Improved Repeatability of the Estimation of Pulsatility of Inferior Vena Cava

Improved repeatability of the estimation of pulsatility of inferior vena cava

Luca Mesin[a,*], Tatiana Giovinazzo[a], Simone D'Alessandro[a], Silvestro Roatta[b], Alessandro Raviolo[c], Flavia Chiacchiarini[c], Massimo Porta[c], Paolo Pasquero[c]

[a]*Mathematical Biology and Physiology, Department of Electronics and Telecommunications, Politecnico di Torino, Torino, Italy*
[b]*Integrative Physiology Lab, Department of Neuroscience, University of Torino, Torino, Italy*
[c]*Department of Medical Sciences, University of Torino, Torino, Italy*

Abstract

The inferior vena cava (IVC) shows variations of cross-section over time (referred to as pulsatility) induced by different stimulations, like as breathing and heartbeats. The amplitude of these pulsations is affected by the volume status of the patient and can be investigated by ultrasound (US) measurements. Thus, the caval index (CI), i.e., an index of pulsatility of IVC based on US visualization, was proposed as a non-invasive indirect measurement of the volume status. However, the methodology is not standardized, operator-dependent and affected by movements of the vein and non-uniform pulsatility. We introduced a software that processes a B-mode US video-clip to track IVC movements and estimate the CI on an entire portion of the vein. This new method is here compared to the standard approach in terms

*Corresponding Author: Luca Mesin, Dipartimento di Elettronica e Telecomunicazioni, Politecnico di Torino, Corso Duca degli Abruzzi, 24 - 10129 Torino - Italy; Email, luca.mesin@polito.it; Phone, +39 011.090.4085

of repeatability of the estimated CI. Furthermore, the cardiac and respiratory contributions to IVC pulsatility are separated, avoiding the confounding effects of their asynchronous summation to provide two additional selective pulsatility indexes. We report on the variability of CI estimation over the following factors: different respiratory cycles or heart pulsations, longitudinal sections of the vein and intra/inter observer reproducibility. Our method allows to reduce the variability of CI assessment, providing a step toward its standardization.

Keywords: Inferior vena cava, Ultrasound, Tracking, Repeatability, Volume status

---

Introduction

Pulsatility of the diameter of the inferior vena cava (IVC), estimated from ultrasound (US) measurements, is a non-invasive procedure, widely adopted to assess the intravascular volume status both in healthy subjects and conditions of altered volemic status in patients. However, measurement techniques are not standardized (Wallace et al. (2010)), as both recordings along longitudinal (Barbier et al. (2004); Brennan et al. (2006); Fields et al. (2011); Feissel et al. (2004); Grant et al. (1980); Kircher et al. (1990); Lyon et al. (2005); Moreno et al. (2019)) or transversal sections (Blehar et al. (2009); Chen et al. (2010); Moreno et al. (2019)) of the vein are used. Different recommendations have been proposed on where to measure the vein diameter along a longitudinal section (Wallace et al. (2010); Resnick et al. (2011)). However, since the pulsatility of the vessel is not uniform along its longitudinal axis (Mesin et al. (2015, 2019b)), CI values vary considerably in the literature in both healthy and pathologic conditions and, as a result, diagnostic recommendations are also non homogeneous (Zhang et al. (2014)).

The pulsations of the vessel during the respiratory cycle are used to measure the caval index (CI, Blehar et al. (2012)). However, the movements of the vein relative to the transducer during the respiratory cycle give an additional contribution to the variability of CI. Indeed, M-mode registration allows to compute the vein diameter along a fixed line at the end of inspiration and expiration, but, since the IVC moves during respiration, the diameters end up being taken at different points, introducing a possible bias. This is particularly relevant if the vein has an irregular shape, with a variable cross-sectional area (Lichtenstein (2005)) or if the angle between the

M-mode line and the vein changes considerably during its movements. In addition, respiration cycles may differ between each other and change among subjects (e.g., breathing can be diaphragmatic, thoracic or a combination of both), inducing changes in the IVC dynamics (Kimura et al. (2011)). In order to minimize movements of the vein during respiration, variations of the IVC section was investigated during voluntary apnoea, thus bringing forward the effect of cardiac activity on IVC pulsatility (Folino et al. (2017); Nakamura et al. (2013)), which is otherwise poorly detectable on M-mode representation. However, this technique cannot be easily applied in clinics.

