Journal of Physiology: Heart and Circulatory Physiology. 2000. 278(6): H2039-H2049

Rieger A, Stoll R, Kreuzfeld S, Behrens K, Weippert M. Heart rate and heart rate variability as indirect markers of surgeons' intraoperative stress. International Archives of Occupational and Environmental Health. 2014; 87(2): 165-74

Ruiz-Padial E, Sollers JJ 3rd, Vila J, Thayer JF. The rhythm of the heart in the blink of an eye: emotion-modulated startle magnitude covaries with heart rate variability. Psychophysiology. 2003; 40(2): 306-13

Russoniello CV, Zhirnov YN, Pougatchev VI, Gribkov EN. Heart rate variability and biological age: implications for health and gaming. Cyberpsychology, Behavior and Social Networking. 2013; 16(4): 302-8

Sacre JW, Jellis CL, Marwick TH, Coombes JS. Reliability of heart rate variability in patients with type 2 diabetes mellitus. Diabetic Medicine. 2011; 29(7): e33-40

Sauvet F, Jouanin JC, Langrume C, Van Beers P, Papelier Y, Dussault C. Heart rate variability in novice pilots during and after a multi-leg cross-country flight. Aviation, Space, and Environmental Medicine. 2009; 80(10): 862-9

Schäfer A, Vagedes J. How accurate is pulse rate variability as an estimate of heart rate variability? A review on studies comparing photoplethysmographic technology with an electrocardiogram. International Journal of Cardiology. 2013; 166(1): 15-29

Schubert C, Lambertz M, Nelesen RA, Bardwell W, Choi JB, Dimsdale JE. Effects of stress on heart rate complexity – a comparison between short-term and chronic stress. Biological Psychology. 2009; 80(3): 325-32

Sookan T, McKune AJ. Heart rate variability in physically active individuals: reliability and gender characteristics. Cardiovascular Journal of Africa. 2012; 23(2): 67-72

Streitberger K, Steppan J, Maier C, Hill H, Backs J, Plaschke K. Effects of verum acupuncture compared to placebo acupuncture on quantitative EEG and heart rate variability in healthy volunteers. Journal of Alternative and Complementary Medicine. 2008; 14(5): 505-13

Tarvainen MP, Niskanen JP. 2012. Kubios HRV version 2.1. User's Guide. University of Eastern Finland, Kuopio. http://kubios.uef.fi/media/Kubios_HRV_2.1_Users_Guide.pdf [accessed 26.03.14]

Taş-Cebe G, Cummings C. A single-blinded controlled trial comparing corticosteroid injection, acupuncture and home-based exercises for the treatment of primary contracted (frozen) shoulder. Journal of the Acupuncture Association of Chartered Physiotherapists. 2013 (Autumn); 25-38

Thalheimer W, Cook S. 2002. How to calculate effect sizes from published research articles: a simplified methodology. http://education.gsu.edu/coshima/EPRS8530/Effect_Sizes_pdf4.pdf [accessed 01.10.13, but no longer available at this site]

Thayer JF, Hall M, Sollers JJ 3rd, Fischer JE. Alcohol use, urinary cortisol, and heart rate variability in apparently healthy men: Evidence for impaired inhibitory control of the HPA axis in heavy drinkers. International Journal of Psychophysiology. 2006; 59(3): 244-50

Uchida S, Kagitani F, Hotta H. Neural mechanisms of reflex inhibition of heart rate elicited by acupuncture-like stimulation in anesthetized rats. Autonomic Neuroscience. 2010; 157(1-2): 18-23

Umetani K, Singer DH, McCraty R, Atkinson M. Twenty-four hour time domain heart rate variability and heart rate: relations to age and gender over nine decades. Journal of the American College of Cardiology. 1998; 31(3): 593-601

Vaillancourt DE, Newell KM. Changing complexity in human behavior and physiology through aging and disease. Neurobiology of Aging. 2002; 23(1): 1-11

Vuksanović V, Gal V. Nonlinear and chaos characteristics of heart period time series: healthy aging and postural change. Autonomic Neuroscience. 2005; 121(1-2): 94-100

Wang L, Valentini J, Sugimoto K, Cheng W, Cheng G, Geng H, Gaischek I, Kuang H, Litscher G. Biomedical teleacupuncture between China and Austria using heart rate variability, Part 1: poststroke patients. Evidence-Based Complementary and Alternative Medicine. 2011;2011:782489

Wong JS, Lu WA, Wu KT, Liu M, Chen GY, Kuo CD. A comparative study of pulse rate variability and heart rate variability in healthy subjects. Journal of Clinical Monitoring and Computing. 2012; 26(2): 107-14

Wu SB, Zhou MQ, Zhou YP, Wang KM, Wang YL, Chen YN, Cao J. [Study on the mechanism of electroacupuncture of different acupoint groups in the treatment of cerebrocardiac syndrome in rats]. Zhen Ci Yan Jiu. 2009; 34(5): 315-8, 328

Yang FS, Hong B, Tang QY. (2001). Approximate entropy and its application in biosignal analysis. In: Akay M. (Ed.). Nonlinear Biomedical Signal Processing: Dynamic Analysis and Modeling. II. New York: IEEE Press, pp. 72-91

Yang ZK, Wu ML, Xin JJ, He W, Su YS, Shi H, Wang XY, Hu L, Jing XH, Litscher G. Manual acupuncture and laser acupuncture for autonomic regulations in rats: observation on heart rate variability and gastric motility. Evidence- Based Complementary and Alternative Medicine. 2013; 2013: 276320. doi: 10.1155/2013/276320

Zulli R, Nicosia F, Borroni B, Agosti C, Prometti P, Donati P, De Vecchi M, Romanelli G, Grassi V, Padovani A. QT dispersion and heart rate variability abnormalities in Alzheimer's disease and in mild cognitive impairment. Journal of the American Geriatrics Society. 2005; 53(12): 2135–39

PubMed searches:

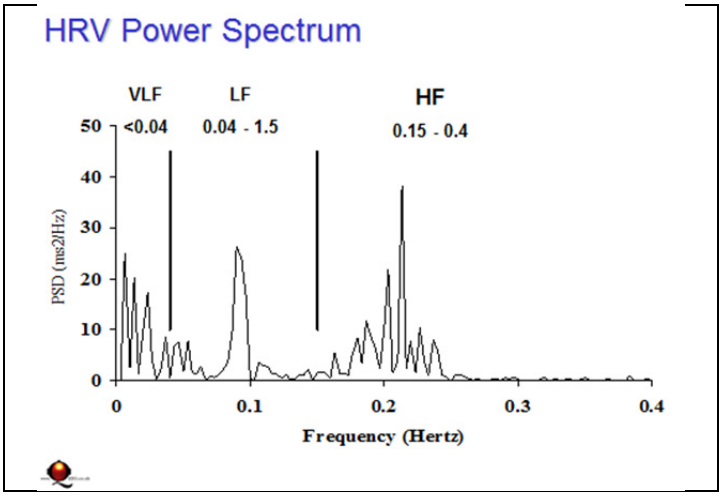
1. '("heart rate variability" OR HRV OR PRV OR "pulse rate variability") AND acupuncture': 117 (27.02.14).

2. Estimate of total number of studies in PubMed (28.02.14), using 'a* OR b* OR … z*': 15583797. If add '… AND acupuncture': 13364; but as 'acupuncture' alone results in 20265 hits, presumably total = (20265 x 15583797) / 13364, or 23631072, of which 20265 is 0.09%.

APPENDICES

Appendix A. LF peak frequency

In principle, LF peak frequency should approach 0.1 Hz (respiratory rhythm) with increased relaxation:



Changes in LF pk were assessed from Kubios.

In Pilots 1 and 3, mean LF pk moved further from 0.1 Hz over the course of the session, and in Pilot 2 LF pk approached 0.1 Hz more closely. In all Pilots, the direction of change varied with Hz, ID and V, but with little agreement across different Pilots.

EO1 to EO2.

	EO1	EO2	closer to 0.1 (SD dec)	significant
P1	0.077 (0.031)	0.075 (0.023)	n (y)	n
P1 Hz			2.5 y (y) 10 n (y)	n
P1 Amp				n/a
P1 ID			4 y 3 n	n
P1 V			v1 n v2 y	n

ID	EO1	EO2	closer to 0.1 Hz (SD dec)	increases
2185	0.056641	0.070313	y (n)	y
5611	0.095703	0.103516	y (y)	y
7032	0.058594	0.089844	y (-)	y
8311	0.070313	0.050781	n (-)	n
8680	0.113282	0.048828	n (y)	n
8875	0.066407	0.095703	y (n)	y
8954	0.064454	0.060547	n (y)	n
All			4 y 3 n	4 y 3 n

Pilot factors	EO1	EO2	closer to 0.1 (SD dec)	significant
P2	0.087 (0.031)	0.087 (0.031)	y (y)	n
P2 Hz			2.5 n (y) 10 y (n)	n
P2 Amp				n/a
P2 ID			4 y 8 n	y
P2 V			v1 n v2 y v3 n v4 n	y (baseline difference)

ID	EO1	EO2	closer	increases
3235	0.093756	0.086427	n (n)	n
4290	0.121088	0.108887	y (n)	n
5044	0.080078	0.079102	n (n)	n
5115	0.051758	0.057943	y (y)	y
5453	0.091797	0.072266	n (y)	n
6899	0.101559	0.086919	n (y)	n
7112	0.057617	0.088542	y (n)	y
7338	0.087891	0.085286	n (n)	n
7501	0.09668	0.111328	n (y)	y
7815	0.107422	0.113281	n (n)	y
7904	0.055664	0.0625	y (y)	y
9960	0.092773	0.081055	n (y)	n
All			4 y 8 n	5 y 7 n

Pilot factor	EO1	EO2	closer to 0.1 (SD dec)	significant
P3	0.101 (0.025)	0.083 (0.027)	n (n)	y
P3 Hz			2.5 n (y) 10 n (n)	y
P3 Amp				n/a

