Cardiovascular and cardiorespiratory coupling analyses: a review

Steffen Schulz¹, Felix-Constantin Adochiei², Ioana-Raluca Edu³, Rico Schroeder¹, Hariton Costin⁴, Karl-Jürgen Bär⁵ and Andreas Voss¹

¹Department of Medical Engineering and Biotechnology, University of Applied Sciences Jena, Jena, Germany
²Faculty of Electronics, Telecommunications and Information Technology, Gh. Asachi Technical University of Iasi, Iasi, Romania
³Faculty of Electrical Engineering, Politechnica University of Bucharest, Bucharest, Romania
⁴Faculty of Medical Bioengineering, Gr. T. Popa University of Medicine and Pharmacy, Iasi, Romania
⁵Department of Psychiatry and Psychotherapy, University Hospital Jena, Jena, Germany

Recently, methods have been developed to analyse couplings in dynamic systems. In the field of medical analysis of complex cardiovascular and cardiorespiratory systems, there is growing interest in how insights may be gained into the interaction between regulatory mechanisms in healthy and diseased persons. The couplings within and between these systems can be linear or nonlinear. However, the complex mechanisms involved in cardiovascular and cardiorespiratory regulation very likely interact with each other in a nonlinear way. Recent advances in nonlinear dynamics and information theory have allowed the multivariate study of information transfer between time series. They therefore might be able to provide additional diagnostic and prognostic information in medicine and might, in particular, be able to complement traditional linear coupling analysis techniques. In this review, we describe the approaches (Granger causality, nonlinear prediction, entropy, symbolization, phase synchronization) most commonly applied to detect direct and indirect couplings between time series, especially focusing on nonlinear approaches. We will discuss their capacity to quantify direct and indirect couplings and the...
direction (driver–response relationship) of the considered interaction between different biological time series. We also give their basic theoretical background, their basic requirements for application, their main features and demonstrate their usefulness in different applications in the field of cardiovascular and cardiorespiratory coupling analyses.

1. Introduction

The analysis of causal and non-causal relationships within and between dynamic systems has become more and more a topic of great interest in different fields of science, for example, economics, neuroscience, physics or physiology. Especially in the medical field, the understanding of driver–response relationships between regulatory systems and within subsystems is of growing interest. In particular, the detection and quantification of the strength and direction of couplings are two major aspects of investigations for a more detailed understanding of physiological regulatory mechanisms. The cardiovascular and cardiorespiratory systems are characterized by a complex interplay of several linear and nonlinear subsystems [1]. Interactions of these physiological subsystems within the cardiovascular system can be described as closed loops with feed-forward (FF) and feed-back (FB) mechanisms. On the one hand, blood pressure changes detected by baroreceptors lead to changes in heart rate (HR) regulation through the arterial baroreflex control loop, and on the other hand, heart rate variations (HRVs) affect blood pressure via the Windkessel function [2]. Interactions within the cardiorespiratory system are mainly reflected in the respiratory sinus arrhythmia (RSA), the rhythmic fluctuation of cardiac cycle intervals (RR interval) in relation to respiration. RSA is commonly described as an alteration between inspiratory HR acceleration and expiratory HR deceleration under normal physiological conditions [3]. Under these circumstances, two major mechanisms are discussed: the central influence of respiration on vagal cardiac motoneurons and the impact of respiration on intrathoracic pressure and stroke [3–6]. For the analyses of the cardiovascular and cardiorespiratory regulatory systems, as well as the quantification of their interactions, a variety of methods have been proposed. Commonly applied linear approaches include cross-correlation analysis in the time domain and cross-spectral power density or coherence analysis in the frequency domain, both of which are used to investigate the interrelationships between two time series. However, linear approaches might be insufficient to quantify nonlinear structures and the complexity of physiological (sub)systems. Therefore, approaches from nonlinear time-series analysis seem to be more suited to capture complex interactions between time series. These approaches are partly based on the notion of Granger causality (GC), implying that if one time series has a causal influence on a second time series, then the knowledge of the past of the first time series is useful to predict future values of the second time series [7]. In biomedical applications, evaluation of causality is commonly performed by looking for directional dependencies within a set of multiple time series measured in the physiological system under investigation [8]. Here, causality is defined in terms of predictability and uses the directionality of time to determine a causal ordering of dependent time series, incorporating both direct and indirect causal influences from one process to another. This definition can be applied to bivariate (two time series) and multivariate (more than two time series) analysis. In the case of multivariate analysis, it is possible to differentiate between direct coupling (from one time series to another) and indirect coupling (effects mediated by one or more other time series). Direct coupling between two time series \( x_1 \) and \( x_2 \) exists if \( x_1 \rightarrow x_2 \) or \( x_2 \rightarrow x_1 \) (figure 1a), whereas indirect coupling occurs when two time series \( x_1 \) and \( x_2 \) cause a third common time series \( x_3 \) mediated by one of the two time series and also when both time series were caused by a third common time series \( x'_3 \) mediated by one of the two time series \( x_1 \) or \( x_2 \) (direct coupling: \( x_1 \rightarrow x_2; \ x_2 \leftrightarrow x_3; \ indirect \ coupling; \ x_1 \rightarrow x_3 \ mediated \ by \ x_2 \); figure 1b). The coupling definitions generalize the causality definitions by accounting for both forward and backward interactions [9].
In this review, we describe the approaches (Granger causality, nonlinear prediction, entropy, symbolization, phase synchronization) most commonly applied to detect direct and indirect couplings between time series, especially focusing on nonlinear approaches. We then discuss their usefulness for detecting linear and nonlinear interdependencies and summarize their theoretical background as well as basic requirements/conditions for their application. Then, we report results of coupling analyses in cardiovascular and cardiorespiratory studies. Finally, we compare the presented approaches.

2. Methods

The most applied approaches used to assess direct and indirect couplings were grouped using traditional domain classification into five classes: GC, nonlinear prediction, entropy, symbolization and phase synchronization, describing their main properties and requirements for application. For a better understanding of how these approaches work, electronic supplementary material is available. The electronic supplementary material gives basic descriptions of principles of classifications and examples of introduced approaches.

(a) Granger causality

Wiener [10] defined causality between two time series in a statistical framework by saying that: ‘for two simultaneously measured time series, one series can be called causal to the other if we can better predict the second time series by incorporating knowledge of the first’ [11, p. 145]. Later, Granger adapted this concept to the context of stochastic processes in economics to analyse linear time series based on an autoregressive (AR) model. Today, this concept is known as GC. GC for two processes \( X_1(t) \) and \( X_2(t) \) is defined as: \( X_1(t) \) has causal influence on \( X_2(t) \); \( X_1(t) \rightarrow X_2(t) \) if the knowledge of the past of both \( X_2(t) \) reduces the variance of the prediction error of \( X_2(t) \) in comparison with the knowledge of the past of \( X_2(t) \) alone (the past and the present cause the future but not vice versa) [7,12]. GC can be assessed by linear and nonlinear methods as described below.

(i) Linear methods

All linear methods assess GC based on a parametric multivariate autoregressive (MAR) model and can be split into methods that assess GC in the time or in the frequency domain.

Time domain. Geweke [13] was the first to propose a linear bivariate time-series approach assessing linear Granger causality (LGC) based on the prediction error variance of the time series associated with a statistical test for causality [12]. If process \( X_2 \) has no causal influence on process \( X_1 \), then \( \text{LGC}_{X_2 \rightarrow X_1} \) becomes close to zero, which means that the knowledge of past values of \( X_2 \) does not improve the prediction of \( X_1 \), but if \( \text{LGC}_{X_2 \rightarrow X_1} > 0 \), then a causal influence can be assumed. High values of \( \text{LGC}_{X_1 \rightarrow X_2} \) and \( \text{LGC}_{X_2 \rightarrow X_1} \) indicate a bidirectional coupling or a FB relationship between the two time series [14].
Bassani et al. [15] estimated the direct causal coupling by applying two GC approaches (the F-test and the Wald test) along the baroreflex in two anaesthesiological procedures. The advantages of conducting two GC tests are that

— they need not assume that the cardiovascular control mechanisms occur along specific temporal scales as is the case with testing GC in the frequency domain;
— the percentage of false GC detections can be rigorously controlled by assigning the type I error probability accepted by the tests; and
— the distribution of the statistic that assesses GC under the null hypothesis assuming no causal relationship between the two series follows a classical statistical distribution (in the case of the F-test, the F-distribution; and for the Wald test, the chi-squared distribution), thus allowing analytical calculation of the critical value above which the null hypothesis is rejected [15].