We reported on successfully tracking IVC movements in long-axis US scans while estimating its diameter in each frame, along a direction moving together with the vein (Mesin et al. (2015)). This method has a lower computational cost than other advanced image processing techniques applied to US images (Yang et al. (2008); Yeung et al. (1998); Krupa et al. (2007)) and provides a more precise estimation of the IVC local pulsatility with respect to standard measurements, based on a fixed M-mode line (Mesin et al. (2015)). However, a possible problem is that pulsatility along a single section of the IVC may be not representative of the dynamics of the whole vessel. Some parts of the vein are anchored to nearby structures (e.g., the diaphragm or vein inlets) and show smaller pulsatility than other portions. For example, lower pulsatility was reported at the level of the diaphragm compared to more caudal sites (Wallace et al. (2010)). These observations were confirmed in Mesin et al. (2015) (Figure 9), showing that diameter variations along distinct directions (moving together with the vein) resulted in considerably different pulsatility. Lack of consensus about where to measure

diameters (Wallace et al. (2010); Resnick et al. (2011)) and the non-uniform behaviour of the vessel are likely to contribute to the non-homogeneous assessments of IVC pulsatility in the literature (Weekes et al. (2012)). Thus, we recently proposed a new algorithm that tracks the movements and computes the diameter of different sections of a whole portion of the IVC (Mesin et al. (2019b)). Here, we compare this innovative method to the standard approach and report on the repeatability of information extracted from different measurements on the same subjects.

Materials and Methods

Automated detection of the IVC borders

US video-clips were processed using the algorithm proposed in Mesin et al. (2019b), which allows to obtain a continuous measurement of IVC borders along an entire portion of the vessel after compensating for possible movements. The algorithm was implemented in MATLAB R2018a (The Mathworks, Natick, Massachusetts, USA).

The user is asked to indicate the location of the vein in the first frame (Figure 1A). Moreover, as shown in Figure 1B, on the same frame, he chooses two reference points to be tracked (to account for IVC movements and deformations) and the most proximal/distal sections (defining the portion of the IVC of interest, which was between the confluence of the hepatic veins into the IVC and the caudate lobe of the liver). Finally, the locations of the borders of the vein along the most proximal line are indicated. The software is then ready to process the video-clip. It distributes uniformly N lines in the portion of IVC indicated by the user (N=21 in this paper) and automat-

ically detects the borders of the vein along these lines (Figure 1C). For each frame, the location and direction of the N lines are updated depending on the movements of the reference points. In this way, the superior and inferior borders of the vein are estimated in the IVC portion of interest.

Subjects

US data were recorded from 10 healthy volunteers (5 females, 5 males; mean±std age 30±13 years, height 172±12 cm, weight 63±11 kg) with a SonoSite M-Turbo system (SonoSite, Bothell, USA[1]; frame rate 30 Hz, resolution of about 0.42 mm per pixel, 256 gray levels) equipped with a convex 2-5 MHz probe. Two-dimensional (B-mode) longitudinal views of the IVC were taken with a subxifoideal approach, with the subject in the supine position during relaxed normal breathing. The study was approved by the Ethics Committee of the University of Turin and complies with the principles of the Declaration of Helsinki. All subjects provided written informed consent for the collection of data and subsequent analysis.

Experimental set-up and protocol

The experimental protocol is illustrated in Figure 2. Three operators performed the US scans: one expert (PP), one in training (AR) and one beginner (FC), with balanced arrangement of their order. An operator started by taking 3 measurements of IVC diameters (as defined below) using standard methodology in M-mode. Then, a 15s video-clip was recorded, allowing for at least three respiratory cycles. After the first recording, the subject was

[1]M-Turbo Ultrasound System - User manual, http://www.sonosite.com/downloads/M-Turbo_UG_P07662.pdf

asked to stand up for one minute to minimize any changes of the IVC due to remaining in the supine position for a prolonged time (Folino et al. (2017)). Then, the subject was asked to lie down again supine and a new acquisition was taken by a second operator and, after standing up again, by a third one. The whole procedure was repeated a second time, obtaining six video-clips for each subject.

Indexes extracted from the data

Different indexes were taken from each measurement, in order to test their repeatability. Three manual measurements in M-mode were taken before registering the video-clips. The operator chose three respiratory cycles. For each of them, the maximum and minimum vein diameters (D_{max} and D_{min}, respectively) were indicated, and the (manual) CI was computed as

$$CI = \frac{D_{max} - D_{min}}{D_{max}} \tag{1}$$

The video-clips were then processed to estimate the IVC borders as detailed above. Notice that the position of each point of the border is indicated by time series (location along x and y directions, one value per frame). These time series were low pass filtered with a 4 Hz cut-off, in order to remove high frequency and quantization noises (this filter and the ones mentioned below were of Butterworth type, order 4 and used in both directions to remove phase distortion and delay, Mesin et al. (2019b)). Then, the borders of the IVC were estimated from the confluence of the hepatic veins into IVC to 4 cm in distal direction (Figure 1D). Specifically, from the estimated borders, the IVC midline was computed. It was then approximated by a parabolic function. The location of the confluence of the hepatic veins into the IVC was

indicated by the user (SA, who was not an echographer) on the first frame of the video-clip. This point was orthogonally projected on the IVC midline and represented the starting point from which other 4 points were automatically estimated, with 1 cm curvilinear distance from each other along the IVC midline. Thus, 5 points were obtained, 0 to 4 cm distant from the confluence of the hepatic veins into the IVC, projected on the midline of the vein. Then, the sections orthogonal to the IVC midline passing from each such points were considered (Mesin et al. (2019b); Pasquero et al. (2015)) and the IVC diameters in these sections were computed by interpolation from the estimated vein borders (see Mesin et al. (2019b) for details). These five diameters are further considered in the following.