P3 ID			0 y 4 n	n
P3 V			v1 n v2 n v3 v3 n v4 n	n

ID	EO1	EO2	closer	increases
2185	0.105469	0.083869	n (n*)	n
5611	0.083984	0.092529	n (n)	y*
8680	0.118164	0.079834	n* (y*)	n*
8954	0.09668	0.075195	n* (n)	n*
All			0 y 4 n	1 y 3 n

* Agreement with P1

Appendix B. Differences with stimulation frequency (**Hz**) in the three Pilots

Pilot 1

Hz (by **Visit**)

2.5 vs 10 Hz	RR	SDNN	RMS SD	HFpwr	LF/HF	ApEn	SampEn	D_2	*N* signif
Both visits	ns	ns	0.024 [0.020] (0.009)	(0.039)	ns	ns	0.007 [0.007] (0.005)	ns	2 (3)
Visit 1	ns	<0.001 [0.007] (<0.001)	0.002 [0.006] (0.002)	(0.027)	0.012 [0.025]	0.010 [0.017] (0.011)	<0.001 [0.001] (<0.001)	ns	5 (5)
Visit 2	ns	ns	ns	ns	0.023	ns	ns	ns	1 (0)

T-test; equal variances not assumed (Mann-Whitney U test); 2-tailed significance

This suggests an order effect, with more differences (or, in general, more variation?) in Visit 1 compared with Visit 2.

Pilot 1

Hz (by **Loc**)

No difference between HRV measures for the two frequencies were significant, at any of the 6 Locations.

Pilot 1

Hz (by **ID**)

Significant differences between HRV measures for the two frequencies, by ID.

2.5 vs 10 Hz	RR	SDNN	RMS SD	HFpwr	LF/HF	ApEn	SampEn	D_2	*N* signif
2185	0.016 [0.016] (0.016)			(0.025)	0.005 [0.036] (0.004)				2 (3)
5611	<0.001 [0.001] (0.004)		0.003 [0.012] (0.004)	<0.001 [0.004] (0.004)	0.008 [0.045] (0.004)			0.003 [0.013] (0.004)	5 (5)
7032/ 8311	<0.001 [0.001] (0.004)				0.024 (0.016)		0.006 [0.013] (0.010)	<0.001 [0.007] (0.004)	4 (4)
8680	<0.001 [0.001] (0.004)		0.001 [0.004] (0.004)	0.001 [0.002] (0.004)	0.010 (0.004)	0.038 (0.037)		0.011 (0.010)	6 (6)
8875	0.009 [0.015] (0.025)			0.006 [0.014] (0.010)	0.014 (0.004)		0.009 [0.020] (0.010)	(0.025)	4 (5)
8954	<0.001 [0.002] (0.004)	0.001 [0.037] (0.004)	0.001 [0.010] (0.004)	0.009 [0.042] (0.004)			<0.001 [0.002] (0.004)	0.001 [0.010] (0.004)	6 (6)
N signif	6 (6)	1 (1)	3 (3)	4 (5)	5 (5)	1 (1)	3 (3)	4 (5)	**27 (29)**

T-test [Bootstrap; equal variances assumed] (Mann-Whitney U test); 2-tailed significance

Here SDNN and ApEn were the measures least sensitive to differences in stimulation frequency, and RR the most sensitive (followed by LF/HF and then HFpwr and D_2, tied). 8680 and 8954 were the participants most sensitive to stimulation frequency difference, and 2185 the participant least sensitive to this difference.

Pilot 2

(EA segments only)

Hz (by **Visit**)

2.5 vs 10 Hz	RR	SDNN	RMS SD	HFpwr	LF/HF	ApEn	SampEn	D_2	*N* signif
ALL visits						[0.049] (0.039)			
Visit 1									0 (0)
Visit 2	0.006 [0.011] (0.011)	0.003 [0.004] (0.003)	0.001 [0.002] (0.001)	0.005 [0.010] (0.004)				0.018 [0.012] (0.024)	5 (5)
Visit 3		0.039 (0.010)	0.019 [0.027] (0.001)	0.010 [0.018] (0.010)				<0.001 [0.001] (0.001)	4 (4)
Visit 4		0.038		0.024		0.041			3 (2)

		[0.036] (0.019)		[0.038]		[0.035] (0.032)			
N signif	1 (1)	3 (3)	2 (2)	3 (2)		1 (1)		2 (2)	**12 (11)**

1-way ANOVA [Bootstrap; equal variances assumed] (Kruskal-Wallis test); 2-tailed significance

The lack of significant difference in Visit 1 as against the other Visits is intriguing, suggesting a possible cumulative effect.

Pilot 2

(EA segments only)

Hz (by **Loc**)

2.5 vs 10 Hz	RR	SDNN	RMS SD	HFpwr	LF/HF	ApEn	SampEn	D_2	*N* signif
ALL	ns	ns	ns	ns	ns	(0.039) [0.049]	ns	ns	
B	<0.001 [0.001] (<0.001)								1 (1)
L					0.015 [0.047] (0.008)				1 (1)
R	0.046 [0.049] (0.039)	0.001 [0.005] (0.002)	0.007 [0.011] (0.005)		0.004 [0.012]				4 (3)
T		0.045 [0.042]	0.028 [0.026]	0.009 [0.011]		<0.001 [0.001] (<0.001)			4 (1)
N signif	2 (2)	2 (1)	2 (1)	1 (0)	2 (1)	1 (1)	0 (0)	0 (0)	**10 (6)**

In Pilot 2, there were more frequency-dependent differences when stimulation was applied at R and T than at B and L.

Pilot 2

(EA segments only)

Hz (by **ID**)

Significant differences between frequencies for the various participants in Pilot 2.

2.5 vs 10 Hz	RR	SDNN	RMS SD	HFpwr	LF/HF	ApEn	SampEn	D_2	*N* signif
3235					0.049		0.010 [0.013] (0.021)	0.07 [0.024]	3 (1)
4290			0.004 [0.005]						1 (1)

			(0.005)						
5044									0 (0)
5115	0.027 [0.046] (0.036)							0.031 [0.037] (0.027)	2 (2)
5453							0.001 [0.006] (0.002)		1 (1)
6899				0.003 [0.002] (0.004)	0.002 [0.042] (0.004)				2 (2)
7112					0.038 (0.046)				1 (1)
7338	0.001 [0.005] (0.001)								1 (1)
7501	0.001 [0.001] (0.005)	(0.036)	0.002 [0.003] (0.005)	0.018 [0.035] (0.012)				(0.016)	3 (5)
7815						0.029 [0.049] (0.027)			1 (1)
7904				[0.049]					0 (0)
9960		0.019 [0.027] (0.036)						0.015 [0.040] (0.006)	2 (2)
N signif	3 (3)	1 (2)	2 (2)	2 (2)	3 (2)	1 (1)	2 (2)	3 (3)	**17 (17)**

T-test [Bootstrap; equal variances assumed] (Mann-Whitney U test); 2-tailed significance, asymptotic

Here 7501 shows most frequency-dependent differences in HRV measures, 5044 and 7904 fewest such differences. RR and D_2 (and LF/HF) are the HRV measures here most sensitive to frequency effects, ApEn the least sensitive. 7501 is the participant most sensitive to stimulation frequency, 5044 and 7904 the least sensitive.

Pilot 3

(EA segments only)

Hz (by **Visit**)

2.5 vs 10 Hz	RR	SDNN	RMS SD	HFpwr	LF/HF	ApEn	SampEn	D_2	*N* signif
ALL visits	ns	ns	ns	ns	ns	0.019 (0.019)	ns	ns	
Visit 1	<0.001 (0.021)	0.002 (0.021)	0.001 (0.021)	0.009 (0.021)	0.009 (0.021)			0.008 (0.021)	6 (6)
Visit 2									0 (0)
Visit 3						0.025 (0.021)	0.036 (0.021)		2 (2)
Visit 4									

N signif	1 (1)	1 (1)	1 (1)	1 (1)	1 (1)	1 (1)	1 (1)	1 (1)	**8 (8)**

Pilot 3

(TEAS segments only)

Hz (by **Visit**)

2.5 vs 10 Hz	RR	SDNN	RMS SD	HFpwr	LF/HF	ApEn	SampEn	D_2	*N* signif
ALL	ns	ns	ns	ns	ns	ns	ns	ns	
Visit 1	0.001 (0.021)	0.009 (0.021)	0.021 (0.021)	0.026 (0.021)				<0.001 (0.021)	5 (5)
Visit 2	0.026 (0.021)		0.041						2 (1)
Visit 3		(0.021)	(0.021)	(0.021)		0.031 (0.021)	0.007 (0.021)		2 (5)
Visit 4	(0.043)								0 (1)
N signif	2 (3)	1 (2)	2 (2)	1 (2)	0 (0)	1 (1)	1 (1)	1 (1)	**9 (12)**

T-test [Bootstrap; equal variances not assumed] (Mann-Whitney U test); 2-tailed significance, asymptotic

There are more significant changes in Visit 1 than visits 2 or 4 (and for the EA segments, than visit 3 as well), suggesting a possible order effect, but many possible factors could account for this.

Pilot 3

(EA *and* TEAS segments)

2.5 vs 10 Hz	RR	SDNN	RMS SD	HFpwr	LF/HF	ApEn	SampEn	D_2	*N* signif
ALL	ns	ns	ns	ns	ns	0.026 [0.016] (0.032)	ns	ns	1 (1)
Visit 1	<0.001 [0.001] (0.001)	0.002 [0.046] (0.001)	0.002 [0.044] (0.001)	0.006 [0.049] (0.001)	(0.021)			<0.001 [0.001] (0.001)	5 (6)
Visit 2	0.004 [0.017] (0.005)								1 (1)
Visit 3		0.032 (0.001)	0.033 (0.021)	(0.006)		<0.001 [0.019] (0.001)	<0.001 [0.003] (0.002)		4 (5)
Visit 4			0.044 (0.021)						1 (1)
N signif	2 (2)	2 (2)	3 (3)	1 (2)	0 (1)	2 (2)	1 (1)	1 (1)	**12 (14)**

T-test [Bootstrap; equal variances assumed] (Mann-Whitney U test); 2-tailed significance, asymptotic

The same order effect is evident here as when the EA and TEAS segments are considered separately.