These approaches based on predictability improvement (PI) were later extended to account for the effect of latent confounders such as respiration by Porta et al. [16]. To do this, an exogenous signal was added to the bivariate AR closed-loop model to evaluate the bias induced on causality when the exogenous signal source was disregarded. In addition, Porta et al. [17] proposed a multivariate dynamical adjustment (MDA) modelling approach (open loop, OLMDA; closed loop, CLMDA) to assess the strength of the baroreflex as well as of direct and indirect cardiopulmonary couplings in contrast to the two previously mentioned approaches [15,16]; causal coupling was assessed using factorization signals into partial process decompositions, thus allowing the assessment of both direct and indirect couplings. The coupling strength in the MDA class is estimated as the variance, as indicated by the contribution of the partial process to the total variance.

Frequency domain. the most prominent linear approaches to calculate GC in the frequency domain are the partial directed coherence (PDC) [18,19] and the directed transfer function (DTF) [20]. These approaches are based on a fitted AR model and presuppose the stationarity of signals in the time interval under investigation [21].

PDC. This is a parametric approach based on m-dimensional MAR processes with order p. It can detect direct and indirect causal information transfer because it measures exclusively direct effects between signals in multivariate dynamic systems. Based on the Fourier transformation of the coefficient matrix, the PDC function quantifies the strength of the causal coupling from Xj to Xi as a function of frequency f. The PDC is normalized between 0 and 1, in that way, the direct influence from process Xj to process Xi is inferred by PDC \( \neq 0 \) (PDC = 0 when Xj does not cause Xi at frequency f, PDC = 1 when all causal influences originating from Xj at frequency f are directed towards Xi [8,14,19]. For PDC, a significance level was introduced to ensure reliable detection of the direct information flow [22].

Faes & Nollo [8] proposed the utilization of an extended MAR model to investigate either instantaneous and lagged effects (ePDC) or lagged effects (iPDC) only. They showed that the presence of instantaneous correlations may produce misleading profiles of PDC, whereas ePDC and iPDC provided a correct interpretation of extended and lagged causality, suggesting that ePDC and iPDC are more interpretable than PDC when applied to known cardiovascular and neuronal data. Milde et al. [23] introduced a time-variant version of PDC (tvPDC) for short-term multivariate data analysis which is based on a time-variant MAR model in combination with a Kalman filter for model parameter estimation. The Fourier transform of the AR parameters is used to define tvPDC. Milde et al. demonstrated that tvPDC avoids misinterpretations in HRV analyses and quantifies the partial correlative interaction properties between respiratory movements and RSA.

DTF. This enables the determination of directed causal interactions between two signals in relation to all other signals of the analysed system by applying on a MAR model using the transfer matrix to describe the causal information transfer [20,24]. Moreover, DTF measures both direct and indirect effects from one series to another, and for this reason, a differentiation between direct and indirect causal interactions or both is not possible, thereby leading to a greater
number of interactions than are actually present [25]. The DTF is normalized such that describes the ratio between the inflow from signal 2 to signal 1, to all the inflows of the activity to the destination signal 1. The DFT takes values between 0 and 1 (DTF = 1, most of the signal 1 consists of the signal 2; DTF = 0, almost no flow from signal 2 to signal 1 at frequency f) [20,24].

Korzeniewska et al. [24] proposed the directed DTF, an extension of the DTF in combination with partial coherence, that is able to estimate direct causal relations between signals. The short-time directed DTF (SDTF) [11,21] enables the construction of a time-variant GC and can be used to visualize the dynamics of transmissions using the DTF in combination with a short window spectral technique [26].

The main differences and similarities between PDC and DTF are [12,14,18,19,25]:

— DTF uses the transfer matrix, PDC uses the coefficient matrix;
— DTF cannot distinguish between direct or indirect causal information transfer, or both, whereas PDC can distinguish between both direct and indirect causal information transfer;
— PDC is more robust and efficient than DTF;
— PDC is normalized to the outflow of information (the structure that receives the signal), whereas DTF is normalized to the total inflow of the information (the structure that sends the signal);
— DTF and PDC both depend on the reliability of the fitted MAR model (i.e. optimal model order, epoch length);
— a significance level has to be used for both to avoid spurious interactions;
— PDC and DTF can be sensitive in detecting interactions in nonlinear multivariate systems under particular circumstances;
— DTF and time-varying PDC detect various types of time-varying influences (sensitive for time-resolved investigations of non-stationary data); and
— in bivariate cases, both PDC and DTF are reduced to the causal coherence introduced by Porta et al. [27].

(ii) Nonlinear methods

Granger mentioned that, in practice, it is usually not possible to use completely optimum predictors, unless all sets of series are assumed to be normally distributed, because such optimum predictors may be nonlinear in complicated ways. It would seem obvious to simply use linear predictors and to accept the above definitions under this assumption of linearity [7]. The most limiting factor for the assessment of nonlinear GC is the choice of the model because it must be appropriately matched to the dynamics of the investigated biosignals. However, there are some promising approaches that can quantify causality in nonlinear signals, such as methods based on nonlinear global model identification and methods based on local linear models.

Methods based on nonlinear global model identification were proposed by Faes et al. [28], who introduced an approach for the detection of coupling as well as the causality between two time series, based on nonlinear autoregressive (NAR) and nonlinear autoregressive exogenous (NARX) models. This approach is more accurate and sensitive in detecting imposed GC conditions and was more accurate than the least-squares methods based on the Akaike information model order criterion. The main advantages are that it can be applied to short data records, only a few parameters have to be defined and a more realistic physiological interpretation of the investigated system is possible. To assess GC by means of NARX models, they examined the mean-squared prediction error, which ranges between 0 and 1, where 0 represents a fully predictable time series and 1 a fully unpredictable time series. Causality from \( y \) to \( x \) can be investigated by reversing the input-output roles of the two series and by calculating the absolute and normalized relative PI obtained by the NARX model compared with the NAR model prediction, which resulted from the inclusion of \( y \) samples in the prediction of \( x \).
Another approach was introduced by Riedl et al. [29] based on nonlinear additive autoregressive (NAARX) models with external inputs fitted to bivariate time series for a model-based causal coupling analysis. They showed, if the additional external input led to a significant reduction in the variance of the predicting error, then the external input could be said to have a causal influence on the response variable. Here, the models are fitted separately to the time series, and the different AR predictors, as well as external ones, are compared with each other with regard to the best prediction. Thus, the improvement of the prediction was measured by a cross-validation criterion, which is weighted by the cost of adding a new predictor or increasing the roughness of the estimated curve. This approach also allows the quantification of the strength and the morphology of the selected couplings. Therefore, the variance of the separate single step prediction of each external input is calculated and normalized by dividing by the variance of the response variable. The resulting values indicate the strength of coupling between various external inputs and the response (increased values = increased coupling).

Methods based on local linear models were proposed by Chen et al. [30], who introduced a conditional-extended GC model using a spatial reconstruction of the joint dynamics of nonlinear multivariate time series. This is associated with various delays and is able to determine whether the causal relation between two nonlinear signals \((x, y)\) is coupled directly or mediated by another process [12]. Here, the extended Granger causality index (EGCI) \(\Delta y \rightarrow x\) was introduced as a function of \(\delta\) (neighbourhood size). If EGCI is less than 1, then it implies that \(y\) (or \(x\)) has causal influence on \(x\) (or \(y\)). For linear systems, \(\Delta y \rightarrow x\) will stay roughly the same as \(\delta\) becomes smaller, whereas for nonlinear systems, \(\Delta y \rightarrow x\) in the small \(\delta\) limit reveals the true nonlinear causal relation that may or may not be captured at the full attractor level [30]. In the multivariate case, a conditional-extended GC index was additionally proposed to distinguish between direct and indirect causal relations.

