Assessing causality in normal and impaired short-term cardiovascular regulation via nonlinear prediction methods

BY GIANDOMENICO NOLLO 1,*, LUCA FAES 1, RENZO ANTOLINI 1 AND ALBERTO PORTA 2

1 Department of Physics, University of Trento, 38050 Povo, Trento, Italy
2 Department of Technologies for Health, Galeazzi Orthopaedic Institute, University of Milan, 20161 Milan, Italy

We investigated the ability of mutual nonlinear prediction methods to assess causal interactions in short-term cardiovascular variability during normal and impaired conditions. Directional interactions between heart period (RR interval of the ECG) and systolic arterial pressure (SAP) short-term variability series were quantified as the cross-predictability (CP) of one series given the other, and as the predictability improvement (PI) yielded by the inclusion of samples of one series into the prediction of the other series. Nonlinear prediction was performed through global approximation (GA), approximation with locally constant models (LA0) and approximation with locally linear models (LA1) of the nonlinear function linking the samples of the two series, on patients with neurally mediated syncope and control subjects. Causality measures were evaluated in the two directions (from SAP to RR and from RR to SAP) in the supine (SU) position, in the upright position after head-up tilt (early tilt, ET) and after prolonged upright posture (late tilt, LT). While the trends for the GA, LA0 and LA1 methods were substantially superimposable, PI elicited better than CP the prevalence of causal coupling from RR to SAP during SU. Both CP and PI noted a marked decrease in coupling in both causal directions in syncope subjects during LT, documenting the impairment of cardiovascular regulation in the minutes just preceding syncope.

Keywords: global nonlinear prediction; k-nearest neighbours local nonlinear prediction; neurally mediated syncope; nonlinear Granger causality; heart rate and arterial pressure variability; out-of-sample prediction

1. Introduction

In the study of complex systems, it is not only important to detect interactions but also to identify driver–response relationships, and thus to assess causality. The issue of assessing and quantifying directional interactions from the analysis of bivariate time series is indeed receiving significant attention in the scientific
Applications to physiological systems are widespread, with particular emphasis on the characterization of causality among neural (Pereda et al. 2005) and cardiac (Nollo et al. 2002, 2005; Porta et al. 2002; Rosenblum et al. 2002; Faes & Nollo 2006; Faes et al. 2006, 2008a) signals. With regard to the cardiovascular application, evaluating causality in the short-term regulation of heart period and arterial pressure is crucial to disentangle coupling mechanisms. The heart period and the arterial pressure are indeed interdependent and continuously regulated by the concurrent action of the feedback baroreflex, a neural reflex producing variations of heart period in response to parallel changes of arterial pressure and of the mechanical feed-forward mechanisms, whereby heart period modifications induce changes in the cardiac output and consequently in the arterial blood pressure (Malpas 2002). This closed-loop beat-to-beat regulation is influenced by several pathologies, as different degrees of unbalancing of the mutual interaction between heart period and arterial pressure have been associated, for example, with myocardial infarction (Nollo et al. 2002), spinal cord injury (Legramante et al. 2001) and neurally mediated syncope (Faes et al. 2006). In particular, neurally mediated syncope is one of the most common types of syncope, and is defined as a transient loss of consciousness and postural tone consequent to a sudden decrease in blood pressure and heart rate (Brignole et al. 2001). Although the exact nature of the pathophysiological mechanisms that trigger the syncope events is still poorly understood (Mosqueda-Garcia et al. 2000), the involvement of the closed-loop cardiovascular regulation is widely acknowledged.

It is commonly accepted that a measure of causal interdependence of bivariate time series should quantify to what extent the variability of one series depends on the variability of the other. With implicit reference to this general notion, a common way to assess causal dependencies is to define a directional index of coupling and to exploit its asymmetry over the two causal directions to assess the directionality of coupling. Among the variety of approaches (e.g. exploiting state-space correspondence; Arnhold et al. 1999; Quiroga et al. 2002), phase synchronization (Rosenblum & Pikovsky 2001; Rosenblum et al. 2002) or information theory (Hlavackova-Schindler et al. 2007), the cross-prediction method quantifying the predictability of one of the two series starting from samples of the other series has been widely used to assess directional coupling in short and noisy physiological time series (Schiff et al. 1996; Le Van Quyen et al. 1999; Faes & Nollo 2006).

A specific definition of causality in bivariate time series is that originally proposed by Wiener (1956) and later formalized by Granger (1969) in the context of linear regression modelling of stochastic processes. According to this definition, the series $x$ ‘Granger causes’ the series $y$ when the past values of $x$ contain information that helps predict $y$ above and beyond the information contained in past values of $y$ alone. As a result, nonlinear Granger causality approaches quantify the predictability improvement (PI) by comparing the mixed prediction of a time series with its self-prediction (Ancona et al. 2004; Chen et al. 2004; Faes et al. 2008a).