Pilot 3

(EA segments only)

Hz (by **Loc**)

No significant differences were found for any Location between the two frequencies.

Pilot 3

(TEAS segments only)

Hz (by **Loc**)

2.5 vs 10 Hz	RR	SDNN	RMS SD	HFpwr	LF/HF	ApEn	SampEn	D_2
ALL	ns	ns	ns	ns	ns	ns	ns	ns
B								
L					0.019 (0.021)			
R								
T								

T-test [Bootstrap; equal variances not assumed] (Mann-Whitney U test); 2-tailed significance, asymptotic

Pilot 3

(EA segments only)

Hz (by **ID**)

2.5 vs 10 Hz	RR	SDNN	RMS SD	HFpwr	LF/HF	ApEn	SampEn	D_2	*N* signif
2185	0.030 (0.021)				(0.043)				1 (2)
5611									0 (0)
8680		(0.043)		(0.043)					0 (2)
8954				0.039 (0.043)					1 (1)
N signif	1 (1)	0 (1)	0 (0)	1 (2)	0 (1)	0 (0)	0 (0)	0 (0)	**2 (5)**

T-test [Bootstrap; equal variances not assumed] (Mann-Whitney U test); 2-tailed significance, asymptotic

Both 2185 and 8680 showed significant differences in two HRV measures for stimulation frequency; 5611 showed none.

Pilot 3

(TEAS segments only)

Hz (by **ID**)

2.5 vs 10 Hz	RR	SDNN	RMS SD	HFpwr	LF/HF	ApEn	SampEn	D_2	*N* signif
2185									0 (0)
5611					(0.043)				0 (1)
8680									0 (0)
8954									0 (0)
N signif	0 (0)	0 (0)	0 (0)	0 (0)	0 (1)	0 (0)	0 (0)	0 (0)	**0 (1)**

T-test [Bootstrap; equal variances assumed] (Mann-Whitney U test); 2-tailed significance, asymptotic

Clearly, using TEAS, there was little significant frequency dependence effect, whereas there was slightly more when using EA.

Pilot 3

(EA *and* TEAS segments)

Hz (by **ID**)

2.5 vs 10 Hz	RR	SDNN	RMS SD	HFpwr	LF/HF	ApEn	SampEn	D_2	*N* signif
2185	0.006 [0.013] (0.021)				0.045 (0.021)				2 (2)
5611									0 (0)
8680			0.049						1 (0)
8954				0.005 [0.010] (0.016)	0.025 [0.023] (0.021)				2 (2)
N signif	1 (1)		1 (0)	1 (1)	2 (2)	0 (0)	0 (0)	0 (0)	**5 (4)**

T-test [Bootstrap; equal variances assumed] (Mann-Whitney U test); 2-tailed significance, asymptotic

These results are comparable to those for EA alone. Again, there is little association between these results and those for Pilot 1, although 8954 appears quite responsive to the frequency difference in both Pilots.

Compilation of results for all Pilots: HRV measures (Hz)

Diff%

10 Hz – 2.5 Hz	Highest abs diff%	Lowest abs diff%
P1	RMS SD SDNN	RR SampEn
P2	HFpwr D_2	SampEn SDNN
P3 EA	HFpwr RMS SD	RR D_2
P3 TEAS	LF/HF SDNN	SampEn RR
mean	HFpwr LF/HF	RR D_2

Measures that tend to show greatest differences for the two frequencies in more than one Pilot are SDNN (2), RMS SD (2) and HFpwr (2), and those showing least differences are RR (3) and SampEn (3).

However, replacing LF/HF with HF/LF:

10 Hz – 2.5 Hz	Highest abs diff%	Lowest abs diff%
P1	RMS SD SDNN	RR **HF/LF**
P2	**HF/LF** HFpwr	SampEn SDNN
P3 EA	HFpwr RMS SD	RR **HF/LF**
P3 TEAS	SDNN HFpwr	SampEn RR
mean	HFpwr SDNN	RR D_2

Measures that tend to show greatest differences for the two frequencies in more than one Pilot are now SDNN (2), RMS SD (2) and HFpwr (3), and those showing least differences are RR (3), HF/LF (2) and SampEn (2).

Significance of differences in HRV measures according to stimulation frequency.

2.5 vs 10 Hz	RR	SDNN	RMS SD	HFpwr	LF/HF	ApEn	SampEn	D_2	*N* signif
P1			0.024 [0.020] (0.009)	(0.039)			0.007 [0.007] (0.005)		2 (3)
P2						[0.049] (0.039)			1 (1)

P3 EA						0.019 (0.019)			1 (1)
P3 TEAS									0 (0)

T-test [Bootstrap; equal variances not assumed] (Mann-Whitney U test); 2-tailed significance, asymptotic.

No measures demonstrate significant differences over all Pilots, although ApEn demonstrates significant differences in two separate Pilots. Using **HF/LF** instead of LF/HF did not change this for any Pilot.

Compilation of results for all Pilots: Hz (by Visit)

Pilot	2.5 vs 10 Hz	RR	SDNN	RMS SD	HFpwr	LF/HF	ApEn	SampEn	D_2	*N* signif
P1	V 1		<0.001 [0.007] (<0.001)	0.002 [0.006] (0.002)	(0.027)	0.012 [0.025]	0.010 [0.017] (0.011)	<0.001 [0.001] (<0.001)		5 (5)
	V 2					0.023				1 (0)
P2[a]	V 1									0 (0)
	V 2	0.006 [0.011] (0.011)	0.003 [0.004] (0.003)	0.001 [0.002] (0.001)	0.005 [0.010] (0.004)				0.018 [0.012] (0.024)	5 (5)
	V 3		0.039 (0.010)	0.019 [0.027] (0.001)	0.010 [0.018] (0.010)				<0.001 [0.001] (0.001)	4 (4)
	V 4		0.038 [0.036] (0.019)		0.024 [0.038]		0.041 [0.035] (0.032)			3 (2)
P3 EA	V 1	<0.001 (0.021)	0.002 (0.021)	0.001 (0.021)	0.009 (0.021)	0.009 (0.021)			0.008 (0.021)	6 (6)
	V 2									0 (0)
	V 3						0.025 (0.021)	0.036 (0.021)		2 (2)
	V 4									
P3 TEAS	V 1	0.001 (0.021)	0.009 (0.021)	0.021 (0.021)	0.026 (0.021)				<0.001 (0.021)	5 (5)
	V 2	0.026 (0.021)		0.041						2 (1)
	V 3		(0.021)	(0.021)	(0.021)		0.031 (0.021)	0.007 (0.021)		2 (5)
	V 4	(0.043)								0 (1)

T-test [Bootstrap] (Man-Whitney), 2-tailed asymptotic. a. Bootstrap equal variances not assumed.

In P1, P2 and P3 (EA), there are more significant differences between measures for the two frequencies in V1 than subsequent visits, but in P3 (TEAS), most significant differences are found in V1 *and* V3.

Compilation of results for all Pilots: Hz (by Loc)

2.5 vs 10 Hz	RR	SDNN	RMS SD	HFpwr	LF/HF	ApEn	SampEn	D_2
ALL	ns	ns	ns	ns	ns	0.026 [0.016] (0.032)	ns	ns
B	ns	ns	ns	ns	ns	ns	ns	ns
L	ns	ns	ns	ns	ns [0.043] (0.046)	ns	ns	ns
R	ns	ns	ns	ns	ns	ns	ns	ns
T	ns	ns	ns	ns	ns	ns	ns	ns

1-way ANOVA [Bootstrap] (Kruskal-Wallis test)

There are few convincing patterns of interaction between Hz and Loc across all Pilots. Comparison between the Pilots is a little more revealing:

Pilot	2.5 vs 10 Hz	RR	SDNN	RMS SD	HFpwr	LF/HF	ApEn	SampEn	D_2	*N* signif
P1	B									
	L									
	R									
	T									
P2	B	<0.001 [0.001] (<0.001)								1 (1)
	L					0.015 [0.047] (0.008)				1 (1)
	R	0.046 [0.049] (0.039)	0.001 [0.005] (0.002)	0.007 [0.011] (0.005)		0.004 [0.012]				4 (3)
	T		0.045 [0.042]	0.028 [0.026]	0.009 [0.011]		<0.001 [0.001] (<0.001)			4 (1)
P3 EA	B									
	L									
	R									
	T									

P3 TEAS	B									
	L					0.019 (0.021)				1 (1)
	R									
	T									

It is interesting that the only Pilot showing any significant results is P2, where the same location was stimulated consistently for 20 minutes per session, whereas in P1 and P3, various locations (6 in P, 2 in P3) were stimulated for only 5 minutes each per session. In P2, more measures showed significant differences for the two frequencies when stimulation was at R or T points, rather than L or B.

Compilation of results for all Pilots: Hz (by ID)

Sensitivity to differences in stimulation frequency.

Pilot	Least sensitive measure	Most sensitive measure	Least sensitive to stim frequency	Most sensitive to stim frequency
P1 (TEAS)	SDNN ApEn	RR	2185	8680 8954
P2 (EA)	ApEn	RR D_2	*5044* *7904*	*7501*
P3 (EA)[a]	RMS SD ApEn SampEn D_2	HFpwr	5611	2185 8680
P3 (TEAS)[b]				
P3 (EA & TEAS)	ApEn SampEn D_2	LF/HF	5611	2185 8954

a. Data poor; b. Data insufficient

Although data is sparse, there are some parallels here in responsiveness for two of those who took part in both Pilot 1 and Pilot 3 (8680 and 8954). That there are fewer significant differences overall in Pilot 3, compared to in Pilot 1, might indicate a habituation or learning process, or a difference when measures are recorded during, rather than after, the intervention.