Ancona et al. [31] and Marinazzo et al. [32] used a radial basis function (RBF) approach to assess nonlinear GC in the context of PI between nonlinear bivariate time series. This means that in the frame of a linear regression model, if the prediction error of the first time series is reduced by including values from the second time series, then the second time series is said to have a causal influence on the first time series. Ancona & Stramaglia [33] demonstrated that not all nonlinear prediction approaches are suitable to evaluate GC between two time series because they should be invariant if statistically independent variables are added to the set of input variables (they were unable to quantify how much knowledge of the other time-series counts to improve prediction error). They stated that any prediction scheme providing a nonlinear extension of GC should satisfy the following property: (P1) if \(Y\) is statistically independent of \(X\) and \(x\), then \(\epsilon_X = \epsilon_{XY}\); if \(X\) is statistically independent of \(Y\) and \(y\), then \(\epsilon_Y = \epsilon_{YX}\) \((x, y, X, Y = \text{stochastic variables}; \epsilon_X = \text{the prediction error when} x \text{is predicted solely on the basis of knowledge of its past values; similarly for \(\epsilon_Y\)}\). Ancona et al. proposed a class of nonlinear models that satisfy P1 based on an RBF-like approach to assessing nonlinear GC [31,32]. Later, they introduced an approach to quantify nonlinear GC based on kernel Hilbert spaces providing a statistically robust tool to assess driver–response relationships [34,35]. This approach performs linear GC in the feature space of suitable kernel functions, assuming an arbitrary degree of nonlinearity and also fulfills the good properties of linear models in the nonlinear case. The problem of overfitting the models was handled by exploiting the geometry of reproducing kernel Hilbert spaces. The kernel algorithms work by embedding data in a Hilbert space, and searching for linear relations in that space (Hilbert spaces were spaces of kernel functions) [35]. The proposed approach has the following features:

- the nonlinearity of the regression model can be controlled by choosing the kernel function;
- the problem of false causalities, which arises as the complexity of the model increases, is addressed by a selection strategy of the eigenvectors of a reduced Gram matrix; and
- reduction of unknown model parameters by considering a relatively small number of (possibly nonlinear) mixtures of parameters thus bounds model complexity and ensures that the accuracy of different models is a rough approximation to their evidence [35].
(b) Nonlinear prediction

Nonlinear prediction approaches are based on cross-prediction and are similar to those based on PI in their underlying methodological framework, but differ from PI in that they do not measure GC, but rather a causality concept that exploits asymmetry of cross-predictability (CP) when performed over the two possible directions of interaction between two series. Farmer & Sidorowich [36] were the first to introduce a concept of prediction, called the $k$-nearest neighbour prediction, which was later integrated into different approaches.

Based on local linear prediction (the nearest neighbour local linear approximation [36]), an extension of a nonlinear bivariate prediction approach for the investigation of causal interdependencies between two time series $(x, y)$ with a specific out-of-sample cross-validation approach (to avoid overfitting) was introduced for short-term time series [37]. This method was based on the principle of the mutual prediction method, which provides a measure for the coupling strength and coupling directionality between two time series [38]. The bivariate prediction model is defined using knowledge of similar patterns in the first time series to predict the current values of a second time series. By exploiting the input and output series $x$ and $y$, the relationship between a pattern of samples of $x$ and a synchronous sample of $y$ was approximated using a linear polynomial whose coefficients were estimated applying an equation system, including the nearest neighbour patterns in $x$ and the corresponding samples in $y$. Finally, an index describing the level of predictability (LP) was defined $(LP = 1$, when $y$ is completely predictable given patterns of length $L$ in $x$; $LP = 0$ (also negative)) when $y$ is completely unpredictable in relation to $x$ indicating the complete uncoupling. The advantage of this predictor is that it allows the possibility to obtain dependable estimates of predictability without constraining the embedding of the series when dealing with short-term time series as well as the predictor being less biased (overfitting) [37]. In 2008, three different mutual nonlinear prediction approaches ($k$-nearest neighbours) were compared: the cross-prediction, the mixed prediction and the PI to test their ability to assess the coupling strength and directionality of the interactions in bivariate time series [39]. Based on simulations and real physiological data (cardiovascular), it was found that cross-prediction is valuable for quantifying the coupling strength and PI for determining the directionality of interactions in short and noisy bivariate time series.

These approaches have the following properties:

— cross-prediction quantifies interdependence in terms of the predictability of one of the two series estimated using state space vectors of the other series;
— mixed prediction quantifies interdependence in terms of the predictability of a series using state vectors that contain samples of both series; and
— PI quantifies interdependence in terms of the increase in predictability yielded by mixed prediction compared with self-prediction [38,39].

(c) Entropy

Methods based on entropies have in common that they analyse a putative information transfer between signals/time series. The concept of entropy addresses the uncertainty or predictability of signals. Greater entropy values reflect higher uncertainty and lower predictability. The concept of Shannon entropy $(H)$ was introduced by Shannon [40], quantifying the information content within a time series. $H(x)$ describes the statistical properties of a time series $x$ (stationary) and represents a measure of uncertainty of a time series based on probabilities.

$$H(x) = - \sum_{i=1}^{M} p(x_i) \log_2 p(x_i),$$

where $p(x_i)$ is the probability distribution of $i$th bin of the time series $x$ and $M$ is the total number of all bins.

Mutual information (trans-information). The concept of mutual information (MUI) is based on the determination of the Shannon entropies $H_x$ and $H_y$ as well as the joint entropy $H_{xy}$, where $p(x)$
and \( p(y_j) \) are probability distributions of \( x \) and \( y \) and \( p(x_i, y_j) \) is the joint probability distribution of both time series,

\[
\text{MUI}_{xy}[\text{bit}] = H_x + H_y - H_{xy} = \sum_{i,j=1}^{M} p(x_i, y_j) \log_2 \left( \frac{p(x_i, y_j)}{p(x_i)p(y_j)} \right).
\]

MUI analysis is applied to detect and quantify non-directional linear and nonlinear interdependences within one time series (univariate) or between different (bi- and multivariate) time series. MUI measures the information that \( x \) and \( y \) share in units called ‘bits’ because of the application of \( \log_2 \) [41]. Large values of MUI represent strongly dependent time series, and small values indicate nearly independent ones (MUI = 0, if \( x \) and \( y \) are completely independent). Moreover, MUI is symmetric (MUI_{xy} = MUI_{yx}), non-negative (MUI_{xy} ≥ 0) and bounded from above by \( \min[H_x, H_y] \) [42,43].

**Cross-conditional entropy.** Porta et al. introduced the cross-conditional entropy (CE_{xy}) based on the conditional entropy (CE) as a modification of the Shannon entropy. CE_{xy} quantifies the degree of coupling between two normalized time series (\( x,y \)) and represents a measure of the complexity of \( x \) with respect to \( y \) [44],

\[
\text{CE}_{x/y}(L) = - \sum_{L=1} \sum_{l=1} p(y_{l-1}) \sum_{t/L-1} p(x(t)/y_{l-1}) \log p \left( \frac{x(t)}{y_{l-1}} \right),
\]

with pattern length \( L \), joint probability \( p(y_{l-1}) \) of the pattern \( y_{l-1}(t) \) and conditional probability \( p(x(t)/y_{l-1}) \) of the sample \( x(t) \), given the pattern \( y_{l-1} \). CE_{xy} is a process of sorting and counting mixed patterns. It describes the amount of information included in the sample \( x(t) \) when the pattern of \( L - 1 \) samples of \( y_{l-1}(t) \) is given. CE_{xy} is intended as the pattern formed by the past \( L - 1 \) samples of \( y \), i.e. \( y_{l-1}(t) = (y(t-1), \ldots, y(t-L+1)) \). CE_{xy} has the following inherent properties:

— it is equal to zero when a sufficient number of samples of \( y \) allow complete prediction of \( x \);
— it is high and constant if \( x \) and \( y \) are independent processes; and
— it decreases towards a value between these extremes when the knowledge of \( y \) is useful to partially estimate \( x \) [45].

In addition, the synchronization index: \( \chi_{x,y} = 1 - \min(\text{UF}_{x,y}) \) (with \( \text{UF} \) the uncoupling function) quantifying the maximum amount of information exchanged between the two time series [44] can be estimated. The larger the synchronization index, the more coupled the two time series are.