In the present study, we exploit the nonlinear prediction approach set in Faes et al. (2008b) to assess causality in the short-term interactions between heart period and arterial pressure, during both normal and impaired cardiovascular regulation. The study considers and compares two approaches for assessing
causality, i.e. cross-prediction and PI, as well as two methods for performing mutual nonlinear prediction, i.e. global approximation (GA) and local approximation (LA) of the nonlinear function linking the samples of the two series. Neurally mediated syncope is considered to be a model of dynamic impairment of the cardiovascular regulatory mechanisms. Predictability indices were evaluated in a group of syncope patients and in a control group of healthy subjects during head-up tilt (HUT), in order to clarify whether these indices can be helpful in typifying the physiopathology of syncope and in providing clinical information.

2. Methods

(a) Nonlinear prediction of bivariate time series

Let \{X(n)\} and \{Y(n)\}, \(n=1, \ldots, N\), be two stationary time series simultaneously measured. Since the two series usually do not span the same range of values, they are constrained to zero mean and unitary variance, obtaining the normalized time series \{x(n)\} and \{y(n)\}, \(n=1, \ldots, N\). Nonlinear prediction methods (Casdagli 1989) perform the estimation of one of the two series approximating an assumed nonlinear function that maps a properly chosen combination of the present and past values of the two series into the current value of one of them. Specifically, to predict the series \(y\), the following functional nonlinear relationship is assumed:

\[
y(n) = f(z_{P+Q+1}(n)),
\]

\[
z_{P+Q+1}(n) = [y(n-1), \ldots, y(n-P), x(n), x(n-1), \ldots, x(n-Q)],
\]

where \(z_{P+Q+1}(n)\) is a delay vector containing \(P\) past samples of \(y\) as well as the current and \(Q\) past samples of \(x\). Nonlinear prediction methods derive an approximation \(\hat{f}\) of the nonlinear function in equation (2.1), allowing prediction of the current sample of \(y\), i.e. calculation of \(\hat{y}(n)\). Nonlinear prediction of \(x\) can be performed reversing the roles of the two series, to obtain the predicted series \(\hat{x}(n)\).

Different prediction schemes can be followed by selecting the samples of the two series to be included into the delay vectors \(z\) (Faes et al. 2008b). In particular, self-prediction, cross-prediction and mixed prediction of the series \(y\) are performed, respectively, using only samples taken from \(y\) \((P\geq 1, Q=-1\) in equation (2.1)), using only samples taken from \(x\) \((P=0, Q\geq 0)\) and using samples taken from both \(x\) and \(y\) \((P\geq 1, Q\geq 0)\). Besides the adopted prediction scheme, two main approaches to nonlinear prediction can be followed depending on the method used for approximating the nonlinear function \(f\) in equation (2.1). Global nonlinear prediction makes use of all the available data from the two series to derive a GA of \(f\) that is valid for the whole mixed state space. On the contrary, local nonlinear prediction uses only a part of the data, ad hoc selected in the neighbourhood of the current state \(z_{P+Q+1}(n)\), and performs an LA of \(f\) that holds only in a limited region of the mixed state space. The GA approach performs an explicit representation of \(\hat{f}\) in which nonlinear relationships are assumed and weighed by a set of coefficients to be estimated from the data. The most used LA approach evaluates \(\hat{f}\) from the \(k\)-Euclidean nearest neighbours of the current state \(z_{P+Q+1}(n)\).
implementing either locally constant models or locally linear models that perform, respectively, a zeroth order LA (LA0) and a first-order LA (LA1) of \( f \). Nonlinear prediction methods based on GA, LA0 and LA1 have been formalized in the past (Farmer & Sidorowich 1987; Korenberg 1989; Sugihara & May 1990), and recently developed to perform self-prediction, cross-prediction and mixed prediction over short and noisy bivariate time series (Porta et al. 2007; Wang et al. 2007; Faes et al. 2008a,b). We refer to appendix A for a detailed description of the algorithms implemented in this study to perform linear and nonlinear prediction.

(b) Evaluation of prediction

The degree of predictability of the series \( y \) can be quantified by taking the squared correlation between its original values, \( y(n) \), and predicted values, \( \hat{y}(n) \), yielding the index \( r^2_y \) (Porta et al. 2007). This index ranges between 0 and 1, indicating, respectively, full unpredictability and full predictability of \( y \). The same procedure can be used to estimate the predictability of \( x \), calculating the squared correlation \( r^2_x \). Moreover, since the estimated degree of predictability depends on the adopted prediction scheme, in the following we will indicate with \( (SP_x, SP_y) \), \( (CP_x, CP_y) \) and \( (MP_x, MP_y) \) the squared correlation estimated for the series \( x \) and \( y \) after performing self-prediction, cross-prediction and mixed prediction, respectively. In this study, evaluation of prediction was performed according to \( M \)-fold cross-validation (see appendix A for details). Cross-validation was chosen since it prevents overfitting while allowing the exploitation of the maximum amount of available data for the training of the predictor.