The pulsatility of the IVC in each section was described by the (automated) CI, defined as

$$CI_{auto} = \frac{\max(D) - \min(D)}{\max(D)} \tag{2}$$

where D indicates the estimated diameter time series (in a specific section). Local maxima and minima were computed for each respiration cycle (Figure 3A). Thus, an estimate of CI was obtained for each respiratory cycle and for each section considered. As in the case of the manual CI estimation, the CIs of 3 respiratory cycles were selected. In the cases in which more than 3 cycles were present in the video-clip, the CIs closer to their mean across different cycles were selected. After testing the repeatability across respiration cycles, the estimated CIs were averaged. A CI accounting for the overall pulsatility of the considered portion of the vein was also considered (indicated as CI_{global}): it was obtained by averaging the estimates across different sections.

Additional indexes of pulsatility were obtained after further processing the diameter time series estimated by our software. The vein dynamics was considered as the sum of two components, reflecting the stimulation induced by respiration and heartbeat (Mesin et al. (2019a)). The two components were separated as follows: the effect of respiration was computed by low pass filtering the whole diameter time series with a cut-off frequency of 0.4 Hz. The cardiac contribution was computed by high pass filtering the whole diameter time series with a cut-off frequency of 0.8 Hz. Then, the following additional indexes were estimated, as shown in Figure 3.

- The respiratory caval index (RCI), applying the same formula (2) to the respiration component only.
- The cardiac caval index (CCI), applying the same formula (2) to the cardiac component only.

Also for these two indexes, stimulation cycles were selected: 3 respiration cycles and 10 heartbeats were included. Moreover, the subscript global was added to indicate their average across different sections (RCI_{global} and CCI_{global}).

Assessment of repeatability and discriminability

Different indicators were used to assess the repeatability of each index (manual and automated CI, CCI, RCI) extracted from the 6 measurements performed by the operators.

- Coefficient of variation (CoV), defined as the ratio between the standard deviation and the mean of the estimates. It gives an indication

of the agreement of an index extracted from different measurements in the same conditions. It was used to test variations due to different respiration cycles, different sections and different experimental sessions (intra- and inter-operator).

- Intraclass correlation coefficient (ICC). It is defined as

$$ICC = \frac{var(S)}{var(S) + var(M) + var(E)} \tag{3}$$

where $var(S)$, $var(M)$ and $var(E)$ indicate the variability due to either different subjects or measurements (i.e., experimental sessions) and the residual error, respectively (Bartko (1966)). It was used to test intra- and inter-operator variability. Notice that the ICC is equal to 1 if the whole variability is due to the differences between subjects, whereas no variability is due neither to the measurements nor to errors (always the same value is obtained).

An index of discrimination was also studied, in order to avoid the possible case in which an index is repeatable only because it always takes similar values, even considering different subjects. The Fisher ratio was used. It measures the linear discrimination between two sets of values as

$$FR = \frac{(\mu_1 - \mu_2)^2}{\sigma_1^2 + \sigma_2^2} \tag{4}$$

where μ_k and σ_k^2 (with $k = 1, 2$) are the mean and the variance of the kth sets, respectively. The sets to be compared were constituted by the 6 values of a specific index extracted from the different measurements on each subject. The mean of the Fisher ratios measuring the discrimination of each pair of subjects was used as overall discriminability indicator.

Finally, analysis of variance (ANOVA) was used to investigate the different sources of variability. The manual CI and CI_{global} (i.e., the automated CI obtained averaging across different sections) were compared with an ANOVA (normality of residuals was assessed by Lilliefors test), investigating the variability induced by the following factors: subject, operator, repetition and respiration cycle. Some paired post-hoc tests for significant variations among couples of variables were performed by either t-tests or Wilkoxon signed rank tests (depending on the output of the Lilliefors normality test). The significance level was set to $p = 0.05$.

Summary of investigated indexes

The following indexes are considered.