**Transfer entropy.** Schreiber [46] proposed an information-theoretical approach named transfer entropy (TE) to distinguish between driving and responding elements and to detect asymmetries in the interaction and to quantify the extent to which the dynamics of one process influences the conditioned transition probabilities of another. TE measures GC with the prediction improvement approach and extends the concept of Shannon entropy by taking into account the probabilities of transitions rather than static probabilities. Generally, if no information flow from process \( Y \) to process \( X \) exists, then the state of \( Y \) has no influence on the transition probabilities on \( X \). To analyse the dynamics of the shared information between the two processes, the deviation from the generalized Markov property is measured with its transition probabilities \( p(x_{n+1}^{(k)}, y_{n}^{(l)}) = p(x_{n+1}^{(k)}, y_{n}^{(l)}), \) where \( l \) is the conditioning state from process \( Y \). The Kullback–Leibler divergence is used to measure the deviation from the generalized Markov property leading to TE,

\[
\text{TE}_{Y \rightarrow X} = \sum p \left( x_{n+1}^{(k)}, y_{n}^{(l)} \right) \log \left( \frac{p \left( x_{n+1}^{(k)}, y_{n}^{(l)} \right)}{p \left( x_{n+1}^{(k)} \right)} \right).
\]
The most important feature of TE is that it is asymmetric under exchange of $X$ and $Y$ ($\text{TE}_{X \rightarrow Y} \neq \text{TE}_{Y \rightarrow X}$). The direction of coupling and information flow between coupled processes can be determined more adequately in comparison with MUI and CE [47,48]. The main advantage of TE is its ‘model-free’ approach [49]. A model-free causality statistic can be defined as the conditional MUI between the past of one process and the future of another process, given the knowledge about the past of the latter. Moreover, it can be shown that, under proper conditions, TE is equivalent to the conditional MUI [49]. In addition, it must be noted that conditional entropies are estimated directly from sampled probability distributions; results will vary with the estimation technique applied. Thus, a naive estimation of TE, for example, by partitioning the state space is problematic and might fail to converge to the correct result. In practice, more sophisticated techniques such as kernel or $k$-nearest neighbour estimators will be required [48,50]. Furthermore, there are close similarities between TE and GC as entirely formal equivalence. This has been demonstrated in the case of Gaussian stochastic processes by Barnett et al. [50], thus bridging AR and information-theoretic approaches to data-driven causal inference. GC approaches are typically implemented within a framework of MAR models, but imply many assumptions about how to model the data. Thus, the main problem of a parametric approach is the model misspecification. On the other hand, TE may present severe difficulties of empirical application, despite being theoretically ‘model agnostic’.

Vakorin et al. [49] introduced the partial transfer entropy (pTE), a GC measure based on a multivariate version of TE that quantifies causality between two nodes of an interacting network. They found that pTE is a more sensitive technique for identifying robust causal relations than its bivariate equivalent and demonstrated the confounding effects of the variation in indirect coupling on the detectability of robust causal links. Verdes [48] applied the concept of TE to multivariate time series, thus proposing a model-free, information-theoretical tool for examining weakly dependent time series that compare the relative influence of two or more external dynamics trigger on a given system. The TE approach was extended by Staniek & Lehnertz [51] by using a technique of symbolization to estimate TE, called symbolic transfer entropy (STE). STE is a robust and computationally fast method to quantify the dominating direction of information flow between time series from structurally identical and non-identical coupled systems. Recently, Faes et al. [52,53] introduced an enhanced version based on the corrected CE and a sequential procedure for non-uniform embedding to assess nonlinear GC in multivariate time series. It is actually a modified estimation of TE. This approach quantifies causality from one time series to another as the amount of information flowing directly from the first to the second time series, while accounting for the effects of all other time series in the multivariate representation, and separates direct and indirect causal effects.

Investigators have the choice to choose the most suitable approach for their data analysis. Numerical issues aside, analytical equivalence underlines the essential point that under Gaussian assumptions, GC has a natural interpretation as TE and vice versa [50].

**Symbolization**

Methods based on symbolization enable a coarse grain quantitative assessment of short-term dynamics of time series. The direct analysis of successive signal amplitudes is based on discrete states (symbols).

*Joint symbolic dynamics (JSD).* JSD was introduced by Baumert et al. [54] and is based on the analysis of bivariate dynamic processes by means of symbols [55]. JSD considers short-term beat-to-beat changes, allowing the assessment of overall short-term cardiovascular and cardiorespiratory couplings. Therefore, a bivariate sample vector $X$ of two time series $(x,y)$ is transformed into a bivariate symbol vector $S$ where $n$ are beat-to-beat values,

$$X = \{(x_n, y_n)^T\} \xrightarrow{n=0,1,...,x \in R} S = \{(s(x_n), s(y_n))^T\} \xrightarrow{s=0,1,...,s \in 0,1}.$$


For symbol transformation, a given alphabet $A = \{0, 1\}$ was used, where symbol ‘1’ represented increasing amplitude values and symbol ‘0’ decreasing and unchanged amplitude values (JSD2). Afterwards, short patterns (words $w$) of symbol sequences with a word length of three symbols were formed. For the quantification of JSD2 from each word type, the normalized joint probability ($p(w_{ij})$) of occurrence was estimated using an $8 \times 8$ word distribution density matrix $W$ ranging from $(000,000)^T$ to $(111,111)^T$. In addition, from the matrix $W$, the probabilities of all single type occurrences $p(w_{xy})$, the sum of each column $c_j$, the sum of each row $r_i$ and the Shannon entropy as a measure of the overall complexity and probability of occurrence of each column and each row can be computed.

Recently, we introduced a new high-resolution version of JSD that is characterized by three symbols which are formed on the basis of a threshold ($l \neq 0$) and which clusters the coupling behaviour into eight word type families (HRJSD) for the quantification of cardiovascular couplings. Thus, it circumvented the problems encountered by JSD2 to distinguish between decreases and steady state, as well as between small and large changes of autonomic regulation owing to $l = 0$ and $A = \{0, 1\}$. It is also impossible to differentiate between noise, artefacts (e.g. generated by undersampling or ectopic events) and fluctuations that arise from autonomic regulation. Both approaches have the main advantages that they are not sensitive to non-stationary time series and are capable of capturing nonlinear bivariate couplings by a simple procedure.

*Symbolic coupling traces (SCTs).* A JSD extension, the SCT was introduced by Wessel *et al.* [56]. It is based on the analysis of structural patterns and enables the detection of the direction (bidirectional) of time-delayed couplings in short-term bivariate time series. Using the JSD2 algorithm, two time series $x(t)$ and $y(t)$ were transformed into symbol sequences $s_x(t)$ and $s_y(t)$ using also the alphabet $A = \{0, 1\}$, and afterwards, series of word $w_x(t)$ and $w_y(t)$ of length $l = 3$ were formed. In contrast to JSD2, a delay-time probability matrix $\Pi(\tau) = (p_{ij}(\tau))$ was estimated describing how word $W_j$ would occur in $w_x$ at time $t$ and $W_j$ would occur in $w_y$ at time $(t + \tau)$ with $p_{ij}$ as the joint probabilities of the words.

For the quantification of SCTs, only the symmetric and diametric traces of the bivariate word distribution (BWD) matrix were used, thereby excluding random effects and including only significant coupling information. Here, tree indices can be calculated: the trace $T$ of the matrix $\Pi(\tau)$ representing the fraction of both time series, which are structurally equivalent (symmetrical influences) to each other at lag $\tau$. The trace $\bar{T}(\tau)$ describes the fraction of each signal, which is structurally diametric at lag $\tau$ [57]. Both parameters vary from 0 to 1 and comprise the diagonals of the BWD only. Finally, the difference $\Delta T = T - \bar{T}$ can be calculated to determine the exact detection of lags (delayed couplings) between two time series. The lags $\tau$ should be limited to $20 \leq \tau \leq 20$ (sampling units) in order to focus on short-time-delayed dependencies only. The main advantages of SCTs are their ability to detect delayed coupling (time lags), its applicability to moderately noisy time series (less than 10 dB) and its insensitivity to non-stationarity.