(c) Regularity and Granger causality evaluated from predictability measures

When cross-validation is implemented, the predictability does not increase indefinitely as a function of the vector length. Hence, measures of predictability can be defined as the maximum of the squared correlation obtained varying within a prescribed range the number \( P \) and \( Q \) of samples used for prediction. Accordingly, the time series are described with the embedding yielding the maximum predictability (Porta et al. 2007). This choice prevents the use of large values for \( P \) and \( Q \), which in short time series could be associated with an ill-conditioned model identification. When a single series is predicted only from its past values through a self-prediction approach, the maximum self-predictability (SP) is taken as a measure of regularity (Porta et al. 2007)

\[
SP_y = \max_{P=1, \ldots, P_{\text{max}}} SP_y(P), \quad SP_x = \max_{P=1, \ldots, P_{\text{max}}} SP_x(P). \tag{2.2}
\]

When a series is predicted also using the samples of the other series through cross- or mixed-prediction approaches, directional measures of interdependence between the two series can be derived (Faes et al. 2008b). Specifically, the cross-prediction approach can be exploited to measure the causal coupling from one series to the other as the maximum cross-predictability (CP; Schiff et al. 1996; Faes & Nollo 2006)

\[
CP_{x \rightarrow y} = \max_{Q=0, \ldots, Q_{\text{max}}} CP_y(Q), \quad CP_{y \rightarrow x} = \max_{P=0} CP_x(Q). \tag{2.3}
\]
The PI approach measures the causal coupling from one series to the other as the maximum normalized increase in predictability yielded by mixed prediction compared with self-prediction (Ancona et al. 2004; Chen et al. 2004; Faes et al. 2008a)

\[
\begin{align*}
\text{PI}_{x \rightarrow y} &= \max_{P=1,\ldots,P_{\text{max}}, \quad Q=0,\ldots,Q_{\text{max}}} \frac{\text{MP}_y(P, Q) - \text{SP}_y(P)}{1 - \text{SP}_y(P)}, \\
\text{PI}_{y \rightarrow x} &= \max_{P=1,\ldots,P_{\text{max}}, \quad Q=0,\ldots,Q_{\text{max}}} \frac{\text{MP}_x(P, Q) - \text{SP}_x(P)}{1 - \text{SP}_x(P)}.
\end{align*}
\]

(2.4)

3. Experimental protocol and data analysis

The study population consisted of nine patients (38 ± 19 years old) with neurally mediated syncope as established by the positive response to prolonged HUT test (progressive hypotension accompanied by reflex bradycardia leading to faint). Ten healthy volunteers with no previous history of syncope, presyncope or cardiovascular diseases, age matched with the syncope patients (31 ± 12 years old, \(p > 0.05\) versus syncope) were considered to be the control group. None of the control subjects fainted during the test. HUT was performed in a maximally controlled environment, after at least 5 hours of fasting. After 15 min of rest in the supine (SU) position, the table was tilted to the 60° position for a maximum of 30 min. If syncope occurred, the table was lowered to the SU position and the test was labelled as positive. The experimental protocol was approved by the Ethical Committee of the S. Chiara Hospital, Trento, Italy, where the research was conducted. All syncope patients and control subjects provided written informed consent.

The electrocardiogram (ECG) (lead II) and non-invasive arterial pressure measured at a finger level (Finapres) were continuously monitored during the test and acquired with a 1 kHz sampling rate and 12 bit precision. The variability series of the RR interval and the systolic arterial pressure (SAP) were offline measured on a beat-by-beat basis as the temporal distance between two consecutive R peaks in the ECG recognized through a template-matching algorithm (Nollo et al. 1992) and the maximum arterial pressure value within each detected heart beat, respectively. Three windows corresponding to peculiar epochs of the test, each lasting \(N=300\) samples, were considered: one during the last 5 min of resting in the SU position and two during tilt, at approximately 2 min (early tilt: ET) and approximately 15 min (late tilt: LT) from the onset of the tilting manoeuvre. In syncope patients, the LT window was chosen immediately before the decrease in SAP associated with the syncopal event (15 ± 8 min). The selected windows were cleaned up from artefacts and detrended by a zero-phase IIR high-pass filter.

After measuring mean and variance for the time-domain characterization, each RR and SAP series was normalized by subtracting the mean and then dividing it by the standard deviation, thus obtaining the dimensionless series \(r(n)\) and \(s(n), \ n=1, \ldots, 300\). Nonlinear prediction was repeated varying the numbers \(P\) and \(Q\) of samples used for prediction, respectively, in the ranges \(\{1, \ldots, 12\}\) and \(\{0, \ldots, 12\}\) for both GA and LA approaches. Local nonlinear
prediction was performed using a number of neighbour vectors \( k = 1.5 \times (P + Q + 2) \) for the LA0 approach (Kanters et al. 1994) and \( k = 50 \) for the LA1 approach (Porta et al. 2007). To provide a comparison between linear and nonlinear approaches, we also performed global linear prediction, discarding nonlinear model terms in the GA approach and broadening to the maximum size the neighbourhood in the LA1 approach. Evaluation of the predictors was carried out according to 10-fold cross-validation (\( M = 10 \) segments; test set size: \( N/M = 30 \) samples). The analysis was performed calculating, for GA, LA0 and LA1 approaches, the SP of detrended and normalized RR and SAP series (indices \( \text{SP}_r \) and \( \text{SP}_s \), equation (2.2)), and the causal coupling along the two pathways of the RR–SAP loop measured either as CP (indices \( \text{CP}_{s \rightarrow r} \) and \( \text{CP}_{r \rightarrow s} \), equation (2.3)) or PI (indices \( \text{PI}_{s \rightarrow r} \) and \( \text{PI}_{r \rightarrow s} \), equation (2.4)).