1. Manual CI, which is a variable depending on the following factors: respiration cycle (3 cycles considered), subject (10 volunteers) and experimental session (6 sections, which could be further split into 3 operators repeating twice the experiment). The average across the respiration cycles was also considered.
2. CI_{auto}, RCI_{auto} and CCI_{auto}, depending on the following factors: respiration cycle (3 cycles considered) or heartbeat in the case of CCI_{auto} (10 beats considered), subject, section (5 locations, measured in terms of the distance from the hepatic veins) and experimental session. The average across the respiration cycles/heartbeats was also considered.
3. CI_{global}, RCI_{global} and CCI_{global}, obtained by averaging the previous indexes across the sections (obtaining a global index for the vein tract under study), so that they depend on respiration cycle or heartbeat (the

latter in the case of CCI_{global}), subject and experimental session. The average across the respiration cycles/heartbeats was also considered.

Results

Figures 4-7 show different contributions to the variability of the estimates of some indexes reflecting the pulsatility of IVC. For clarity, a single source of variability is considered in each figure (respiration, longitudinal section, experimental session and intra-/inter-operator variability, respectively) and only some indexes are shown. The whole database is fully explored with the statistical analysis shown in Tables 1-3.

Variability of CI in subsequent breaths

Figure 4A shows the changes in IVC diameter exhibited in a representative subject at rest. The tracings refer to different IVC sections, located at 0, 2 and 4 cm distal to the confluence of hepatic veins into the IVC. Notice that the sections exhibit different average diameter and different amplitude of oscillatory components of cardiac and respiratory origin. For example, at the confluence of the hepatic vein, the algorithm estimated different respiration cycles with CIs varying in the range 18%-28% and with a CoV equal to 19% (indicating the variability of the CI estimations across different respiration cycles). Figure 4B shows the CoV of the estimations of the CIs assessed on single respiratory cycles, extracted from the whole dataset. This CoV, expressing the variability observed over consecutive respiratory cycles, was calculated for all trials (obtaining 60 values of CI_{auto}[2]) and for each IVC

[2]60 values of CoV are obtained as we considered 10 subjects for 6 experimental sessions.

section. In addition, for comparison, the same figure also includes the CoV of CI_{global} and CI_{manual}. Notice that the median variability with respect to different respiration cycles (in terms of CoV) is about 15% when considering the standard (manual) method, about 5% when considering single sections tracked by the automated method Mesin et al. (2015) (CI_{auto}) and lower than 3% when considering the global CI (averaged over all IVC sections, CI_{global}; Wilkoxon signed rank test indicated that the CoV of manual and global CI were statistically different).

Variability of CI with longitudinal position

All the following figures show indexes obtained by averaging estimations on different respiration cycles.

Figure 5 shows the variability of CI estimation across different sections along the IVC. The dependence of IVC pulsatility along the longitudinal position is visible in 5A for the different subjects (CI_{auto} is shown averaged over all 6 experimental sessions). Notice that there is no location showing larger or lower pulsatility, being the patterns very different among the subjects. The dependence of CI on position can be relevant: e.g., in subject number 7, CI_{auto} decreases from about 40% to 10%, moving caudally by 3 cm from the confluence of the hepatic veins into IVC; conversely, in subject 8, CI increases from about 50% to 70%, over the same distance.

The variability of CI_{auto} along the considered IVC tract was quantified by its CoV. One estimation of CoV was obtained for each experimental session, obtaining 6 values for each subject which are shown in Figure 5B. On average, it is as high as 30% (which means that the range of variation is larger than

the mean value[3]).

Variability of CI, RCI and CCI over the different experimental sessions

For the different indexes (now including also RCI and CCI), the CoV was computed over the 6 experimental sessions, thus providing a measure of repeatability of the assessment for each subject.

This evaluation was conducted separately for the different positions along the IVC in order to compare automated and manual assessments. As illustrated in Figure 6, none of the sections along the IVC exhibits a CoV significantly smaller than the others. Moreover, it can be observed that i) manual and automated (over single section) assessments have similar variability (6A); ii) removing the respiratory component improves repeatability (6B and 6D); iii) filtering out the cardiac component does not improve repeatability (6C and 6D); iv) a relevant reduction in CoV of CI_{auto} is obtained by calculating the CI over the entire longitudinal portion of IVC (CI_{global}). Statistically significant differences were found between the manual CI and CCI_{global} and between CI_{global} and RCI_{global}.