Of the HRV measures, it appears that APEn is consistently insensitive to frequency-induced differences, but that RR may be the most sensitive measure. These findings are consistent with those in **Table E1** below (for Bonferroni post hoc tests).

Interestingly, 5611 reported subjective changes that differed with the two frequencies used, but this is not really supported by the HRV findings.

Appendix C. Differences with visit (**V**) in the three Pilots

Pilot 1

Visit 1 vs visit 2	RR	SDNN	RMS SD	HFpwr	LF/HF	ApEn	SampEn	D_2	*N* signif
	ns	(0.022)	0.022 (0.037)	ns	(0.001)	ns	0.010 (0.001)	ns	2 (4)

T-test; equal variances not assumed (Mann-Whitney U test); 2-tailed significance

Pilot 2

(EA segments only)

Among visits	RR	SDNN	RMS SD	HFpwr	LF/HF	ApEn	SampEn	D_2
	ns	ns	ns	0.035	ns	ns	ns	ns

1-way ANOVA; equal variances not assumed (Kruskal-Wallis test); 2-tailed significance

Pilot 3

(EA *and* TEAS segments)

Among visits	RR	SDNN	RMS SD	HFpwr	LF/HF	ApEn	SampEn	D_2
ALL	(0.048)	ns	ns	ns	ns	ns	ns	ns

(Kruskal-Wallis test); 2-tailed significance

Compilation of results for all Pilots (Visit)

	RR	SDNN	RMS SD	HFpwr	LF/HF	ApEn	SampEn	D_2	*N* signif
P1[a]		(0.022)	0.022 (0.037)		(0.001)		0.010 (0.001)		2 (4)
P2[b]				0.033 (0.035)	0.029				2 (1)
P3 EA[b]									
P3 TEAS[b]									

a. T-test (Mann-Whitney U test); b. 1-way ANOVA (Kruskal-Wallis test). 2-tailed significance.

There are more significant differences among visits in P1 than in the other Pilots.

Appendix D. Differences with stimulation location (**Loc**) in the three Pilots

Pilot 1

The Kruskal-Wallis test showed no significant differences among Locations in Pilot 1.

Pilot 2

The Kruskal-Wallis test showed no significant differences among Locations in Pilot 2.

Pilot 3 (EA)

The Kruskal-Wallis test showed no significant differences among Locations in Pilot 3 (EA).

Pilot 3 (TEAS)

The Kruskal-Wallis test showed no significant differences among Locations in Pilot 3 (TEAS).

No significant differences among Locations were found in any Pilot using 1-way ANOVA or Kruskal-Wallis test with Bootstrap.

Appendix E. Differences with participant (**ID**) in the three Pilots

Pilot 1

among IDs	RR	SDNN	RMS SD	HFpwr	LF/HF	ApEn	SampEn	D_2
	<0.001 (<0.001)	<0.001 (<0.001)	<0.001 (<0.001)	0.001 (<0.001)	0.001 (<0.001)	0.030 (0.028)	0.013 (0.012)	<0.001 (<0.001)

1-way ANOVA (Kruskal-Wallis test)

Non-significant differences between participants (post-hoc Bonferroni test, with default settings)

Among IDs	RR	SDNN	RMS SD	HFpwr	LF/HF	ApEn	SampEn	D_2	*N* **non-**signif
2185	5611 7032 8954	5611 8311 8680 8954	5611 8311 8680 8954	5611 7032 8311 8680 8954	5611 7032 8311 8680 8875 8954	5611 7032 8311 8680 8875	5611 8311 8680 8875 8954	7032 8954	34 (34)
5611	8680 8954	7032 8311 8680 8954	7032 8311 8680 8954	7032 8311 8680 8954	7032 8311 8875 8954	7032 8311 8680 8875 8954	7032 8311 8680 8875 8954	8311 8680 8875	31 (38)
7032		8311	8311	8311	8311	8311	8680		22 (33)

		8680 8875	8680 8875 8954	8680 8875 8954	8680 8875 8954	8680 8875 8954	8875 8954		
8311		8680 8954	8680 8875 8954	8680 8875 8954	8680 8875 8954	8680 8875 8954	8680 8875 8954	8875	18 (36)
8680	8875 8954	8954	8954	8954	8875 8954	8875 8954	8875 8954		11 (36)
8875			8954	8954	8954	8954	8954		5 (23)
8954									0 (34)
N **non-**signif	7	14	17	18	20	20	19	6	**121 (234)**

Thus the least sensitive HRV measures to differences among participants in Pilot 1 were LF/HF and ApEn, the most sensitive being RR and D_2. Of the participants, 8875 conformed least to the overall group response pattern (only 23 non-significant differences), and 5611 conformed most (38 non-significant differences).

Pilot 2

(EA segments only)

among IDs	RR	SDNN	RMS SD	HFpwr	LF/HF	ApEn	SampEn	D_2
	<0.001 (<0.001)	<0.001 (<0.001)	<0.001 (<0.001)	<0.001 (<0.001)	<0.001 (<0.001)	<0.001 (<0.001)	<0.001 (<0.001)	<0.001 (<0.001)

1-way ANOVA (Kruskal-Wallis test)

These results, identical for the two tests used, are highly significant.

Non-significant differences between participants (post-hoc Bonferroni test, with default settings).

Among IDs	RR	SDNN	RMS SD	HFpwr	LF/HF	ApEn	SampEn	D_2	*N* **non-**signif
3235	6899 9960	5115 7338 9960	5044 5115 5453 7112 9960	5044 5115 5453 7112 7501 7904 9960	4290 5115 5453 6899 7112 7338 7815 7904 9960	4290 5044 5453 6899 7501 7815 7904 9960	4290 5453 6899 7112 7815 9960	5044 5115 7338 7501 7815 9960	46 (46)
4290	5044 5453 7338 7501	5044 7501	7815	7815	5115 5453 6899 7112	5044 5453 6899 7338	5115 5453 6899 7112	5044 7501 7815 9960	35 (38)

					7338 7815 7904 9960	7501 7815 7904 9960	7815 7904 9960		
5044	5453 7338 7501 7904	7501 7815	5115 7501 9960	5115 5453 6899 7112 7501 7904 9960		5453 7112 7338 7501 7815 7904 9960	7338 7501	5115 7338 7501 7815 9960	30 (38)
5115	7112 7815 7904	5453 7112 7904 9960	7112 7501 9960	5453 6899 7112 7501 7904 9960	5453 6899 7112 7338 7815 7904 9960	7112 7338 7501 7815 7904 9960	5453 6899 7112 7815 7904 9960	7112 7338 7501 9960	39 (50)
5453	7904	6899 7112 7904	6899 7112 7904	6899 7112 7501 7904	6899 7112 7338 7815 7904 9960	6899 7501 7815 7904 9960	6899 7112 7815 7904 9960	6899 7904	29 (45)
6899		7112 7904	7904	7112 7501 7904	7112 7338 7815 7904 9960	7815 9960	7112 7815 7904 9960	7904	18 (36)
7112	7815	7904	7904	7501 7904	7338 7815 7904 9960	7338 7501 7815 7904 9960	7815 7904 9960	5115 7904	19 (44)
7338	7501	7815 9960	7501 7815	7815	7815 7904 9960	7501 7815 7904 9960	7501	7501 7815 9960	17 (35)
7501	9960	7815	9960	7904 9960		7815 7904 9960		7815 9960	10 (39)
7815	7904	7338 7501			7904 9960	7904 9960	7904 9960	9960	10 (46)
7904					9960	9960	9960		3 (47)
9960									0
N **non-**signif	18	22	20	33	45	51	37	30	**256**

Thus in Pilot 2 the least sensitive HRV measures to differences among participants were LF/HF and ApEn, the most sensitive RR, SDNN and RMS SD. Of the participants, 7338 conformed least to the

overall group response pattern (only 35 non-significant differences), and 5115 conformed most (50 non-significant differences).

Pilot 3

(EA segments only)

among IDs	RR	SDNN	RMS SD	HFpwr	LF/HF	ApEn	SampEn	D_2
	<0.001 (<0.001)	0.043 (<0.001)	0.016 (0.002)	ns (0.001)	ns	ns	<0.001 (<0.001)	<0.001 (0.003)

1-way ANOVA (Kruskal-Wallis test; asymptotic significance)

Non-significant differences between participants (post-hoc Bonferroni test, with default settings)

Among IDs	RR	SDNN	RMS SD	HFpwr	LF/HF	ApEn	SampEn	D_2	*N* **non**-signif
2185	8954	5611 8680 8954	5611 8680 8954	5611 8680 8954	5611 8680 8954	5611 8680 8954	5611 8680 8954	8680 8954	21 (21)
5611		8680 8954	8680 8954	8680 8954	8680 8954	8680 8954		8680	11 (17)
8680	8954			8954	8954	8954	8954		5 (18)
8954									0 (18)
N **non**-signif	2	5	5	6	6	6	4	3	

Here the least sensitive HRV measures to differences among participants were again HFpwr, LF/HF and ApEn, the most sensitive being RR. Of the participants, 5611 conformed least to the overall group response pattern (17 non-significant differences), and 2185 conformed most (21 non-significant differences).