Statistical analysis was performed in order to assess the significance of: (i) differences among indices during the three epochs of the considered experimental protocol, (ii) differences among the three prediction methods in evaluating SP, and (iii) the differences between the causal couplings evaluated in the two directions. While analysis (iii) performed a single comparison (\( \text{CP}_{s \rightarrow r} \) versus \( \text{CP}_{r \rightarrow s} \) or \( \text{PI}_{s \rightarrow r} \) versus \( \text{PI}_{r \rightarrow s} \)), analyses (i) and (ii) were conducted according to a multiple testing procedure, performing the comparisons (SU versus ET; ET versus LT) and (GA versus LA0; GA versus LA1; LA0 versus LA1), respectively. Multiple testing was carried out through the Holm method (Holm 1979), a sequential Bonferroni approach that is more powerful and equally protective against type I errors than the classical Bonferroni correction. In all analyses, comparisons were performed by the Wilcoxon signed-rank test; non-parametric statistics was used to account for the non-Gaussian distribution of the indices over subjects. Statistical significance was assessed with a nominal significance level \( \alpha = 0.05 \).

4. Results

Representative examples of the variability series of the RR interval and SAP selected during the three epochs for a healthy control subject and a syncope patient are shown in figure 1. Trends of the time-domain parameters are summarized in Table 1. Both control subjects and syncope patients showed the typical response to HUT, documented by the maintenance of the mean SAP values accompanied by the tachycardia (significant decrease in the mean RR) induced by an assumption of the upright position. The two groups showed an increase in the SAP variance from SU to ET, statistically significant for the control subjects, and a decrease in the mean RR from ET to LT, statistically significant for the syncope patients. Moreover, patients exhibited, as a forewarn of syncope, a significant decrease in the mean SAP moving from ET to LT.

SP, CP and PI derived from the series depicted in figure 1 are shown in figure 2. The SP of both the RR and SAP series was always lower when calculated using the LA0 approach than using GA or LA1 approaches. Regardless of the approach used, in both subjects, the predictability of the RR interval increased from SU to ET, and remained high moving to LT. The predictability of SAP increased from SU to ET, particularly in the syncope patient, and decreased during LT. In the control subject, the CP increased.
markedly going from SU to ET and to LT in both causal directions. In the syncope patient, the increase in CP with tilt was followed by an abrupt decrease during the epoch immediately preceding syncope. The PI was almost constant in the three epochs for the healthy control, while it dropped to low values during the LT epoch for the syncope patient.

Table 1. Time-domain characterization of RR and SAP variability series during tilt test. (Values are means ± standard deviation (s.d.) over 10 control subjects and nine syncope patients. SU, supine; ET, early tilt; LT, late tilt. Statistical analysis: *significance of ET versus SU or LT versus ET.)

<table>
<thead>
<tr>
<th>control subjects</th>
<th>syncope patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>SU</td>
<td>ET</td>
</tr>
<tr>
<td>mean RR (ms)</td>
<td>831 ± 78</td>
</tr>
<tr>
<td>s.d. RR (ms)</td>
<td>39 ± 21</td>
</tr>
<tr>
<td>mean SAP (mm Hg)</td>
<td>119.1 ± 17.8</td>
</tr>
<tr>
<td>s.d. SAP (mm Hg)</td>
<td>4.3 ± 1.4</td>
</tr>
</tbody>
</table>

Figure 1. Representative examples of (i–iii) RR and (iv–vi) SAP series measured during the three time windows selected for the analysis (SU, supine (i) and (iv)); ET, early tilt ((ii) and (v)); LT, late tilt (iii) and (vi)) for (a) a control subject and (b) a syncope patient.

Phil. Trans. R. Soc. A (2009)
The trends described in figure 2 were found to be statistically significant by the analysis extended to the whole groups, as summarized in figures 3–5. Figure 3 shows that the SP was always significantly lower when evaluated by the LA0 approach, while GA and LA1 approaches yielded comparable predictability. In both control subjects and syncope patients, the RR interval predictability increased significantly from SU to ET, and remained unchanged during LT, independently of the nonlinear prediction approach. The SAP predictability was substantially unchanged during the three epochs, except for a reduction from ET to LT documented in syncope patients using LA0 nonlinear prediction.

As shown in figure 4, the three nonlinear prediction approaches yielded similar results in terms of CP calculated over the two causal directions. In control subjects, both the directional coupling measures $\text{CP}_{s\rightarrow r}$ and $\text{CP}_{r\rightarrow s}$ increased progressively with time, and the increase was marked and statistically significant moving from SU to ET. In syncope patients, the two measures increased from SU to ET and decreased from ET to LT, with variations that were statistically significant in the causal direction from SAP to RR interval and less evident in the opposite direction. In both groups and in all epochs, no significant differences were observed between $\text{CP}_{s\rightarrow r}$ and $\text{CP}_{r\rightarrow s}$.