**(e) Phase synchronization**

Rosenblum *et al.* [58,59] proposed an approach based on phase synchronization, or the directionality index $d$, for the detection and quantification of coupling directions of weakly coupled self-sustained bivariate time series, even when the interaction between the two time series is too weak to induce synchronization. The term ‘phase synchronization’ is used to denote the state when a relation only between the phases $(\Phi_1, \Phi_2)$ of interacting signals sets in, but the amplitudes remain chaotic and nearly uncorrelated [60]. In contrast to the other proposed methods for examining signal amplitudes, this approach examines directly the oscillation phases. The idea behind this approach is that if a signal 1 is driven by signal 2, then the evolution of $\Phi_1$ also depends on $\Phi_2$; in other words, prediction of $\Phi_1$ from its previous values can be improved by taking into account the prehistory of $\Phi_2$ only if signal 2 drives signal 1 [58]. This means that weak coupling affects the phases of interacting time series (oscillators), whereas the amplitudes of those oscillators remain practically unchanged and the dynamics of the interacting signals can
be reduced to those of two phases, $\Phi_{1,2} = \omega_{1,2} + \varepsilon_{1,2} f_{1,2}(\Phi_{1,1}, \Phi_{2,1}) + \xi_{1,2}$, where random terms $\xi_{1,2}$ describe the noisy perturbations, small parameters $\varepsilon_{1,2} \ll \omega_{1,2}$ characterize the strength of the coupling, functions $f_{1,2}$ are $2\pi$ periodic and $\omega_{1,2}$ are the natural frequencies of the two oscillators. The main advantages of this approach are that it can be applied to both noisy and chaotic time series. It reveals whether the interaction is bi- or unidirectional and quantifies the degree of asymmetry of bidirectional coupling. In addition, the feasibility of a posterior estimation of coupling direction is possible and it can widely be used in biomedical studies where the parameters of systems typically cannot be assessed or only measurements under free running conditions are possible [59].

In the context of analysing instantaneous phases, Paluš & Stefanovska [61] introduced an approach based on the conditional MUI method for the detection of the direction of coupling from phases of weakly coupled oscillators. This approach is able to distinguish between uni- or bidirectional couplings and to quantify the degree of asymmetry in bidirectional coupling. Schäfer et al. [62] used the concept of phase synchronization of chaotic oscillators to analyse irregular non-stationary and noisy bivariate time series using the cardiorespiratory synchronogram (CRS) to detect different synchronous states $(n:m)$ and transitions between the two time series and to distinguish between different periods of synchronization using their instantaneous phases. Eckmann et al. [63] introduced the method of recurrence plots (RPs) to visualize the recurrences of a dynamical system in its phase space. Bivariate cross-recurrence plots (CRPs) are extensions of RPs, and can be used to analyse the nonlinear dependencies between two different systems by comparing their states. A CRP is essentially assumed as a generalization of the linear cross-correlation function. To quantify CRPs, indices of complexity were introduced mainly based on diagonal structures in CRPs. CRPs can find nonlinear interrelations from bivariate time series, whereas linear correlation tests cannot [64].

Table 1 summarizes approaches used to assess direct or non-indirect coupling with their method-specific characteristics, their similarities and differences, their basic requirements/conditions for application and their main features.

### 3. Applications of cardiovascular and cardiorespiratory coupling analyses

#### (a) Granger causality

##### (i) Linear methods in the time and frequency domain

Bassani et al. [15] applied two GC approaches from systolic blood pressure (SBP) to RR (F-test and Wald test) in two anaesthesiological studies involving the volatile administration of sevoflurane plus remifentanil or the intravenous administration of propofol plus remifentanil. They found that both approaches detected similar percentages of GC from SBP to RR, thus suggesting that both procedures performed similarly in assessing GC in the RR–SBP variability interactions during general anaesthesia. In addition, they stated that GC from SBP to RR was preserved during general anaesthesia, thus allowing the use of techniques exploiting spontaneous variability for the assessment of baroreflex sensitivity (BRS), and that volatile and intravenous anaesthetic similarly affected the RR–SBP relationship.

Porta et al. [17] evaluated the strength of the causal relationship from SBP to RR of the baroreflex and from respiration (RESP) to RR in the cardiopulmonary pathway. They applied an MDA (OLMDA and CLMDA) modelling approach to 19 healthy subjects to assess the coupling strength during progressive sympathetic activation induced by graded head-up tilt. Both OLMDA and CLMDA models revealed that baroreflex coupling exhibited a significant positive linear correlation with tilt table angles that progressively increased with tilt table inclination ($15^\circ, 30^\circ, 45^\circ, 60^\circ$ and $90^\circ$). They concluded that the gradual sympathetic activation provoked by graded head-up tilt is accompanied by an involvement of baroreflex becoming increasingly important with higher tilt table inclination. Only the CLMDA model indicated that cardiopulmonary coupling owing to the direct link from RESP to HR gradually decreased with tilt...
Table 1. Summary of different coupling approaches with their method-specific characteristics. cv, cardiovascular; cr, cardiorespiratory; eeg, electroencephalogram; ---, ‘no information provided’.

<table>
<thead>
<tr>
<th>coupling approach</th>
<th>classes</th>
<th>directed</th>
<th>non-directed</th>
<th>short term</th>
<th>long term</th>
<th>signals</th>
<th>stationarity required</th>
<th>specifics and restrictions</th>
</tr>
</thead>
<tbody>
<tr>
<td>F-test and Wald test [15]</td>
<td>linear Granger causality; time domain</td>
<td>x</td>
<td>—</td>
<td>x</td>
<td>—</td>
<td>cv, cr</td>
<td>yes</td>
<td>no assumption about the temporal scale of the oscillations under examination is needed; performances are not affected by differences in complexity between signals and sharpness; frequency of dominant oscillations does not influence the rate of false positives</td>
</tr>
<tr>
<td>multivariate dynamical adjustment modelling [17]</td>
<td>linear Granger causality; time domain</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>—</td>
<td>cv, cr</td>
<td>yes</td>
<td>main feature uses unique characteristics of the model structure; variance normalization of the prediction error prevents overestimating coupling strength</td>
</tr>
<tr>
<td>partial directed coherence [8,19,23]</td>
<td>linear Granger causality; frequency domain</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>cv, cr, eeg</td>
<td>yes*</td>
<td>mainly depends on the estimation of the model parameters (model order $p$) of MAR and window length (overfitting, instability); MAR model assumes that the underlying process is stationary; * tvPDC suitable for non-stationary signals</td>
</tr>
<tr>
<td>directed transfer function [20,24,26]</td>
<td>linear Granger causality; frequency domain</td>
<td>x</td>
<td>—</td>
<td>x</td>
<td>x</td>
<td>cv, cr, eeg</td>
<td>yes*</td>
<td>depends on the estimation of the model parameters; MAR model assumes that the underlying process is stationary; * SDIF suitable for non-stationary signals</td>
</tr>
</tbody>
</table>