Pilot 3

(TEAS segments only)

among IDs	RR	SDNN	RMS SD	HFpwr	LF/HF	ApEn	SampEn	D_2
	<0.001 (0.001)	0.013 (0.002)	0.004 (0.003)	0.008			<0.001 (0.001)	0.001 (0.018)

1-way ANOVA (Kruskal-Wallis test; asymptotic significance)

Non-significant differences between participants (post-hoc Bonferroni test, with default settings)

Among IDs	RR	SDNN	RMS SD	HFpwr	LF/HF	ApEn	SampEn	D_2	*N* **non**-signif
2185	8680 8954	5611 8954	5611 8954	5611 8680 8954	5611 8680 8954	5611 8680 8954	8680 8954	8680 8954	19 (19)
5611		8680	8680	8680	8680	8680		8680	11 (16)

		8954	8954	8954	8954	8954			
8680	8954			8954	8954	8954	8954	8954	6 (18)
8954									0 (19)
N **non**-signif	3	4	4	6	6	6	3	4	**36 (62)**

Here the least sensitive HRV measures to differences among participants were HFpwr, LF/HF and ApEn, the most sensitive are RR and SampEn. Of the participants, 5611 conformed least to the overall group response pattern (16 non-significant differences), and 2185 and 8954 conformed most (19 non-significant differences).

Pilot 3

(EA *and* TEAS segments)

among IDs	RR	SDNN	RMS SD	HFpwr	LF/HF	ApEn	SampEn	D_2
	<0.001 (<0.001)	<0.001 (<0.001)	<0.001 (<0.001)	0.005 (<0.001)	ns	ns	<0.001 (<0.001)	<0.001 (<0.001)

Non-significant differences between participants (post-hoc Bonferroni test, with default settings)

Among IDs	RR	SDNN	RMS SD	HFpwr	LF/HF	ApEn	SampEn	D_2	*N* **non**-signif
2185	8954	5611 8954	5611 8954	5611 8954	5611 8680 8954	5611 8680 8954	8680 8954	8680 8954	17 (17)
5611		8680 8954	8954	8680 8954	8680 8954	8680 8954		8680	10 (15)
8680	8954				8954	8954			3 (12)
8954									0 (16)
N **non**-signif	2	4	3	4	6	6	2	3	**30 (60)**

Here the least sensitive HRV measures to differences among participants were again LF/HF and ApEn, the most sensitive being RR and SampEn. Of the participants, 8680 conformed least to the overall group response pattern (12 non-significant differences), and 2185 conformed most (17 non-significant differences).

Compilation of results for all Pilots (ID)

among IDs	RR	SDNN	RMS SD	HFpwr	LF/HF	ApEn	SampEn	D_2
P1	<0.001 (<0.001)	<0.001 (<0.001)	<0.001 (<0.001)	0.001 (<0.001)	0.001 (<0.001)	0.030 (0.028)	0.013 (0.012)	<0.001 (<0.001)
P2	<0.001 (<0.001)	<0.001 (<0.001)	<0.001 (<0.001)	<0.001 (<0.001)	<0.001 (<0.001)	<0.001 (<0.001)	<0.001 (<0.001)	<0.001 (<0.001)
P3 EA	<0.001 (<0.001)	0.043 (<0.001)	0.016 (0.002)	ns (0.001)	ns	ns	<0.001 (<0.001)	<0.001 (0.003)
P3 TEAS	ns	<0.001 (0.001)	0.013 (0.002)	0.004 (0.003)	0.008 (ns)	ns	ns	<0.001 (0.001)

1-way ANOVA (Kruskal-Wallis test, 2-way asymptotic significance)

Here, those measures most likely to indicate significant differences among participants across all Pilots are SDNN, RMS SD and D_2; those least likely are ApEn, followed by RR, LF/HF and SampEn.

Table E1. Comparing *non*-significant differences in HRV measures and participants in Pilots 1-3 (post-hoc Bonferroni test)

Pilot	Least sensitive measure	Most sensitive measure	Most conformist response	Least conformist response
P1 (TEAS)	LF/HF ApEn	RR D_2	5611	*8875*
P2 (EA)	LF/HF ApEn	RR SDNN RMS SD	*5115*	*7338*
P3 (EA)	HFpwr LF/HF ApEn	RR	2185	5611
P3 (TEAS)	HFpwr LF/HF ApEn	RR SampEn	2185 8954	5611
P3 (EA & TEAS)	LF/HF ApEn	RR SampEn	2185	8680

In *italics*: participants in only one Pilot.

Here, there is agreement between the two Pilots that LF/HF and ApEn are the least sensitive measures to detect differences between individual participants, RR being the most sensitive. However, there is no consistent pattern of least or most conformist for those who took part in both Pilot studies (2185, 5611, 8680 and 8954), indicating that sensitivity to stimulation frequency effects will change on different occasions.

Appendix F. Differences with baseline state (**B**) in the three Pilots

HRV measures at baseline (EO1) were compared to the median for the whole sample (all participants, in all segments) , transformed into binary numbers (1 if > median, 0 if ≤ median), and relabelled as RR-ini, SDNN-ini, etc.. In addition, peak LF frequency was coded into 1s and 0s in the same way (as 'LFpk-ini'), and also relative to 0.1 (see Appendix A), as 'LFpk-0.1'.

Number of significant differences in HRV values during stimulation segments dependent on initial state, using the Mann-Whitney U test (2-tailed asymptotic significance):

Pilot 1	RR	SDNN	RMS SD	HFpwr	LF/HF	ApEn	SampEn	D_2	***N* All**
RR-ini	0.000	0.000	0.000	0.000	0.000	0.135	0.706	0.000	6
SDNN-ini	0.000	0.000	0.000	0.000	0.368	0.107	0.003	0.000	6
RMSSD-ini	0.000	0.000	0.000	0.000	0.235	0.864	0.003	0.000	6
HFpwr-ini	0.000	0.016	0.003	0.002	0.277	0.496	0.177	0.000	5
LF/HF-ini	0.451	0.028	0.954	0.846	0.000	0.157	0.170	0.016	3
ApEn-ini	0.063	0.581	0.057	0.040	0.000	0.033	0.173	0.180	3
SampEn-ini	0.012	0.115	0.566	0.685	0.364	0.014	0.025	0.064	3
D_2-ini	0.000	0.000	0.008	0.000	0.205	0.553	0.837	0.000	5
LFpk-ini	0.000	0.000	0.000	0.000	0.630	0.658	0.658	0.000	5
LFpk-0.1	0.022	0.804	0.232	0.161	0.019	0.640	0.173	0.040	3
***N* All**	8	7	6	7	4	2	3	8	**45**

Thus initial values of RR, SDNN, RMS SD, HFpwr, D_2 and LFpk (but not LFpk-0.1) differentiated between values in stimulation segments of over half the HRV measures. Those measures that were most affected by these initial values were RR and D_2, whereas ApEn and SampEn were least affected.

Correlation ratio *eta* (η)

Pilot 1	RR	SDNN	RMS SD	HFpwr	LF/HF	ApEn	SampEn	D_2	**mean**
RR-ini	**0.772**	0.276	0.469	0.379	0.413	0.144	0.034	0.563	0.381
SDNN-ini	0.480	0.636	0.651	0.511	0.045	0.168	0.387	0.461	0.417
RMSSD-ini	0.415	0.593	0.555	0.439	0.011	0.018	0.378	0.420	0.354
HFpwr-ini	0.443	0.286	0.280	0.274	0.134	0.082	0.168	0.569	0.280
LF/HF-ini	0.059	0.262	0.001	0.063	0.421	0.236	0.170	0.331	0.193
ApEn-ini	0.138	0.107	0.200	0.154	0.212	0.247	0.158	0.085	0.163
SampEn-ini	0.313	0.241	0.162	0.156	0.265	0.277	0.235	0.157	0.226
D_2-ini	0.544	0.330	0.144	0.151	0.190	0.119	0.017	**0.838**	0.292
LFpk-ini	0.411	0.473	0.375	0.352	0.076	0.010	0.022	0.573	0.287
LFpk-0.1	0.201	0.057	0.053	0.002	0.175	0.054	0.138	0.298	0.122
mean	0.397	0.356	0.315	0.275	0.196	0.145	0.174	0.444	**0.288**

The three comparisons in red are those where baseline was associated with values of the same measure in the stimulation segments; those in **bol**d are where *eta* (η) ≥0.750; those underlined are the maxima in their rows. Mean *eta* (η) in the bottom row was based on the 9 '-ini' baseline measures (excluding LFpk-0.1). LFpk-0.1 was excluded from further analysis as having the lowest mean *eta* (η) across all measures.

Mean *eta* (η) was largest for **SDNN**, then **RR**, of the initial values, and lowest for ApEn and LF/HF. For the values in stimulation segments, it was largest for **RR** and **SDNN**, and lowest for ApEn and SampEn.

Significant differences in HRV values during stimulation segments. Median values were calculated for both P2 and P3 together.

Pilot 2	RR	SDNN	RMS SD	HFpwr	LF/HF	ApEn	SampEn	D_2	***N* All**
RR-ini	0.000	0.360	0.111	0.850	0.850	0.000	0.057	0.381	2
SDNN-ini	0.566	0.000	0.000	0.000	0.743	0.339	0.000	0.000	5
RMSSD-ini	0.164	0.000	0.000	0.000	0.493	0.292	0.000	0.000	5
HFpwr-ini	0.512	0.000	0.000	0.000	0.003	0.899	0.000	0.000	6
LF/HF-ini	0.259	0.151	0.000	0.000	0.000	0.277	0.302	0.059	3
ApEn-ini	0.111	0.000	0.002	0.020	0.004	0.000	0.003	0.046	7
SampEn-ini	0.010	0.000	0.000	0.000	0.543	0.934	0.000	0.000	6
D_2-ini	0.959	0.000	0.000	0.000	0.010	0.064	0.011	0.000	6
LFpk-ini	0.000	0.000	0.002	0.004	0.965	0.008	0.018	0.178	6
***N* All**	3	7	8	8	4	3	7	6	**46**

Here initial values of SDNN, RMS SD, HFpwr, ApEn, SampEn, D_2 and LFpk differentiated between values in stimulation segments of over half the HRV measures. Those measures that were most affected by these initial values were RMS SD and HFpwr, whereas RR and LF/HF were least affected.