Intra- and inter-operator variability of CI assessment

Figure 7 shows a comparison between the CoV of manual CI and global automated estimation (CI_{global}). Intra-operator variability was computed using the two repetitions of the measurement by the specific operator considered. Inter-operator variability was computed from the average CI obtained

[3]Assume a Gaussian distribution of the estimates of CI along the sections: the range is about 4 times the standard deviation of the estimates. Thus, if CoV is 30%, the range is about 120% of the mean.

by the operators (averaging the two repeated measurements) from each subject. The spread of the estimates obtained from the same subject was lower for the automated method for 9 subjects out of 10 (a statistically significant difference is indicated by the Wilcoxon signed rank test applied to the standard deviations of the estimates obtained using either the manual or the automated CIs; the CoV of manual and global CI were not statistically different, instead). Most of the repeated manual measurements of each operator were quite similar (mean intra-operator CoV equal to 28%), but the estimations varied a lot among different operators (mean inter-operator CoV equal to 35%). The automated measurements were more stable and showed similar intra- and inter-operator variabilities (mean CoV equal to 24 and 18%, respectively).

tistical analysis

The statistical analysis of our data is shown in Tables 1-3. Table 1 shows the ANOVA, comparing the manual CI and CI_{global}. Notice that the total variability of CI is larger when using the standard clinical approach. Moreover, as indicated by the F statistics, a slightly higher percentage variability is obtained considering different subjects when using the automated method instead of the standard one (so that a better discrimination of different subjects can be obtained using the automated algorithm). On the other hand, a lower variability is obtained using the automated method in different experimental sessions (when pooling together the factors repetition and operator, results not shown) and respiration cycles (even if the variations induced by the respiration cycle are not significant). Splitting the experimental sessions into the factors repetition and operator, we notice that the variations on

different repetitions were quite small (and not significant), whereas larger (significant) differences were found considering different operators (in line with the inter- and intra-operator CoV discussed above). Moreover, smaller variations over different repetitions were found for the standard approach, whereas those induced by different operators were smaller for the automated approach. Thus, the automated approach provides measurements that are more stable across different operators, whereas, by the standard approach, the echographers obtained twice similar values, which were however different from those of the colleagues, indicating a possible bias.

Tables 2 and 3 show respectively the ICC and the Fisher ratio of the caval indexes computed either by the standard or the automated method (manual CI, CI_{global}, CCI_{global} and RCI_{global}). Intra-operator values were computed considering only the estimates obtained by each operator, separately; inter-operator values were obtained by grouping together the estimates of the same operator. Notice that the most experienced operator obtained quite high values of ICC and Fisher ratio, considering both the standard method and the indexes extracted from the video-clips that he recorded. The CIs measured with the standard method had a correlation with those estimated by our software using the corresponding video-clips (i.e., those registered after the M-mode assessment) which was found to be related to the experience: FC, AR and PP (i.e., the operators in order of increased experience) showed a

correlation coefficient of 36.2%, 58.1% and 70.8%, respectively[4]. The second operator in order of experience (AR) had a personal technique to measure the CI in M-mode (further commented in the Discussion section) which allowed him to get similar values in repeated measurements by the standard approach, so that his ICC and Fisher ratio are quite high. Notice that the estimates of CI obtained by the automated method are more consistent across different operators (inter-operator ICC about 70%, whereas it is about 61% for the standard estimation). High values of ICC were obtained also for the estimation of CCI, lower values for RCI (in line with Figure 6). Notice also that the video-clips acquired by the most experienced operator allowed to get more repeatable estimates of the automated indexes (this indicates the importance of acquiring good video-clips to get repeatable results also from the automated processing). The results on intraclass correlation are in line with those shown by the Fisher ratio: indeed, a larger repeatability of the estimation of the pulsatility of each subject allows to better discriminate between different subjects.

[4]The following definition of correlation coefficient is used:

$$C = \frac{\sum_n (x[n] - \bar{x})(y[n] - \bar{y})}{\sqrt{\sum_n (x[n] - \bar{x})^2 \sum_m (y[m] - \bar{y})^2}} \tag{5}$$

where $x[n]$, $y[n]$ are the series to be compared and $\bar{x}$, $\bar{y}$ are their means.

Discussion

Summary

For the first time, repeatability of standard CI estimations was assessed in a group of healthy subjects, the results indicating rather poor values in terms of both intra- (mean CoV=28%, ICC in the range 49-82%) and inter-operator variability (mean CoV=41%, ICC=61.5%).

With the help of a semi-automated algorithm analysing 15s lasting video-clips of the IVC in long axis, it was possible to show

1. high variability of the CI over the respiratory pattern (CoV about 5%, whereas it is about 15% for the standard approach),
2. high variability of the CI depending on the longitudinal site of assessment (median of CoV ranging among 10 and 70% for different subjects, after averaging across respiration cycles).

By 1) averaging over consecutive breathing cycles, 2) tracking IVC longitudinal movements and 3) averaging over multiple longitudinal sites, the algorithm offers a more objective and reliable measurement of the CI (here called global CI), reducing the overall variability (intra- and inter-operator mean CoV equal to 24% and 18%, respectively; ICC=70.4%). In addition, the identification of the respiratory and the cardiac oscillatory components may provide new insights and possibilities for the analysis of IVC dynamics, with repeatability performances close to those of the standard CI and global CI, respectively.

Discussion of different sources of variability

The pulsatility of the IVC by the CI estimation is widely used to assess the volemic status in different clinical conditions. However, the measurements are not standardized and the recommendations given in the literature are not univocal (Zhang et al. (2014)).