Figure 2. (i) Self-predictability (SP), (ii) cross-predictability (CP) and (iii) predictability improvement (PI) calculated over the series depicted in figure 1. (a) Control patient and (b) syncope patient. The indices are calculated for the RR (open symbols) and the SAP (filled symbols) series using the GA approach (circles), the LA approach with locally constant models (LA0, triangles) and the LA approach with locally linear models (LA1, squares).
In contrast to CP (figure 4), PI clearly showed differences between the two causal directions (figure 5). Specifically, PI during the SU epoch was found to be significantly lower in the direction from SAP to RR interval (i.e. $PI_{s\rightarrow r}$) than that in the opposite direction (i.e. $PI_{r\rightarrow s}$) independently of the nonlinear prediction approach. This unbalancing, present in both healthy controls and syncope patients, was lost during ET. Moving to the LT epoch, no significant modifications of both $PI_{s\rightarrow r}$ and $PI_{r\rightarrow s}$ indices dropped to low values.

When linear prediction approaches were considered, predictability of RR and SAP series was always slightly lower than that obtained by nonlinear prediction, independently of the prediction method and the physiological condition. Linear prediction revealed similar variations of predictability and causality indices to those observed for nonlinear prediction. However, in syncope patients, the increase in $SP_r$ and of $CP_{s\rightarrow r}$ moving from SU to ET and the decrease in both $CP_{s\rightarrow r}$ and $PI_{s\rightarrow r}$ moving from ET to LT were not statistically significant when calculated by linear prediction.

Figure 3. Box plot (line, median; box, quartiles; whiskers, 10th and 90th percentiles) of the SP of (i) RR ($SP_r$) and (ii) SAP ($SP_s$) series evaluated in (a) control subjects and (b) syncope patients during the three selected epochs (SU, supine; ET, early tilt; LT, late tilt) using the GA approach (white boxes), the LA approach with locally constant models (LA0; grey boxes) and the LA approach with locally linear models (LA1; black boxes). Statistical analysis: *significance of both LA0 versus GA and LA0 versus LA1; # significance of ET versus SU or LT versus ET.

In contrast to CP (figure 4), PI clearly showed differences between the two causal directions (figure 5). Specifically, PI during the SU epoch was found to be significantly lower in the direction from SAP to RR interval (i.e. $PI_{s\rightarrow r}$) than that in the opposite direction (i.e. $PI_{r\rightarrow s}$) independently of the nonlinear prediction approach. This unbalancing, present in both healthy controls and syncope patients, was lost during ET. Moving to the LT epoch, no significant modifications of both $PI_{s\rightarrow r}$ and $PI_{r\rightarrow s}$ were found in the control subjects, while in syncope patients both $PI_{s\rightarrow r}$ and $PI_{r\rightarrow s}$ indices dropped to low values.

When linear prediction approaches were considered, predictability of RR and SAP series was always slightly lower than that obtained by nonlinear prediction, independently of the prediction method and the physiological condition. Linear prediction revealed similar variations of predictability and causality indices to those observed for nonlinear prediction. However, in syncope patients, the increase in $SP_r$ and of $CP_{s\rightarrow r}$ moving from SU to ET and the decrease in both $CP_{s\rightarrow r}$ and $PI_{s\rightarrow r}$ moving from ET to LT were not statistically significant when calculated by linear prediction.
The vector lengths, optimized according to cross-validation, were not modified substantially while varying output series, group and experimental condition. The average values obtained using GA, LA0 and LA1 approximations were, respectively, $P=6, 4, 3$ for self-prediction ($Q=-1$); $Q=7, 6, 5$ for cross-prediction ($P=0$); and $P=2, 3, 1$ and $Q=7, 3, 2$ for mixed prediction. Considering the GA approach in which linear and nonlinear effects were separately modelled (see appendix A), the number of nonlinear model terms (self-prediction, $P_2=3$; cross-prediction, $Q_2=2$; mixed prediction, $P_2+Q_2=3$) was significantly lower than the number of linear terms ($P=6, Q=7, P+Q=8$).

5. Discussion

This study applied nonlinear prediction methods, which are suitable to deal with short and noisy physiological time series (Sugihara 1994; Porta et al. 2007; Faes et al. 2008b), to the characterization of changes in RR and SAP complexity and causality induced by HUT in healthy subjects and subjects prone to neurally mediated syncope. HUT is routinely used to determine the neurocardiogenic aetiology of syncope (Kenny et al. 1986). In humans, the transition from SU to
upright position, after a first phase of full arterial pressure compensation, is followed by a progressive tachycardia characterized by rhythmic oscillations of RR and SAP values at the frequency of 0.1 Hz (Montano et al. 1994). In the search of orthostatic intolerance, this test is prolonged so that, in patients who experienced neurally mediated syncope, the mechanisms leading to faint could be activated and studied. These mechanisms include a progressive SAP decrease associated with changes in the balance between the two main mechanisms determining cardiovascular homeostasis, i.e. the baroreflex feedback and the mechanical feed-forward (Furlan et al. 1998; Faes et al. 2006). HUT is therefore a unique model of dynamic changes of cardiovascular regulation.