Correlation ratio *eta* (η)

Pilot 2	RR	SDNN	RMS SD	HFpwr	LF/HF	ApEn	SampEn	D_2	**mean**
RR-ini	**0.775**	0.040	0.100	0.000	0.074	0.301	0.133	0.054	0.185
SDNN-ini	0.000	**0.779**	0.729	0.617	0.263	0.042	0.500	**0.774**	0.463
RMSSD-ini	0.146	0.691	0.702	0.574	0.240	0.048	0.425	0.655	0.435
HFpwr-ini	0.094	0.623	0.715	0.626	0.059	0.000	0.353	0.694	0.396
LF/HF-ini	0.070	0.072	0.391	0.472	0.415	0.061	0.064	0.146	0.211
ApEn-ini	0.147	0.323	0.174	0.073	0.290	0.346	0.241	0.220	0.227
SampEn-ini	0.140	0.479	0.387	0.355	0.229	0.027	0.415	0.489	0.315
D_2-ini	0.025	0.656	0.567	0.380	0.314	0.133	0.151	0.691	0.365
LFpk-ini	0.384	0.296	0.230	0.267	0.187	0.218	0.200	0.153	0.242
mean	0.198	0.440	0.444	0.374	0.230	0.131	0.276	0.431	**0.315**

Mean *eta* (η) was largest for **SDNN**, then **RMS SD**, of the initial values, and lowest for RR and LF/HF. For the values in stimulation segments, it was largest for **RMS SD** and **SDNN**, and lowest for ApEn and RR.

Significant differences in HRV values during stimulation segments. Median values were calculated for both P2 and P3 together.

Pilot 3 EA	RR	SDNN	RMS SD	HFpwr	LF/HF	ApEn	SampEn	D_2	***N* All**
RR-ini	0.000	0.008	0.036	0.008	0.414	0.669	0.003	0.002	6
SDNN-ini	0.055	0.000	0.000	0.000	0.734	0.327	0.035	0.000	5
RMSSD-ini	0.035	0.000	0.000	0.000	0.187	1.000	0.546	0.001	5
HFpwr-ini	0.001	0.000	0.000	0.000	0.350	0.392	0.007	0.000	6

LF/HF-ini	0.276	0.697	0.312	0.586	0.043	0.484	0.043	0.697	2
ApEn-ini	0.259	0.056	0.150	0.173	0.293	0.938	1.000	0.312	0
SampEn-ini	0.004	0.000	0.000	0.000	0.820	0.044	0.006	0.000	7
D_2-ini	0.001	0.000	0.000	0.000	0.494	0.342	0.004	0.000	6
LFpk-ini	0.699	0.082	0.005	0.009	0.060	0.923	0.562	0.111	2
***N* All**	5	6	7	7	1	1	6	6	**39**

Here initial values of RR, SDNN, RMS SD, HFpwr, SampEn and D_2 differentiated between values in stimulation segments of over half the HRV measures. Those measures that were most affected by these initial values were RMS SD and HFpwr, whereas LF/HF and ApEn were least affected.

Correlation ratio *eta* (η)

Pilot 3 EA	RR	SDNN	RMS SD	HFpwr	LF/HF	ApEn	SampEn	D_2	**mean**
RR-ini	0.740	0.083	0.012	0.048	0.020	0.076	0.585	0.617	0.276
SDNN-ini	0.415	0.484	0.454	0.309	0.133	0.176	0.430	0.634	0.379
RMSSD-ini	0.439	0.457	0.497	0.315	0.246	0.008	0.137	0.609	0.339
HFpwr-ini	0.603	0.448	0.428	0.279	0.051	0.127	0.530	**0.790**	0.407
LF/HF-ini	0.164	0.015	0.111	0.070	0.586	0.113	0.389	0.084	0.192
ApEn-ini	0.192	0.275	0.258	0.198	0.162	0.020	0.004	0.138	0.156
SampEn-ini	0.545	0.520	0.496	0.346	0.078	0.366	0.554	0.702	0.451
D_2-ini	0.632	0.442	0.394	0.296	0.003	0.137	0.537	0.661	0.388
LFpk-ini	0.112	0.205	0.266	0.169	0.428	0.003	0.050	0.277	0.189
mean	0.427	0.325	0.324	0.226	0.190	0.114	0.357	0.501	**0.308**

Mean *eta* (η) was largest for **SampEn**, then **HFpwr**, of the initial values, and lowest for ApEn and LFpk. For the values in stimulation segments, it was largest for **D_2** and **RR**, and lowest for ApEn and LF/HF.

Significant differences in HRV values during stimulation segments.

Pilot 3 TEAS	RR	SDNN	RMS SD	HFpwr	LF/HF	ApEn	SampEn	D_2	***N* All**
RR-ini	0.000	0.001	0.001	0.002	0.785	0.213	0.004	0.000	6
SDNN-ini	0.498	0.004	0.005	0.032	0.734	0.851	0.122	0.046	4
RMSSD-ini	0.624	0.000	0.000	0.005	0.940	0.522	0.498	0.013	4
HFpwr-ini	0.043	0.000	0.000	0.000	0.669	0.815	0.013	0.000	6
LF/HF-ini	0.073	0.640	0.755	0.640	0.276	0.161	0.139	0.243	0
ApEn-ini	0.018	0.815	0.785	0.460	0.697	0.938	0.938	0.350	1
SampEn-ini	0.254	0.003	0.004	0.030	0.939	0.342	0.019	0.063	4
D_2-ini	0.017	0.000	0.000	0.000	0.704	0.569	0.020	0.000	6
LFpk-ini	0.469	0.267	0.176	0.334	0.809	0.961	0.440	0.562	0
***N* All**	4	6	6	6	0	0	4	5	**31**

Here initial values of RR, HFpwr and D_2 differentiated between values in stimulation segments of over half the HRV measures. Those measures that were most affected by these initial values were SDNN, RMS SD and HFpwr, whereas LF/HF and ApEn were least affected.

Correlation ratio *eta* (η)

Pilot 3 TEAS	RR	SDNN	RMS SD	HFpwr	LF/HF	ApEn	SampEn	D_2	**mean**
RR-ini	**0.799**	0.098	0.157	0.044	0.213	0.271	0.529	**0.861**	0.372
SDNN-ini	0.165	0.363	0.353	0.271	0.201	0.134	0.306	0.362	0.269
RMSSD-ini	0.135	0.445	0.510	0.316	0.191	0.184	0.122	0.420	0.290
HFpwr-ini	0.434	0.460	0.500	0.318	0.162	0.038	0.463	0.703	0.385
LF/HF-ini	0.332	0.037	0.067	0.069	0.018	0.190	0.273	0.266	0.157
ApEn-ini	0.424	0.136	0.079	0.123	0.145	0.071	0.018	0.139	0.142
SampEn-ini	0.243	0.395	0.403	0.315	0.190	0.112	0.469	0.381	0.314
D_2-ini	0.495	0.507	0.553	0.353	0.183	0.032	0.415	**0.757**	0.412
LFpk-ini	0.092	0.195	0.228	0.168	0.084	0.020	0.118	0.112	0.127
mean	0.347	0.293	0.317	0.220	0.154	0.117	0.301	0.445	**0.274**

Mean *eta* (η) was largest for **D_2**, then **HFpwr**, of the initial values, and lowest for LFpk and ApEn. For the values in stimulation segments, it was largest for **D_2** and **RR**, and lowest for ApEn and LF/HF.

Summarising significant differences in HRV values during stimulation segments resulting from high or low initial values.

Pilot	initial HRV with most effect	HRV most affected	HRV least affected	initial *eta* (η) largest	initial *eta* (η) least	stim *eta* (η) largest	stim *eta* (η) least
P1	RR, SDNN, RMS SD, HFpwr, D_2, LFpk	RR D_2,	ApEn SampEn	SDNN RR	ApEn LF/HF	RR SDNN	ApEn SampEn
P2	SDNN, RMS SD, HFpwr, ApEn, SampEn, D_2, LFpk	RMS SD HFpwr	RR LF/HF	SDNN RMS SD	RR LF/HF	RMS SD SDNN	ApEn RR
P3 EA	RR, SDNN, RMS SD, HFpwr, SampEn, D_2	RMS SD HFpwr	LF/HF ApEn	SampEn HFpwr	ApEn LFpk	D_2 RR	ApEn LF/HF
P3 TEAS	RR, HFpwr, D_2	SDNN RMS SD HFpwr	LF/HF ApEn	D_2 HFpwr	LFpk ApEn	D_2 RR	ApEn LF/HF

*Association of ID and Visit with baseline values (**B**)*

Results of the *Chi*-square (χ^2) test for Values during stimulation segments:

Pilot 1	ID	df 6	Visit	df 1
RR-ini	22.629	0.001	0.059	ns
SDNN-ini	48.000	<0.001	18.514	<0.001
RMSSD-ini	34.971	<0.001	31.092	<0.001
HFpwr-ini	45.000	<0.001	1.029	ns
LF/HF-ini	59.657	<0.001	0.059	ns
ApEn-ini	48.000	<0.001	18.514	<0.001
SampEn-ini	58.500	<0.001	4.114	0.043
D_2-ini	72.000	<0.001	0.059	ns
LFpk-ini	56.000	<0.001	6.171	0.013
mean	49.417		8.846	

Pilot 2	ID	df 11	Visit	df 3
RR-ini	275.293	<0.001	12.048	0.007
SDNN-ini	289.527	<0.001	7.581	ns
RMSSD-ini	277.031	<0.001	17.592	0.001
HFpwr-ini	216.834	<0.001	13.506	0.004
LF/HF-ini	236.779	<0.001	1.931	ns
ApEn-ini	214.450	<0.001	17.765	<0.001
SampEn-ini	131.500	<0.001	47.245	<0.001
D_2-ini	305.512	<0.001	7.838	0.049
LFpk-ini	169.097	<0.001	49.523	<0.001
mean	235.114		19.448	

Pilot 3	ID	df 3	Visit	df 3
RR-ini	105.600	<0.001	11.733	0.008
SDNN-ini	66.000	<0.001	66.000	<0.001
RMSSD-ini	88.000	<0.001	22.000	<0.001
HFpwr-ini	129.067	<0.001	11.733	0.008
LF/HF-ini	35.200	<0.001	35.200	<0.001
ApEn-ini	58.667	<0.001	58.667	<0.001
SampEn-ini	75.429	<0.001	53.079	<0.001
D_2-ini	97.778	<0.001	30.730	<0.001
LFpk-ini	13.538	<0.001	13.538	0.004
mean	74.364		33.631	

Appendix G. Correlations between HRV measures in the three Pilots

Pilot 1

(Complete sessions, all segments)

Correlations between HRV measures for complete sessions, where Spearman's *rho* ≥0.4; [] indicates significance but *rho* < 0.4. 'All' indicates correlations when both stimulation frequencies were taken together, '2.5 Hz' and '10 Hz' where the data was considered separately for each.