(Continued.)
<table>
<thead>
<tr>
<th>coupling approach</th>
<th>classes</th>
<th>directed</th>
<th>non-directed</th>
<th>short term</th>
<th>long term</th>
<th>signals</th>
<th>stationarity</th>
<th>specific and restrictions</th>
</tr>
</thead>
<tbody>
<tr>
<td>extended Granger causality index [30]</td>
<td>nonlinear Granger causality</td>
<td>x</td>
<td>—</td>
<td>—</td>
<td>x</td>
<td>not proven</td>
<td>yes</td>
<td>number of neighbours and δ neighbourhood size have to be set up; practical usefulness not yet proven</td>
</tr>
<tr>
<td>nonlinear autoregressive exogenous (NARX) models [28]</td>
<td>nonlinear Granger causality</td>
<td>x</td>
<td>—</td>
<td>x</td>
<td>—</td>
<td>cv, cr</td>
<td>yes</td>
<td>NARX model is appropriate only up to third-order nonlinearity, as the number of model terms to be searched becomes excessively large thereafter</td>
</tr>
<tr>
<td>nonlinear additive autoregressive (NAARX) models [29]</td>
<td>nonlinear Granger causality</td>
<td>x</td>
<td>—</td>
<td>x</td>
<td>—</td>
<td>cv, cr</td>
<td>—</td>
<td>applicable only for short-term couplings up to fifth order</td>
</tr>
<tr>
<td>bivariate nonlinear prediction [36,37,39]</td>
<td>nonlinear prediction</td>
<td>x</td>
<td>—</td>
<td>x</td>
<td>—</td>
<td>cv, cr</td>
<td>yes</td>
<td>time series must be normalized; LP depends on the length ( L ) of the patterns and on the number ( K ) of neighbour patterns</td>
</tr>
<tr>
<td>mutual information [41,43]</td>
<td>entropy</td>
<td>—</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>cv, cr</td>
<td>yes</td>
<td>high noise levels and artefacts influence method</td>
</tr>
<tr>
<td>cross-conditional entropy [44,65]</td>
<td>entropy</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>cv, cr</td>
<td>yes</td>
<td>time series must be normalized; for short-term data the corrected conditional entropy has to be used to estimate the UF; without any ( \text{apriori} ) selection of the embedding dimension</td>
<td></td>
</tr>
<tr>
<td>Coupling Approach</td>
<td>Classes</td>
<td>Directed</td>
<td>Non-directed</td>
<td>Short Term</td>
<td>Long Term</td>
<td>Signals Required</td>
<td>Specifics and Restrictions</td>
<td></td>
</tr>
<tr>
<td>---------------------------</td>
<td>---------------</td>
<td>----------</td>
<td>--------------</td>
<td>------------</td>
<td>-----------</td>
<td>------------------</td>
<td>----------------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Transfer entropy</td>
<td>Entropy</td>
<td>x</td>
<td>—</td>
<td>x</td>
<td>x</td>
<td>cv, cr</td>
<td>Yes. Parametric distribution of the processes must be known; time-series analysis: convergence to the continuous entropies often cannot be achieved; depends on bandwidths used in the analysis.</td>
<td></td>
</tr>
<tr>
<td>Joint symbolic dynamics</td>
<td>Symbolization</td>
<td>—</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>cv, cr</td>
<td>No. Coarse-graining without no information about information flow.</td>
<td></td>
</tr>
<tr>
<td>Symbolic coupling traces</td>
<td>Symbolization</td>
<td>x</td>
<td>—</td>
<td>x</td>
<td>x</td>
<td>cv, cr</td>
<td>No. High noise levels (&gt; 10 dB) influence detection of delayed coupling; critical significance level $T_{\text{crit}}$ has to be calculated.</td>
<td></td>
</tr>
<tr>
<td>Directionality index $d$</td>
<td>Phase Synchronization</td>
<td>x</td>
<td>—</td>
<td>x</td>
<td>cv, cr</td>
<td>No</td>
<td>Only the time interval $\tau$ has to be set up for the estimation of the increments of phases and should be related to the periods of oscillation $\tau = T_{\text{osc}}$; appropriate for noisy records; due to non-stationarity, the index $d$ has to computed in a sliding window (e.g. with 3/4 overlap); fitting procedure has to be optimized.</td>
<td></td>
</tr>
<tr>
<td>Directionality index $D$</td>
<td>Phase Synchronization; Entropy</td>
<td>x</td>
<td>—</td>
<td>x</td>
<td>x</td>
<td>cv, cr</td>
<td>Yes for long time lags $\tau$, the directionality detection ability of $D(1, 2)$ is lost; significance level for $D(1, 2)$ must be set up for correct detection (bias, variance) of directionality.</td>
<td></td>
</tr>
</tbody>
</table>
table angle because the indirect link mediated by SBP increased progressively. OLMDA provided a limited view on cardiovascular control, which means it was able to detect the progressive decrease of direct cardiopulmonary coupling, but was unable to detect the increase of indirect cardiopulmonary coupling.

Nollo et al. [69] used the causal coherence approach to calculate the coherence from SBP to RR and from RR to SBP and to compute the gain and phase of the baroreflex transfer function. The method was applied and compared with the non-causal one, to RR and SBP series taken from healthy subjects in the supine position and after a passive head-up tilt. They demonstrated for the low frequency (LF; 0.04–0.15 Hz) spectral component, the enhanced FF coupling, and the blunted FB coupling found at rest indicated the prevalence of non-baroreflex mechanisms. The tilt-manoeuvre-recorded FB influences are stronger than the FF interactions. At the respiratory frequency, the RR–SBP regulation was balanced at rest and shifted towards FB mechanisms after tilt. Thus, they could present evidence that the causal baroreflex gain estimates were always lower than the corresponding non-causal values and decreased significantly from rest to tilt in both frequency bands. They concluded that the importance of non-baroreflex interactions underline the necessity to determine causality in the cross-spectral analysis of the interactions between cardiovascular variables in healthy humans.

Faes et al. [70] investigated the impact of a causal transfer function analysis to characterize closed-loop interactions between cardiovascular and cardiorespiratory signals (RESP → HR, RESP → SBP, SBP → HR) from healthy subjects who were in the supine position and breathing spontaneously. They found that in the cases of RESP → HR and RESP → SBP, the causal and non-causal transfer function estimates were comparable, indicating that RESP, acting as an exogenous signal, sets up an open-loop relationship with SBP and HR. On the contrary, causal and traditional transfer functions from SBP → HR were significantly different, suggesting the presence of a considerable influence in the opposite causal direction. In another study of this group [8], they applied an extended multivariate AR model based on the PDC calculation (ePDC, iPDC) to HR, SBP and RESP time series to analyse short-term cardiovascular and cardiorespiratory interactions in healthy subjects (breathing spontaneously, standing in a 60° upright position after passive head-up tilt). They found that compared with generalized PDC, the ePDC and iPDC functions show congruent trends in other directions (RESP → SBP and HR → SBP) and different trends in other directions (RESP → HR and SBP → HR). They presumed that these findings suggest more reliable values for ePDC/iPDC than for generalized PDC because a depressed RSA (lower RESP → HR coupling) and enhanced baroreflex regulation (higher SBP → HR coupling) are expected phenomena for subjects standing in the upright position. They concluded that ePDC and partial PDC are better interpretable than PDC and generalized PDC in terms of known cardiovascular and neural physiologies.

Milde et al. [23] used tvPDC for the multivariate analysis of HRV, RESP and SBP. They demonstrated that respiration-related HRV components can be identified that also occur at other frequencies than the RSA frequency. These additional components are known to be an effect of the ‘half-the-mean-heart-rate-dilemma’ (cardiac aliasing; CA). CA components may contaminate the entire frequency range of HRV and can lead to misinterpretation of the RSA analysis. The tvPDC analysis of full-term neonates and sedated patients revealed these contamination effects and, in addition, that the respiration-related CA components can be separated from the RSA component and the Traube–Hering–Mayer wave. They concluded that tvPDC can be beneficially applied to avoid misinterpretations in HRV analyses as well as to quantify partial correlative interaction properties between RESP and RSA.

(ii) Nonlinear methods

Ancona et al. [31] investigated HR and breath rate of a sleeping human subject suffering from sleep apnoea (10 min) and found that the causal influence of HR on breath is stronger than the reverse, with a directional index $D = 0.76$ (positive and rather large value). This value corresponds to the 0.4 Hz characteristic of the respiratory rhythm.
Marinazzo et al. [32] applied an RBF approach to assess nonlinear GC. They investigated time series of HR and SBP in six congestive heart failure (CHF) patients and six patients affected by sepsis, and found that sepsis patients showed symmetric causal relationships between HR and SBP, whereas CHF patients showed an unbalanced HR and SBP relationship, in which HR significantly dominated SBP.

NAR and NARX models [28] were used for the investigation of HR and SBP time series of young healthy subjects (25 ± 3 years old; 15 min resting supine position and 15 min upright position (tilt table, 60°)). Findings revealed that in the supine position, SBP was significantly more predictable than RR in both NAR and NARX model predictions. After tilt, the predictability of the RR improved significantly. They could demonstrate a bilateral causal relationship between the two signals with the NARX model prediction, compared with the NAR model prediction. Furthermore, they found a significant reduction in the complexity of the dynamics of the two causal signal pathways as the body position was changed from supine to upright. They concluded that the NARX model can be applied to physiological signals to improve understanding of causality and coupling under normal and diseased conditions.

Riedl et al. [29] applied NAARX models with external input in order to characterize the dynamical changes of autonomic regulation (RESP, SBP, diastolic blood pressure (DBP), HR) of healthy pregnant women and subjects suffering from preeclampsia (PE). They found that both groups showed a significant RESP influence on HR (nonlinear form of the RESP influence on the HR was significantly different between the two groups) and blood pressure, an influence of HR on DBP, and for pregnant women in a supine position with relaxed breathing, a dependence of SBP on DBP. The structure of the interactions between HR, SBP, DBP and RESP was equal for both groups; however, there were clear differences in effect size and morphology of several interactions, indicating a neuronal change in patients caused by the resetting of the respiratory gating to more extreme respiratory values. Their final conclusion pointed to a potential role of RESP for understanding the pathogenesis of PE.