(a) **Comparison among different methods to perform nonlinear prediction of cardiovascular series**

We considered two conceptual approaches to perform nonlinear prediction: the GA assumes *a priori* the type of nonlinearity that can be present in the data (Faes et al. 2008a), while the LA approximates an unspecified nonlinearity through locally constant (LA0 method; Sugihara & May 1990) or locally linear (LA1 methods).
method; Farmer & Sidorowich 1987) models. We found that the LA0 approach always provides lower predictability degrees than the GA and LA1 approaches. This suggests that the locally constant approximation is less effective in describing the relationship present within and between short-term cardiovascular dynamics. On the contrary, comparable predictability degrees were found modelling second-order nonlinearities through a GA method, or approximating presumed nonlinearities with locally linear functions. Notwithstanding the differences among approaches, the trends of the measures of complexity and directional coupling during the different phases of the experimental protocol were very similar for all the three nonlinear prediction methods, both in the control subjects and in the syncope patients. Hence, nonlinear prediction appears to be a robust technique to characterize short-term cardiovascular regulation, regardless of the specific technique adopted to approximate nonlinear relationships.

According to the prediction approach set in Faes et al. (2008b), causal interactions between the RR interval and SAP series were assessed by the CP and PI indices. The two measures in most cases showed similar trends in the values of directional coupling, such as their increase in the direction from SAP to RR interval with the assumption of the upright position or their decrement in both directions just before syncope. Nevertheless, some differences distinguished their capability to evaluate causality in the cardiovascular interactions. While the directional coupling given by CP was inclined to be similar when evaluated over the two causal directions, the PI in some cases showed different behaviours. As an example, with the transition from the SU to the upright position the CP increased markedly in both causal directions (figure 4), while the PI showed a tendency to increase from SAP to RR and to decrease from RR to SAP (figure 5). Furthermore, the unbalancing of the cardiovascular regulation with prevalence of the mechanical influences from RR to SAP typical of the SU position (see §5b) was evident only measuring the PI in the two directions (figure 5), while CP asymmetries were not statistically significant (figure 4). These findings support previous validation studies showing that CP, similar to other directional coupling measures based on state-space correspondence, is a good marker of the coupling degree between two variables (Faes et al. 2008b) but might be unable to detect the prevailing direction of interactions (Quiroga et al. 2002). On the contrary, PI calculated as in Chen et al. (2004) and Faes et al. (2008b) was able to detect a driver–response relationship in the analysed short-term cardiovascular interactions.

Our results suggest that nonlinear prediction is more appropriate than linear prediction to describe the cardiovascular dynamics and dynamical interaction, particularly in impaired conditions. Although with a reduced number of model terms with respect to linear prediction, the contribution of nonlinear modelling was determinant in describing causal cardiovascular interactions in syncope patients. In these patients, the increase in the causal coupling from SAP to RR with tilt, as well as its consequent decrease before syncope, was clearly detected by all nonlinear prediction approaches, while these trends appeared to be blunted and not statistically significant using linear prediction. Thus, a nonlinear approach seems recommended to explain physiological behaviours, such as the tilt-induced sympathetic activation and the damaged capability of the heart to compensate for arterial pressure lowering in the minutes just preceding syncope.
(b) **Mechanisms of causal interactions between cardiovascular variability series in normal and impaired conditions**

The proposed nonlinear prediction methods were able to show the different contributions of the baroreflex feedback and mechanical feed-forward mechanisms to cardiovascular regulation in patients and healthy subjects. One major finding was that cardiovascular regulation in the SU position was unbalanced, with a prevalence of the feed-forward effects from RR to SAP, mainly of a mechanical nature (Mullen *et al.* 1997), over the feedback influences from SAP to RR, commonly ascribed to the neural baroreflex mechanisms (Malpas 2002). This finding agrees with the previous results on the relevance of the feed-forward mechanical response of SAP variability to changes of RR, and confirms the need for disentangling it from the baroreflex feedback mechanism in the study of normal and pathological cardiovascular regulatory systems (Saul *et al.* 1991; Legramante *et al.* 2001; Nollo *et al.* 2002, 2005; Porta *et al.* 2002).