To the best of our knowledge, the repeatability of the estimation of the IVC pulsatility has never been assessed previously. However, it would be a very important information, as it could provide an indication of the limits of the method to discriminate the volume status of different patients or in the follow up. In this paper, we explored different sources of variability that may affect the assessment of IVC pulsatility.

- Variation of the depth and modality of respiration, which induce different IVC pulsatility for each breath cycle. Notice that controlling the respiration cycle (e.g., by a spirometer, even if only the respiration depth, not the modality, could be controlled) could possibly reduce this source of variability. Indeed, in the case of mechanically ventilated patients, the respiration cycles are regular and the dynamics of the IVC diameter was found to be useful to detect fluid responsiveness (Feissel et al. (2004)). As an alternative, measuring the pulsatility during a short apnoea, thus caused by the heartbeats only (Folino et al. (2017); Nakamura et al. (2013)), could help to standardize the measurement.

- Variations of the pulsatility in different sections of the vein. These variations were noticed both in longitudinal (Mesin et al. (2015, 2019b)) and transversal scans (Blehar et al. (2012)).

- Variations introduced by the operator. In different measurements, the investigated 2D section can be slightly different. Furthermore, the US probe handled by the operator must follow the movements of the patient during respiration: the ability to follow the movement without affecting the measurement depends on the level of experience of the operator.

In addition, there are variations of the investigated IVC section, due to movements of the vein during an M-mode measurement (as the M-mode registration fixes the considered section in space). Consider that both translation and rotation of the vein with respect to the studied direction are expected to occur in general. The former induces an error in the estimated diameter dependent on the shape of the vein, while rotation affects the estimated diameter even if the vein is a perfect cylinder. The problem is reflected by an error in the estimation of pulsatility, which depends on the range of movements and anatomy of the vein (Mesin et al. (2015)). In this paper, such a problem affected only manual estimations. The automated IVC tracking (introduced in Mesin et al. (2015, 2019b)) allows to remove this source of uncertainty.

The other three sources of variation mentioned above were investigated in this study, considering both the standard manual measurements and the automated estimations provided by the algorithm proposed in Mesin et al. (2019b), which estimates the IVC sections in a whole portion of the vein. Figures 4-7 show repeatability in terms of CoV, so that the variation is measured as the standard deviation of the estimates normalized with respect to their mean.

- The CI (as a measurement of IVC pulsatility) in different respiration cycles had median variation which was about the 15%, 5% and 3% of the mean value, for the manual and the automated methods respectively, either considering a single section or averaging across a portion of the vein (Figure 4). A large variability among different subjects was observed, with the largest variations being about the 90% and the 30%, for the manual and the global automated method (averaging across sections), respectively. The repeatability is much larger for the automated method than considering the clinical standard. For the following discussion, this variability was removed considering the average CI among respiration cycles (for both the manual and the automated method).

- A large variation of CI was observed when considering different sections along the IVC (Figure 5), confirming that the vein pulsations vary a lot, depending on anatomical properties of the vein and of the surrounding tissues (e.g., the presence of anchoring sites). The sections were studied using the automated method, which tracked their motion. The average CoV was about 40%, with great variations among subjects (the one showing the largest differences among sections showed a CoV of about 70%). No section can be considered better than others in terms of repeatability of the estimations: the best one varies among the subjects and also considering different measurements on the same subject. Moreover, a large variability of CI was observed among subjects, without a clear trend of pulsatility when going in proximal or distal direction along the considered longitudinal section of the IVC (extending 4 cm distal from the confluence of the hepatic veins). The

great variability of IVC pulsatility along the cranio-caudal direction can lead to misinterpretation of the overall dynamics of the IVC.

- Considering the measurements of different echographers, we observed a large variability, both among experimental sections (Figure 6) and intra-/inter-operators (Figure 7). The operators had different experience: more than 20 years (PP), 2 years (AR) and less than 1 year (FC). Their procedures in taking the manual measurements were quite different.

  - PP tried to select a direction orthogonal to the IVC midline (Pasquero et al. (2015)). In the average, the measuring site was 2.4 cm from the confluence of the hepatic veins, i.e., close to the centre of the considered portion of IVC.

  - AR took the measurement quite close to the diaphragm, in the average 1.7 cm from the confluence of the hepatic veins (25% of times, the measuring site was at a distance from the confluence of the hepatic veins lower than 1 cm). This procedure helped him in getting stable measurements in different experiments, as there are anatomical references which could be easily found. However, in that region, the vein pulsatility is affected by anchoring tissues and the blood flow from the hepatic vein, so that the accuracy of the measurement could be questionable.