(b) Nonlinear prediction

The study of Faes & Nollo [37] applied a bivariate nonlinear prediction method to simulated and cardiovascular signals from healthy young subjects to characterize the dynamic interaction between short-term spontaneous fluctuations of HR and SBP. They found that in the supine position and given SBP, the predictability of HR was low and influenced by nonlinear dynamics, and that after head-up tilt, the predictability increased significantly. They assumed that their results were related to the larger involvement of the baroreflex regulation from SBP to HR in the upright rather than in the supine position, and to the simplification of the HR–SBP coupling occurring with the tilt-induced alteration of the neural regulation of cardiovascular rhythms.

Comparable results were found by Porta et al. [71], who investigated short-term (300 samples) time series of HR and SBP from (i) healthy subjects (at rest during spontaneous breathing, paced metronome breathing, 80° head-up tilt) and (ii) healthy young subjects (at rest during spontaneous breathing as well as after intravenous administration of two boluses of atropine). They demonstrated that SBP was more predictable than RR and that the predictability of the RR was larger during tilt, controlled respiration at 10 breaths per minute and after a high dose of atropine. In addition, SBP was dominated by linear correlation, and RR exhibited nonlinear dynamics during controlled respiration at 10 bpm and after a low dose of atropine, whereas it was linear during sympathetic activation produced by tilt and after a high dose of atropine.

Nollo et al. [72] applied a mutual nonlinear prediction approach to investigate causal short-term cardiovascular variability (SBP → RR, RR → SBP) of healthy subjects and disease states in the supine position, in the upright position after head-up tilt (early tilt) and after prolonged upright posture (late tilt). They showed that both CP and PI indices showed a marked decrease in coupling in both causal directions in syncope subjects during late tilt, indicating the impairment of cardiovascular regulation in the time interval preceding syncope. In addition, PI elicited the prevalence of causal coupling from RR → SBP during supine better than CP. This decrease was
the main marker of cardiovascular dysregulation preceding syncope. Another major finding was that cardiovascular regulation in the supine position was unbalanced, with a prevalence of the FF effects from HR → SBP, mainly of a mechanical nature, over the FB influences from SBP → HR, commonly ascribed to the neural baroreflex mechanisms.

In a comparative study [39], three approaches based on cross-prediction, mixed prediction and PI were applied to HR and SBP time series (300 samples) in healthy subjects at rest and in the upright position after head-up tilt. They found an increased strength of interaction between HR and SBP after head-up tilt, suggested to be an effect of the activation of the sympathetic nervous system that enhanced primarily the FB baroreflex regulation from SBP to HR and, consequently, the FF influence of HR on SBP. Finally, they concluded that cross-prediction is more suitable than mixed prediction and PI for quantifying the degree of the interaction in bivariate time series.

(c) Entropy

Hoyer et al. [73] investigated 39 patients after acute myocardial infarction (AMI) and 24 healthy controls, applying the auto-mutual information function (AMIF) and the cross-mutual information function (CMIF). They found that the discrimination (univariate, multivariate) between both groups was significantly improved by cardiorespiratory cross-mutual information measures (AMIF, CMIF) in comparison with the standard linear HRV indices. They concluded that these measures quantitatively characterize relevant aspects of cardiorespiratory coordination and might further improve the identification of patients who are at sudden risk of death after AMI and at risk of severe heart failure.

Bahraminasab et al. [74] studied the direction of bivariate couplings of non-invasively recorded 30 min HR and RESP time series in healthy subjects, applying the conditional mutual information method. They found that respiration drives cardiac activity more than vice versa, an outcome that is consistent with direct physiological observations.

Porta et al. [75] applied cross-CE to assess HR and SBP causality of short-term (STHT) and 11 long-term heart transplant (LTHT) recipients and healthy subjects at rest and during graded head-up tilt. They found a lack of causal relation from SBP → HR in STHT recipients and, in LTHT recipients, a gradual shift over time of the causal link from SBP → HR. The head-up tilt protocol induced the progressive shift from the prevalent causal direction from HP → SBP to the reverse causality (i.e. from SBP → HR) with tilt table inclination in healthy subjects. The dependence of causality on tilt table inclination suggested that ‘spontaneous’ BRS, which was estimated using non-causal methods (e.g. spectral and cross-spectral approaches) was more reliable at the highest tilt table inclinations.

Schreiber [46] investigated the breath rate and instantaneous HR of a sleeping human suffering from sleep apnoea and applied time-delayed MUI and TE. The results were that time-delayed MUI was almost symmetric between the two series and that the TE also found information flow in both directions but indicated a stronger information flow from the HR to the breath rate than vice versa. This finding was consistent with the observation that the patient breathed in bursts, which seemed to occur whenever HR crossed some assumed threshold.

The normalized version of the TE [52] was applied to short-term HR, SBP and RESP time series of healthy humans (resting supine position, upright position after head-up tilt). They found that in the supine position, a significant amount of information was unidirectionally transferred from RESP to HR and SBP assumed to be the effect of RESP on the vascular system (mechanical nature and sinus node activity) known as RSA. In addition, they stated that the decrease of information flow from RESP to RR in the upright position is likely to reflect the dampening of RSA consequent to the withdrawal of parasympathetic nervous activity occurring in head-up tilt. The information transfer between RR and SBP was relatively low and balanced over the two causal directions with a slight prevalence from RR to SBP. These findings were interpreted as the FF mechanisms operating in the direction from RR to SBP (mechanical modulation following Starling’s law and Windkessel effects), which is equally important as the well-known FB baroreflex mechanisms. In the upright position, the information transfer from RESP to SBP
remained substantially unchanged, whereas it decreased significantly in the direction from RESP to RR. In the cardiovascular regulatory loop, they could observe a substantial enhancement of the flow of information on the FB direction from SBP to RR, leading to an unbalancing of the RR–SBP regulation. This finding was ascribed to an enhancement of the baroreflex control of HR that occurs concomitantly with the tilt-induced activation of the sympathetic nervous system.

Verdes [48] used the concept of TE for a multivariate time-series analysis of HR, chest volume or respiration force (RF) and blood oxygen concentration (Ox) records of patients suffering from sleep apnoea. They observed that Ox was sensibly affecting the dynamics of respiration in accordance with the mechanism of baroreceptor control. In addition, they found that both dynamics (HR, RF) significantly affected the level of Ox. Furthermore, they showed that a bivariate analysis of HR and RF was insufficient and that it was essential that their interaction with the baroreceptor control system had to be taken into account. Finally, they demonstrated that RSA seemed to be blocked in the case of sleep apnoea.

(d) Symbolization

Applying the JSD2 method, we investigated unmedicated patients suffering from schizophrenia (Szo) and healthy controls (Con) to quantify short-term (30 min, resting conditions) nonlinear bivariate cardiovascular couplings. Indices derived from the diagonals of the word distribution density matrix revealed decreased symmetric (SumSym) word types in Szo ($p < 0.00055$) in comparison with Con and significant increased diametric (SumDiam) in Szo ($p < 0.00055$; figure 2a,b). The symmetric word types (the predominant word type combinations for both groups) represent baroreflex-like response patterns, whereas diametric word types represent oppositional baroreflex patterns [66]. These findings demonstrate that significantly decreased nonlinear bivariate word type combinations (couplings) are a sign of impaired nonlinear baroreflex-mediated coupling patterns of cardiovascular regulation in Szo in comparison with Con. Thus, the results might be interpreted as a decreased parasympathetic tone in Szo.

In addition, by applying JSD2, we investigated Szo and Con to characterize short-term nonlinear bivariate cardiorespiratory couplings (30 min, resting conditions, HR, RESP) and to obtain information on how respiration can influence autonomic regulatory processes in Szo. We found that RESP-related word types were highly significant altered between both groups and that two prominent RESP word types (111, 000) occurred mostly in combination with all other HR word types in both Szo and Con (RESP000: Szo = 25%, Con = 33%; RESP111: Szo = 32%, Con = 37%). These RESP patterns occurred independently from the HR patterns and reflected a sustained behaviour that was uncoupled from HR. All other RESP word types (from the second to the seventh column; figure 2c,d) occurred more frequently in combination with all other HR word types for Szo compared with Con. This result pointed to an increased complexity of short-term cardiorespiratory regulation and might further indicate a more uncoupled and more unspecific cardiorespiratory regulation in Szo. The specific occurrences of predominant RESP patterns could be further considered as a marker for a restricted RSA and were also found in studies by [76,77]. In summary, these results suggested a diminished vagal modulation at the brain stem level or its suppression from higher regulatory centres in Szo.