HRV measure	All	2.5 Hz	10 Hz	2.5 Hz (*N* signif corrs)	10 Hz (*N* signif corrs)
RR	SDNN** RMS SD** HFpwr** [-LF/HF**] [SampEn*] D_2**	[HFpwr*] [SampEn*] D_2**	SDNN** RMS SD** HFpwr** -LF/HF** [SampEn*] D_2**	3 (3)	6 (6)
SDNN	RMS SD** HFpwr** [LF/HF*] [-ApEn**] -SampEn** D_2**	RMS SD** HFpwr** [LF/HF**] [-ApEn**] -SampEn** [D_2*]	RMS SD** HFpwr** [-SampEn*] D_2**	6 (6)	4 (5)
RMS SD	HFpwr** [-ApEn**] [-SampEn**] D_2**	HFpwr** [-ApEn**] -SampEn**	HFpwr** [-LF/HF*] D_2**	3 (4)	3 (5)
HFpwr	[-LF/HF*] [-ApEn*] [-SampEn*] D_2**	[-ApEn*] [-SampEn*] [D_2**]	[-LF/HF*] D_2**	3 (6)	2 (5)
LF/HF	-SampEn**	[-SampEn**]	[ApEn*] -SampEn**	1 (2)	2 (5)
ApEn		[SampEn*]		1 (4)	0 (1)
SampEn				0 (6)	0 (3)
D_2				0 (3)	0 (4)
N signif corrs	21 [16** 5*]	17 [11** 6*]	17 [12** 5*]	17 (34)	**17 (34)**

Pilot 1

(TEAS segments only)

Correlations between HRV measures for TEAS segments, where Spearman's *rho* ≥0.4; [] indicates significance but *rho* < 0.4.

HRV measure	All	2.5 Hz	10 Hz	2.5 Hz (*N* signif corrs)	10 Hz (*N* signif corrs)
RR	SDNN** RMS SD** HFpwr** [-LF/HF**] D_2**	[HFpwr*] D_2**	SDNN** RMS SD** HFpwr** -LF/HF** [SampEn*] D_2**	2 (2)	6 (6)
SDNN	RMS SD** HFpwr** D_2** -Sampen**	RMS SD** HFpwr** [LF/HF*] [-ApEn*] [-SampEn**]	RMS SD** HFpwr** D_2**	5 (5)	3 (4)
RMS SD	HFpwr** [-ApEn*] [-SampEn*] D_2**	HFpwr** -ApEn** -SampEn**	HFpwr** [-LF/HF*] D_2**	3 (4)	3 (5)
HFpwr	[-LF/HF*] D_2**	-SampEn* [D_2*]	[-LF/HF*] D_2**	2 (5)	2 (5)
LF/HF	[ApEn*] -SampEn**	[-SampEn*]	[ApEn*] -SampEn**	1 (2)	2 (5)
ApEn			-SampEn**	0 (2)	1 (2)
SampEn				0 (4)	0 (3)
D_2				0 (2)	0 (4)
N signif corrs	17 [13** 4*]	13 [7** 6*]	17 [13** 4*]	13 (26)	**17 (34)**

Note that this gives the same total as when all session segments are considered, although different correlations are responsible for the same totals.

Correlations between changes in HRV values between session baseline and follow-up segments, where Spearman's *rho* ≥0.4.

HRV measure	All	2.5 Hz	10 Hz	2.5 Hz (*N* signif corrs)	10 Hz (*N* signif corrs)
RR		SampEn**		1 (1)	0 (0)
SDNN	HFpwr** RMS SD**	RMS SD** HFpwr**	RMS SD* HFpwr*	2 (2)	2 (2)
RMS SD	HFpwr**	HFpwr**	HFpwr* D_2*	1 (2)	2 (3)
HFpwr			D_2**	0 (2)	1 (3)
LF/HF				0 (0)	0 (0)
ApEn		SampEn*		1 (1)	0 (0)
SampEn				0 (2)	0 (0)

D_2				0 (0)	0 (2)
N signif corrs	3 [3** 0*]	5 [4** 1*]	5 [1** 4*]	5 (10)	**5 (10)**

** 2-tailed significance at 0.01 level; * at 0.05 level; [significance of Pearson's r].

Pilot 2

(EA segments only)

Correlations between HRV measures for EA segments, where Spearman's *rho* ≥0.4; [] indicates significance but *rho* < 0.4.

HRV measure	All	2.5 Hz	10 Hz	2.5 Hz (*N* signif corrs)	10 Hz (*N* signif corrs)
RR	[RMS SD**] [HFpwr*] [-LF/HF*] [-ApEn**] [SampEn**]	[SDNN*] [RMS SD**] [HFpwr**] [-ApEn**] [D_2**]	[-ApEn**] SampEn**	5 (5)	2 (2)
SDNN	RMS SD** HFpwr** [LF/HF**] [-ApEn**] -SampEn** D_2**	RMS SD** HFpwr** [-ApEn**] -SampEn** D_2**	RMS SD** HFpwr** [LF/HF**] [-SampEn**] D_2**	5 (6)	5 (5)
RMS SD	HFpwr** [-LF/HF*] [-SampEn**] D_2**	HFpwr** [-LF/HF**] [-SampEn**] D_2**	HFpwr** [-SampEn*] D_2**	4 (6)	3 (4)
HFpwr	[-LF/HF**] [-SampEn**] D_2**	-LF/HF** [-SampEn**] D_2**	[-LF/HF**] [-SampEn**] D_2**	3 (6)	3 (5)
LF/HF	[-SampEn*]	[-SampEn*]		1 (3)	0 (2)
ApEn			[-SampEn**] D2*	0 (2)	2 (3)
SampEn	[-D_2**]	[-D_2**]	[-D_2*]	1 (5)	1 (6)
D_2				0 (5)	0 (5)
N signif corrs	20 [16** 4*]	19 [17** 2*]	16 [13** 3*]	19 (38)	**16 (32)**

Correlations between changes in HRV values between session baseline and follow-up segments, where Spearman's *rho* ≥0.4.

HRV measure	All	2.5 Hz	10 Hz	2.5 Hz (*N* signif corrs)	10 Hz (*N* signif corrs)
RR	[SDNN**] RMS SD* HFpwr** [-LF/HF*] [D_2*]	HFpwr*	SDNN** HFpwr**	1 (1)	2 (2)
SDNN	RMS SD** HFpwr** [KF/HF*]	HFpwr** RMS SD** LF/HF*	HFpwr**	3 (3)	1 (2)

RMS SD	HFpwr** [-LF/HF*]	HFpwr**	HFpwr* -LF/HF**	1 (2)	2 (2)
HFpwr				0 (3)	0 (3)
LF/HF	[-SampEn*]			0 (1)	0 (1)
ApEn	[D_2**]		D_2*	0 (0)	1 (1)
SampEn				0 (0)	0 (0)
D_2				0 (0)	0 (1)
N signif corrs	12 [6** 6*]	5 [3** 2*]	6 [4** 2*]	5 (10)	**6 (12)**

** 2-tailed significance at 0.01 level; * at 0.05 level.

Pilot 3

(EA segments only)

Correlations between HRV measures for EA segments, where Spearman's *rho* ≥0.4; [] indicates significance but *rho* < 0.4.

HRV measure	All	2.5 Hz	10 Hz	2.5 Hz (*N* signif corrs)	10 Hz (*N* signif corrs)
RR	SDNN** RMS SD** HFpwr** [-LF/HF*] [-SampEn**] D_2**	RMS SD** HFpwr** -LF/HF* D_2*	SDNN* RMS SD* HFpwr* -SampEn* D_2**	4 (4)	5 (5)
SDNN	RMS SD** HFpwr** -SampEn* D_2**	RMS SD** HFpwr** D_2**	RMS SD** HFpwr** D_2*	3 (3)	3 (4)
RMS SD	HFpwr** -LF/HF* D_2**	HFpwr** D_2**	HFpwr** D_2*	2 (4)	2 (4)
HFpwr	[-LF/HF*] [-SampEn*] D_2**	D_2**	D_2*	1 (4)	1 (4)
LF/HF			[-D_2*]	0 (1)	1 (1)
ApEn	-SampEn**	-SampEn**		1 (1)	0 (0)
SampEn				0 (1)	0 (1)
D_2				0 (4)	0 (5)
N signif corrs	17 [12** 5*]	11 [9** 2*]	11 [4** 8*]	11 (22)	**12 (24)**

Pilot 3

(EA segments only)

Correlations between changes in HRV values between session baseline and follow-up segments, where Spearman's *rho* ≥0.4.