  - FC showed a lower experience than the colleagues, as her measurements required longer time and efforts. In the average, the measuring site was 2.7 cm from the confluence of the hepatic veins

and the distribution of chosen sites was the most dispersed among the colleagues (std of about 1.4 cm, whereas it was 0.94 and 1.15 for PP and AR, respectively).

The ANOVA allows to interpret the different sources of uncertainty in CI estimation and to assess the intra- and inter-operator variability. Our results suggest that the operators had a different consistent bias when taking measurements following the standard procedure. Indeed, their intra-operator estimates were quite consistent (mean CoV=28%), but differed from those of their colleagues (inter-operator CoV=35%). This possibly reflects the different preferred measurement sites of the operators (so that the longitudinal section is similar for the repeated measurements, but different among the three operators). The automated approach, when compared to the standard one, provided smaller inter-operator variability, suggesting that it could contribute to standardizing CI measurements (intra-operator and inter-operator mean CoV equal to 24 and 18%, respectively). Furthermore, the average ICC and Fisher ratio were higher in the CI estimated by the automated method, suggesting that the new approach may allow to better discriminate different subjects. Finally, comparing the standard and automated CI estimations, a direct correlation emerged with operators' experience (the lowest and highest correlation for the least and most experienced echographer, respectively). Hence, the automated method could also be a reference for teaching to novices how to make a manual measurement.

A real time rendering of the identified IVC borders could be a useful feedback to guide the acquisition of a B-mode video-clip. Notice also that the most experienced operator (who made measurements highly correlated to

those of the automated method) selected the M-mode line along the direction mostly orthogonal to the IVC midline: our results further support this choice, already suggested in Pasquero et al. (2015).

RCI and CCI: new indexes estimated by the automated method

As the automated method provides not only local estimates, but time series, more information can be extracted by post-processing. Specifically, the heartbeat and respiratory contributions were separated and additional indexes (CCI and RCI) were computed. Figure 6 shows that RCI has a larger variability than CCI. It is reasonable that the variability is lower when considering an index reflecting the cardiac instead of the breath stimulation. Indeed the effect of the heartbeats is about constant, whereas the respiration cycles can be more variable, so that their effect on different measurements can be important. Moreover, the number of heartbeats is much larger than that of respiration cycles found on the same video-clip, so that more estimations can be averaged when computing CCI than RCI.

Notice that the CoV of the RCI is larger than that of the automated estimation of the CI (CI_{global}), even if the latter is affected by the asynchronous super-position of the heartbeats over the respiration cycles, which introduces a variation in the estimations. However, even if the variability of the estimations of CI is a bit larger than that of the RCI, the mean value is much lower for the latter than the first, so that its CoV is larger. A similar interpretation can be given concerning the results of CCI: the estimates are very stable (with a much lower variability than that of CI), but their absolute values are very small. However, CCI is the index providing the largest intraclass correlation (Table 2) and Fisher ratio (Table 3), indicating that it has

high repeatability and can better discriminate different subjects. Further work is needed to understand how the information provided by these two indexes correlate with the state of the patient (this work investigates only the repeatability of their estimations). For example, we expect that irregular cardiac rhythm may cancel or largely affect the cardiac component, so that the relative weight of the two components could be of help in discriminating some patients.

General comments

The consequence of the large variability of the standard measurement is that clinical CI estimations should be considered with caution (Magnino et al. (2017)). Indeed, problems are expected when the index is used to discriminate between patients with different pathologies: for example, only differences among subjects in the order of 20-30% can be assessed with some confidence. Moreover, it is difficult to monitor a patient in the follow up, as only large variations can be assessed. Finally, clinicians using different approaches in selecting the M-mode line could get different diagnoses.

In order to improve the reliability and repeatability of the estimations, a possible solution is averaging more measurements. Different CIs measured on more respiration cycles can be averaged. In this way, an index is obtained accounting for different vein dynamics, induced by different breath stimulations. Moreover, averaging allows to reduce estimation errors due to small mistakes in measuring on still images the maximal and minimal diameters (which are also affected by the oscillations induced by the heartbeats, which are asynchronously superimposed to those induced by respiration). Furthermore, an average of information from different sections could further improve

the estimation of IVC pulsatility, at the expense of spending time repeating more M-mode investigations along different sections.

Our method allows to average information from different respiration cycles and sections, processing a single US video-clip. This provides a fast and robust overall estimation of the pulsatility in an entire portion of the vein. Here, we show that the averaged estimation provided by our semi-automated method is also more repeatable than the manual assessment. Our results could be considered preliminary, due to the low number of investigated subjects (i.e., 10). However, other indications of the reliability of the information extracted by our automated algorithm are available. For example, the pulsatility of IVC extracted by our algorithm has been recently used to estimate the right atrial pressure, with performances largely superior than those that could be obtained from the manual estimations (Mesin et al. (2019a)). Moreover, works are in progress on the applications on patients, where our algorithm allows to get better discrimination of patients affected by either hypo- or hyper-volaemia.