Baumert et al. [54] applied the concept of JSD2 to analyse HR and SBP time series of pregnant and control women. They found significant differences in five bivariate word types between pregnant and non-pregnant women and concluded that pregnancy has an impact on autonomic control that can be revealed by JSD2 analysis.

To analyse and quantify cardiovascular time-delayed couplings, Suhrbier et al. [57] applied SCTs, to investigate HR and SBP time series during different sleep stages in healthy controls as in normotensive and hypertensive patients with sleep apnoea. They found that SCTs revealed significant different cardiovascular mechanisms, not only between the deep sleep and the other sleep stages, but also between healthy subjects and patients. They found that SCTs was more specific in detecting time delays of directional interactions than standard coupling approaches. They concluded that SCTs may help to determine pathological changes in
cardiovascular regulation as well as the effects of continuous positive airway pressure therapy on the cardiovascular system.

(e) Phase synchronization

Rosenblum et al. [58] investigated the direction in cardiorespiratory coupling (directionality index) in healthy infants. They demonstrated that within the first 6 months of life, there is a tendency to change from a roughly symmetric interaction to a nearly unidirectional one (respiration drives the heart). The directionality index revealed a dependence on the frequency of respiration $f_r$. In the case of $f_r < 0.5$ Hz, the interaction occurred predominantly in one direction, indicating the possible existence of two regimes of interaction. They suggested that the cardiac influence on respiration is weak and frequency independent, whereas the coupling from the respiration to HR is similar to a low-pass filter. They assumed that the dependence of $f_r$ can be explained in terms of two classes of frequency response properties of the pathway that transmits information from the central nervous system to the heart. Thus, at very young ages (during the first week of life), the value of the cut-off frequency appears to be lower. Maturation of the central nervous system as well as cardiovascular adaptation to extrauterine conditions, such as closure of foetal shunts, may account for this finding.
Bartsch et al. [78] studied cardiorespiratory phase synchronization during different physiological stages of sleep (wakefulness, rapid-eye movement (REM) sleep and non-REM sleep) in healthy subjects. They observed that during REM sleep, cardiorespiratory synchronization is suppressed by approximately a factor of 3 compared with wakefulness. On the other hand, during non-REM sleep, it was enhanced by a factor of 2.4, again compared with wakefulness. In addition, they found that these significant differences between synchronization in REM and non-REM sleep are very stable and occur in the same way for males and females, independent of age and body mass index. Their results suggested that the synchronization was mainly due to a weak influence of the breathing oscillator upon the HR oscillator, and that this was disturbed in the presence of long-term correlated noise and superimposed on the activity of higher brain regions during REM sleep.

Schäfer et al. [62] investigated cardiorespiratory synchronization of healthy volunteer high-performance swimmers with a CRS. They found 5:2 locking between the respiratory frequency and the HR during a period of about 300 s, then after a transition period, a regime of 3:1 phase locking was found for about 20 min. Their analysis revealed that subjects with the strongest synchronization had no remarkable RSA, whereas subjects with the highest RSA exhibited no synchronization. Finally, they concluded that phase locking of respiratory and the cardiac rhythms, and respiratory modulation of HR, were two competing aspects of cardiorespiratory interaction.

Censi et al. [79] analysed coupling patterns between RESP and spontaneous rhythms of HRV and blood pressure variability signals applying recurrence quantification analysis to cardiovascular and respiratory data of normal subjects (paced breathing at 0.25, 0.20 and 0.13 Hz). They observed various types of synchronization phenomena, ranging from fully de-coupled to completely synchronized signals. Higher degrees of nonlinear coupling were observed when the respiratory frequency was close to the spontaneous LF rhythm (0.13 Hz), or almost twice the LF frequency (0.2 Hz), whereas weaker coupling was observed when the respiratory frequency was 0.25 Hz.

The above presented studies are, however, not comparable to each other. The reason for this are the different objectives, different investigated diseases, different numbers of enrolled patients, different age–gender distributions of the patients, different time-series lengths, different measuring conditions and different study designs that avoid a direct comparison and a generalized assessment of the results as well as of the applicability.

4. Summary

In this paper, we have presented the most frequently applied linear and nonlinear approaches used to quantify direct and indirect couplings and the direction (driver–response relationship) of these interactions in the cardiovascular and cardiorespiratory systems. We outlined their basic theoretical background, basic requirements for application, features, influencing factors and present application examples in the medical field. One should consider that the application of these approaches cannot be restricted to a single favourable one because this mainly depends on the problem to be solved. There exists no generally superior approach that can solve all problems.

There are some important points to consider when applying these methods to cardiovascular and cardiorespiratory time-series analysis.

— The cardiovascular and cardiorespiratory systems are complex physiological systems interacting in direct or indirect ways. For the investigation of these systems, bivariate approaches are commonly applied. However, it can be assumed that the application of multivariate approaches will be increasingly used instead of bivariate ones because they improve the characterization of causal or non-causal interrelationships.

— Cardiovascular and cardiorespiratory time series (e.g. electrocardiogram, systolic blood pressure and DBP, blood flow, plethysmogram, respiratory frequency, respiration flow) are often noisy and non-stationary or only stationary over short periods.
The assessment of coupling and causality can be performed by applying either linear or nonlinear time-series analysis approaches. While nonlinear methods study complex signal interactions, linear methods favour the frequency domain representation of biological signals (characterization of connectivity between specific oscillatory components).

The method-specific characteristics used to differentiate between direct and indirect information flows and uni- and bidirectional couplings and to assess the coupling strength are not uniformly combined in one approach (table 1).

PDC analyses are most frequently carried out between signals that are scaled on identical sizing systems (e.g. electroencephalogram on a microvolt scale). By contrast, respiration, HRV and SBP are scaled in different metric sizing systems. For such cases, it is recommended to use the generalized PDC as a scale invariant interaction measure [23]. In particular, the PDC can take values arbitrarily close to either one or zero if the scale of the target variable is changed accordingly [80].

The application of time-variant multivariate analysis approaches (i.e. tvPDC) will contribute to an improved understanding of time-variant relationships of the cardiovascular and cardiorespiratory system [23].

For the investigation of pathophysiological conditions, we have to consider that these approaches are partly not validated because they were often applied either on experimental data, in healthy subjects or in subjects with a specific disease. Competitive and representative studies are still missing.

It would be desirable to standardize these approaches in the near future, thus making it possible for researchers to select the right technique for their application. To make the application of coupling approaches more user friendly, the following issues should be addressed:

- reduction of the degrees of freedom;
- reduction and standardization of preconditions (e.g. pre-processing steps, parameter settings, time-series length, model order selection, significance level determination); and
- method validation on large sample sizes of physiological and pathophysiological cases.

In conclusion, linear and nonlinear approaches used to quantify direct or indirect, as well as causal or non-causal relationships, might provide new insights into alterations of the cardiovascular and cardiorespiratory system and possibly will lead to improved knowledge of the interacting regulatory mechanisms under different physiological and pathophysiological conditions. These approaches represent promising tools for detecting information flows in a multivariate sense. They also might be able to provide additional prognostic information in the medical field and might overcome or at least complement other traditional univariate analysis techniques.

The interest in coupling analyses of the cardiovascular and cardiorespiratory system has been growing considerably, and this will therefore lead to an increasing amount of additional applications in the near future, improving our knowledge about interacting regulatory subsystems.

Acknowledgements. This study was partly supported by a grant from the University of Applied Sciences Jena and by the Deutsche Forschungsgemeinschaft DFG (Vo505/8-2). All authors hereby disclose any and all commercial associations that might pose a conflict of interest in connection with the manuscript.

References


4. Eckberg DL. 2009 Point-counterpoint: respiratory sinus arrhythmia is due to a central mechanism versus respiratory sinus arrhythmia is due to the baroreflex mechanism. J. Appl. Physiol. 106, 1740–1742. (doi:10.1152/japplphysiol.91107.2008)