Moreover, the causal coupling analysis documented an enhancement of the interactions between the two cardiovascular variables just after the tilt transition, which was mainly sustained by the increased capability of SAP to drive RR fluctuations (i.e. the baroreflex regulation) owing to the increase in sympathetic tone (Montano *et al.* 1994; Nollo *et al.* 2005). In healthy subjects, this protective mechanism was effective throughout the test, as demonstrated by the tendency towards a further increase in the coupling measured in the two causal directions (particularly the one from RR interval to SAP). On the contrary, in syncope patients, a disappearance of the cardiovascular interactions was noted before syncope, as documented by the dramatic decrease in CP and PI in the minutes preceding the fainting event (figures 4 and 5). This decrease was the main marker of cardiovascular dysregulation preceding syncope, since the overall variability (table 1) and the predictability of the two series (figure 3) remained almost constant with respect to the early period of tilt. Thus, the lack of directional coupling evidenced by mutual nonlinear prediction seems able to forewarn of the failure of cardiovascular regulation prior to the appearance of bradycardia and the fall in blood pressure that lead to fainting. In a previous work conducted on the same dataset using a causal coherence approach, abolishment of baroreflex regulation from SAP to RR prior to fainting was observed along with a maintained feed-forward link from RR to SAP (Faes *et al.* 2006). The present results suggest a complete weakening of cardiovascular regulation, and indicate an involvement of the mechanical feed-forward coupling in the aetiology of neurally mediated syncope. The partial inconsistency between these two findings could be due to the nonlinear nature of the interactions studied with the approaches proposed here, and by the fact that these approaches account for a global response of the system, while causal coherence is a linear tool and focuses analysis on a specific frequency band (low frequency, approx. 0.1 Hz, in Faes *et al.* (2006)).

(c) **Complexity analysis of single cardiovascular series in normal and impaired conditions**

Independently of the adopted nonlinear prediction method, the SP indices were able to follow the changes in cardiovascular control provoked by HUT in healthy subjects and syncope patients. Consistent with previous studies
(Porta et al. 2000, 2007; Faes et al. 2008a), the predictability of the RR interval increased significantly in both groups moving from SU to an upright position. This finding reflects an increase in regularity of the series that can be ascribed to the sympathetic activation consequent to HUT, triggering the heart rate oscillations of approximately 0.1 Hz (Montano et al. 1994). On the contrary, the predictability of the SAP series is high even in the SU position, indicating that the presence of other oscillations besides that at 0.1 Hz (e.g. the one at the respiratory frequency) does not imply an increased complexity, and remains high after HUT when there is mostly the 0.1 Hz oscillation (Porta et al. 2000).

In the control group, the maintenance of the upright position did not give rise to significant variations in the SP of the two cardiovascular variability series, thus reflecting the preservation of the mechanisms of cardiovascular regulation during prolonged HUT. In syncope patients, the complexity of the RR interval was maintained in the period just preceding the syncope event, while a tendency to lower predictability (higher complexity) was observed for the SAP series. This result indicates that the degeneration of cardiovascular control in the minutes preceding syncope is also manifested, even though to a lesser extent with respect to the loss of RR–SAP interactions documented by the coupling analysis, through an increased complexity of SAP variability.

6. Conclusions

The present study shows nonlinear prediction methods to be helpful, non-invasive tools to assess directional coupling mechanisms from bivariate analysis of the short-term spontaneous fluctuations of the heart rate and arterial pressure during normal and diseased conditions. While the majority of short-term cardiovascular variability studies deal with monovariate analysis or non-directional coupling analysis, we show that the inference of causality in cardiovascular interactions provides valuable additional information. On one hand, our results confirm already known physiological behaviours, such as the imbalance of cardiovascular regulation with the prevalence of mechanical feed-forward mechanisms (causality from heart period to arterial pressure) in the SU position and increased feedback activity (causality from arterial pressure to heart period) in the upright position. On the other hand, we point out the usefulness of the proposed causality measures also to characterize alterations in cardiovascular regulation in pathological subjects. In this study, the CP and PI indices calculated through mutual nonlinear prediction were markedly reduced in the minutes preceding neurally mediated syncope. The significance of this result is emphasized by the fact that other common measures, such as standard time-domain quantities, monovariate indices of regularity and traditional linear coupling measures, calculated either in the present or in the previous studies, did not show comparable modifications during HUT testing. Thus, methods based on nonlinear prediction should be useful to complete the understanding of the pathophysiological mechanisms involved in the genesis of neurally mediated syncope. Furthermore, the corresponding causality measures might constitute powerful markers to be included in the medical strategies developed to forewarn of syncope episodes.
Appendix A

The global nonlinear prediction method is based on a GA of the nonlinear function $f$ of equation (2.1) according to the polynomial representation (Wang et al. 2007; Faes et al. 2008a)

$$y(n) = \sum_{i=1}^{P} a_{1}(i)y(n-i) + \sum_{j=0}^{Q} b_{1}(j)x(n-j) + \sum_{i=1}^{P_{2}} \sum_{j=i}^{P_{2}} a_{2}(i,j)y(n-i)y(n-j)$$

$$+ \sum_{i=0}^{Q_{2}} \sum_{j=i}^{Q_{2}} b_{2}(i,j)x(n-i)x(n-j) + \sum_{i=1}^{P_{2}} \sum_{j=0}^{Q_{2}} c_{2}(i,j)y(n-i)x(n-j) + e_{y}(n),$$