HRV measure	All	2.5 Hz	10 Hz	2.5 Hz (*N* signif corrs)	10 Hz (*N* signif corrs)
RR	-SampEn*	ApEn** [-SampEn**]		2 (2)	0 (0)
SDNN	HFpwr**		HFpwr**	0 (0)	1 (1)
RMS SD	HFpwr**	HFpwr*	HFpwr*	1 (1)	1 (1)
HFpwr				0 (1)	0 (2)
LF/HF	-ApEn*		-ApEn**	0 (0)	1 (1)
ApEn	-SampEn*	-SampEn**		1 (2)	0 (1)
SampEn				0 (2)	0 (0)
D_2				0 (0)	0 (0)
N signif corrs	5 [2** 3*]	4 [3** 1*]	3 [2** 1*]	4 (8)	**3 (6)**

** 2-tailed significance at 0.01 level; * at 0.05 level.

Pilot 3

(TEAS segments only)

Correlations between HRV measures for TEAS segments, where Spearman's *rho* ≥0.4; [] indicates significance but *rho* < 0.4.

HRV measure	All	2.5 Hz	10 Hz	2.5 Hz (*N* signif corrs)	10 Hz (*N* signif corrs)
RR	SSDN* RMS SD* D_2**	SDNN** RMS SD** HFpwr** D_2**	D_2*	4 (4)	1 (1)
SDNN	RMS SD** HFpwr** -SampEn* D_2**	RMS SD** HFpwr** D_2**	RMS SD** HFpwr** -SampEn**	3 (4)	3 (3)
RMS SD	HFpwr** [-SampEn*] D_2**	HFpwr** D_2**	HFpwr** -SampEn*	2 (4)	2 (3)
HFpwr	-SampEn** D_2**	D_2**	-SampEn*	1 (4)	1 (3)
LF/HF				0 (0)	0 (0)
ApEn	-SampEn**	-SampEn**		1 (1)	0 (0)
SampEn				0 (1)	0 (3)
D_2				0 (4)	0 (1)
N signif corrs	13 [9** 4*]	11 [11** 0*]	7 [4** 3*]	11 (22)	**7 (14)**

Pilot 3

(TEAS segments only)

Correlations between changes in HRV values between session baseline and follow-up segments, where Spearman's *rho* ≥0.4.

HRV measure	All	2.5 Hz	10 Hz	2.5 Hz (*N* signif corrs)	10 Hz (*N* signif corrs)
RR				0 (0)	0 (0)
SDNN	HFpwr* LF/HF* D_2*	HFpwr** D_2*		2 (2)	0 (0)
RMS SD	HFpwr**		HFpwr*	0 (0)	1 (1)
HFpwr	D_2**	D_2*	D_2**	1 (2)	1 (2)
LF/HF			-ApEn*	0 (0)	1 (1)
ApEn				0 (0)	0 (1)
SampEn				0 (0)	0 (0)
D_2				0 (2)	0 (1)
N signif corrs	5	3 [1** 2*]	3 [1** 2*]	3 (6)	**3 (6)**

** 2-tailed significance at 0.01 level; * at 0.05 level.

Compilation of results for all Pilots (Correlations between HRV measures)

In the Tables below, 'Most' and 'Least' indicate HRV measures with most or least numbers of significant correlations; 'Most –' indicates those measures with most *negative* correlations.

Correlations between HRV measures for stimulation (EA or TEAS) segments, where Spearman's *rho* ≥0.4; [] indicates significance but *rho* < 0.4.

Pilot	Hz	*N*	Most	Most –	Least
P1	2.5 Hz	13 [7** 6*]	SDNN HFpwr (5)	SampEn (4)	RR LF/HF ApEn D_2 (2)
	10 Hz	17 [13** 4*]	RR (6)	LF/HF (4)	ApEn (2)
P2	2.5 Hz	19 [17** 2*]	SDNN HFpwr (6)	SampEn (5)	ApEn (2)
	10 Hz	16 [13** 3*]	SDNN HFpwr (5)	SampEn (5)	LF/HF (2)
P3 EA	2.5 Hz	11 [9** 2*]	RR RMS SD HFpwr (4)	RR LF/HF ApEn SampEn	LF/HF ApEn SampEn (1)

				(1)	
	10 Hz	11 [4** 8*]	RR (5)	RR LF/HF SampEn D_2 (1)	ApEn (0)
P3 TEAS	2.5 Hz	11 [11** 0*]	RR SDNN RMS SD HFpwr D_2 (4)	ApEn SampEn (1)	LF/HF (0)
	10 Hz	7 [4** 3*]	SDNN RMS SD HFpwr SampEn (3)	SampEn (3)	LF/HF ApEn (0)
Summary	2.5 Hz	54 (44** 10*)	RR (2) SDNN (3) RMS SD (2) HFpwr (4) D_2 (1)	RR (1) LF/HF (1) ApEn (2) SampEn (4)	RR (1) LF/HF (3) ApEn (3) sampEn (1) D_2 (1)
	10 Hz	49 (34** 18*)	RR (2) SDNN (2) RMS SD (1) HFpwr (2) SampEn (1)	RR (1) LF/HF (2) SampEn (3) D_2 (1)	LF/HF (2) ApEn (3)

Thus there are more significant correlations overall **for 2.5 Hz than 10 Hz** (with more significant at the 0.01 level and fewer at the 0.05 level), but the proportion of these is not significantly different from that expected by chance.

At both frequencies, SDNN and HFpwr are most often involved in more significant correlations across all Pilots than other measures, and LF/HF and ApEn least often involved. SampEn is involved in more negative correlations with other measures than any other measure, again at both frequencies.

Correlations between changes in HRV values between session baseline and follow-up segments, where Spearman's *rho* ≥0.4.

Pilot	Hz	*N*	Most	Most –	Least
P1	2.5 Hz	5 [4** 1*]	SDNN RMS SD HFpwr SampEn (2)	none	LF/HF D_2 (0)
	10 Hz	5 [1** 4*]	RMS SD HFpwr (3)	none	RR LF/HF ApEn

					SampEn (0)
P2	2.5 Hz	5 [3** 2*]	SDNN HFpwr (3)	none	ApEn SampEn D_2 (0)
	10 Hz	6 [4** 2*]	HFpwr (3)	RMS SD LF/HF (1)	SampEn (0)
P3 EA	2.5 Hz	4 [3** 1*]	RR ApEn SampEn (2)	SampEn (2)	SDNN LF/HF D_2 (0)
	10 Hz	3 [2** 1*]	HFpwr (2)	LF/HF ApEn (1)	RR SampEn D_2 (0)
P3 TEAS	2.5 Hz	3 [1** 2*]	SDNN HFpwr D_2 (2)	none	RR RMS SD LF/HF ApEn SampEn (0)
	10 Hz	3 [1** 2*]	HFpwr (2)	LF/HF ApEn (1)	RR SDNN SampEn (0)
Summary	2.5 Hz	17 (11** 6*)	RR (1) SDNN (3) RMS SD (1) HFpwr (3) ApEn (1) SampEn (2) D_2 (1)	SampEn (1)	RR (1) SDNN (1) RMS SD (1) LF/HF (3) ApEn (2) SampEn (2) D_2 (3)
	10 Hz	17 (8** 9**)	RMS SD (1) HFpwr (4)	RMS SD (1) LF/HF (3) ApEn (2)	RR (3) SDNN LF/HF (1) ApEn (1) SampEn (4) D_2 (1)

Here there are the same numbers of significant correlations overall for 2.5 Hz and 10 Hz (although slightly more significant at the 0.01 level and fewer at the 0.05 level for 2.5 Hz).

At both frequencies, HFpwr is most often involved in more significant correlations across all Pilots than other measures (though at 2.5 Hz, SDNN also appears frequently). At 2.5 Hz, LF/HF is least often involved, and at 10 Hz, SampEn. SampEn is the only measure involved in a negative correlations with other measures at 2.5 Hz (by default, there being few negative correlations at this frequency); at 10 Hz, LF/HF is involved in more negative correlations than the other measures.

Significant correlations (by **Visit**)

Pilot	V1	V2	V3	V4
P1	19 (14** 5*)	15 (12** 3*)		
P2	16 (11*8 5*)	14 (11** 3*)	11 (8** 3*)	14 (11** 3*)
P3 EA	17 (9** 8*)	8	3 (1** 2*)	8
P3 TEAS	9 (4** 5*)	7	5 (4** 1*)	6
P2 & P3	42	29		
All	61	44		

Thus there are more significant correlations in Visit 1 than subsequent visits.

Significant correlations (by **Loc**)

Pilot	Most correlations	Fewest correlations
P1	Bilat [10: 4** 6*]	LLSS [7: 6** 1*]
P2	L [16: 10** 6*]	T [13: 7** 6*]
P3 EA	L [10: 6** 4*]	T [4: 1** 3*]
P3 TEAS	L [11: 6** 5*]	T [4]
All	L most	T least

Although numbers are not greatly different, of the Locs, L tends to show most correlations and T least.

Significant correlations (by **ID**)

Pilot	Most correlations	Fewest correlations
P1	7032 [20: 10** 10*] 8680 [13: 9** 4*]	8311 [3: 1** 2*] **2185** [7: 4** 3*].=
P2	5115 [21: 15** 6*] 5044 [15: 10** 5*]	6899 [3: 2** 1*] 7815 [5: 3** 2*]
P3 EA	8954 [11: 5** 6*] 5611 [10: 2** 8*]	8680 [4: 3** 1*] **2185** [7: 7** 0*]
P3 TEAS	other three [7]	5611 [3]
P2 & P3		
All		

Both 7032 and 8311 attended for only one session, so this is clearly not responsible for whether high or low numbers of correlations are found between the HRV measures. On the other hand, 7032 had a right bundle branch block, so it is possible that those with some cardiovascular pathology, or with generally lower HRV (and perhaps those who are very healthy, with generally higher HRV), show more correlations than those who are at neither extreme.

For 2185, who participated in both P1 and P3, there is some agreement here, although for EA only.