(A 1)

where $e_{y}(n)$ is the prediction error. The vector parameters $a_{1} = \{a_{1}(i), i=1, ..., P\}$ and $b_{1} = \{b_{1}(j), j=0, ..., Q\}$ represent linear (first-order) contributions to $y(n)$, while the matrix parameters $a_{2} = \{a_{2}(i,j), i, j=1, ..., P_{2}\}$, $b_{2} = \{b_{2}(i,j), i, j=0, ..., Q_{2}\}$ and $c_{2} = \{c_{2}(i,j), i=1, ..., P_{2}, j=0, ..., Q_{2}\}$ represent nonlinear (second-order) contributions. Thus, the model orders $P$, $Q$ and $P_{2}, Q_{2}$ in equation (A 1) represent the maximum lags for linear and nonlinear influences, respectively. Nonlinear prediction was performed letting both $P_{2} \geq 1$ and $Q_{2} \geq 0$, while linear prediction was performed imposing $P_{2} = 0$ and $Q_{2} = -1$. Note that the representation of nonlinear effects can be extended by including in equation (A 1) higher order terms. However, for the present application extension to higher than quadratic nonlinearities seems unnecessary, since it has been shown that cardiovascular dynamics can be characterized by nonlinearity lower than third order (Wang et al. 2007; Faes et al. 2008a). The model parameters were identified from the available realization of the two series, after representation in the form $y = H_{y}a_{1} + H_{x}b_{1} + H_{yy}a_{2} + H_{xx}b_{2} + H_{xy}c_{2} + e_{y}$, where $y = [y(1), ..., y(N)]^{T}$ and $e_{y} = [e_{y}(1), ..., e_{y}(N)]^{T}$ are the vectors of the output and the error series; $H_{y}$ and $H_{x}$ are the matrices of the linear self- and cross-terms; and $H_{yy}$, $H_{xx}$ and $H_{xy}$ are the matrices of the quadratic self- and cross-terms (Wang et al. 2007; Faes et al. 2008a). All these matrices constitute the observation matrix $H = [H_{y} H_{x} H_{yy} H_{xx} H_{xy}]$ such that $y = H \cdot d + e_{y}$ with $d = [a_{1}^{T} b_{1}^{T} a_{2}^{T} b_{2}^{T} c_{2}^{T}]^{T}$. The vector $d$ of the model parameters was estimated through traditional least-squares optimization, i.e. $d = (H^{T}H)^{-1}H^{T}y$. Finally, the estimated model parameters were used to predict the output series, i.e. $\hat{y} = H \cdot \hat{d}$.

Local nonlinear prediction performs an LA of the nonlinear map $f$ in equation (2.1), predicting $y(n)$ from only those vectors that are similar to $z_{p+Q+1}(n)$ in the mixed state space. In this study, we performed the $k$-nearest neighbour nonlinear prediction (Farmer & Sidorowich 1987), an approach that evaluates similarity through selection of the $k$ vectors $z_{p+Q+1}(n_{1}), ..., z_{p+Q+1}(n_{k})$ with the lowest Euclidean distance to $z_{p+Q+1}(n)$. To relate the $k$-selected nearest neighbours to their corresponding output samples $y(n_{1}), ..., y(n_{k})$, we implemented local models performing either a zeroth-order LA (LA0; Sugihara & May 1990) or a first-order LA (LA1; Faes & Nollo 2006) of the nonlinear function $f$. LA0 nonlinear prediction was performed estimating the current
sample of \( y \) as a weighted mean of the output samples related to the \( k \)-nearest neighbours: 
\[
\hat{y}(n) = w \cdot y_n, \quad \text{where} \quad y_n = [y(n_1), \ldots, y(n_k)]^T \quad \text{and} \quad w = [w(1), \ldots, w(k)];
\]
the weight \( w(l) \) was related to the inverse of the Euclidean distance between \( z_{p+q+1}(n) \) and \( z_{p+q+1}(n_l) \), \( l=1, \ldots, k \). LA1 nonlinear prediction was performed setting a system of \( k \) linear equations linking each neighbour delay vector \( z_{p+q+1}(n_l) \) to its corresponding output sample \( y(n_l) \): 
\[
y(n_l) = c \cdot z_{p+q+1}(n_l), \quad l=1, \ldots, k.
\]
The linear system was solved by least-squares optimization and the estimated parameter vector \( \hat{c} \) was then used to predict the current sample of \( y \):
\[
\hat{y}(n) = \hat{c} \cdot z_{p+q+1}(n).
\]
Note that, in the LA1 approach, using the maximum \( k \) (i.e. including all vectors into the linear system) corresponds to performing linear prediction, as in this case a single global model is defined and a global linear function \( f \) is approximated.

Both global and local nonlinear predictions were performed through \( M \)-fold cross-validation \((M=10)\). According to this approach, the two time series \( x \) and \( y \) were partitioned in \( M \) consecutive segments, each of size \( N/M \), and prediction was performed \( M \) times. At the \( i \)th iteration \((i=1, \ldots, M)\), the samples of the \( i \)th segment \((x(m) \text{ and } y(m)) \) with \( m=(i-1)N/M+1, \ldots, iN/M \) were kept apart as a test set, and the remaining samples were used as a training set; the training set was used to define the predictor, i.e. to estimate \( \hat{f} \) according to the GA, LA0 or LA1 approaches, while the test set was used to obtain the predicted values of the two series within the \( i \)th segment.

References


Causality in cardiovascular interactions


Phil. Trans. R. Soc. A (2009)