Using an automated method reduces the problems due to subjective interpretations. However, the procedure is still dependent on the quality of the video recorded by the operator, so that the experience of the echographer is still important. In future, the real time rendering of the output of the processing algorithm could provide a feedback to help the operator to acquire a video-clip of good quality. Even considering this limitation of our work (in which the processing was executed off-line), our algorithm allowed to get CI estimations closer to those obtained by the most experienced operator, also when applied to video-clips recorded by a low experience echographer.

Thus, we propose this innovative algorithm as a step towards standardizing measurements of IVC pulsatility.

An instrument applying the algorithm described in this paper was patented by Politecnico di Torino and Universitá di Torino (patent number 102017000006088).

Conclusions

Different sources of variability affect the estimation of IVC pulsatility from US measurements, e.g., the respiration cycles and the selected section of the vein. Our semi-automated algorithm allows to track vein movements and deformations along the long axis, to compute the diameter of different sections orthogonal to the vein and to provide an estimation of pulsatility which is averaged across respiration cycles and sections. The pulsatility estimations of this software were found to be more repeatable than those obtained by the standard approach. This method can provide an important contribution in the standardization of the assessment of IVC pulsatility, with important outcomes expected in the estimation of the central venous pressure and volemic status of patients.

References

Barbier C, Loubieres Y, Schmit C, Hayon J, Ricome J, Jardin F. Vieillard-Baron A. Respiratory changes in inferior vena cava diameter are helpful in predicting fluid responsiveness in ventilated septic patients. Intensive Care Med, 2004;30:17401746.

Bartko J. The intraclass correlation coefficient as a measure of reliability. Psychol Report, 1966;19:3–11.

Blehar D, Dickman E, Gaspari R. Identification of congestive heart failure via respiratory variation of inferior vena cava diameter. Am J Emerg Med, 2009;27:71–75.

Blehar D, Resop D, Chin B, Dayno M, Gaspari R. Inferior vena cava displacement during respirophasic ultrasound imaging. Critical Ultrasound Journal, 2012;4:1–5.

Brennan J, Ronan A, Goonewardena S, Blair J, Hammes M, Shah D, Vasaiwala S, Kirkpatrick J, Spencer K. Handcarried ultrasound measurement of the inferior vena cava for assessment of intravascular volume status in the outpatient hemodialysis clinic. Clin J Am Soc Nephrol, 2006;1:749–753.

Chen L, Hsiao A, Langhan M, Riera A, Santucci K. Use of bedside ultrasound to assess degree of dehydration in children with gastroenteritis. Acad Emerg Med, 2010;17:1042–1047.

Feissel M, Michard F, Faller J, Teboul J. The respiratory variation in inferior vena cava diameter as a guide to fluid therapy. Intensive Care Med, 2004;30:1834–1837.

Fields J, Lee P, Jenq K, Mark D, Panebianco N, Dean A. The interrater reliability of inferior vena cava ultrasound by bedside clinician sonographers in emergency department patients. Acad Emerg Med, 2011;18:98–101.

Folino A, Benzo M, Pasquero P, Laguzzi A, Mesin L, Messere A, Porta M. Roatta S. Vena cava responsiveness to controlled isovolumetric respiratory efforts. Journal of Ultrasound in Medicine, 2017;36:2113–2123.

Grant E, Rendano F, Sevinc E, Gammelgaard J, Holm H, S. G. Normal inferior vena cava: caliber changes observed by dynamic ultrasound. AJR Am J Roentgenol, 1980;135:335–338.

Kimura B, Dalugdugan R, Gilcrease G, Phan J, Showalter B, Wolfson T. The effect of breathing manner on inferior vena caval diameter. Eur J Echocardiogr, 2011;12:120–123.

Kircher B, Himelman R, Schiller N. Noninvasive estimation of right atrial pressure from the inspiratory collapse of the inferior vena cava. Am J Cardiol, 1990;66:493–496.

Krupa A, Fichtinger G, Hager G. Full motion tracking in ultrasound using image speckle information and visual servoin. Proc. ICRA, 2007:2458–2464.

Lichtenstein D. Inferior vena cava. general ultrasound in the critically ill. Berlin: Springer, 2005;23:82.

Lyon M, Blaivas M, Brannam L. Sonographic measurement of the inferior vena cava as a marker of blood loss. Am J Emerg Med, 2005;23:45–50.

Magnino C, Omedé P, Avenatti E, Presutti D, Iannaccone A, Chiarlo M, Moretti C, Gaita F, Veglio F, Milan ARI. Inaccuracy of right atrial pressure estimates through inferior vena cava indices. Am J Cardiol., 2017;120:1667–73.

Mesin L, Albani S, Sinagra G. Non-invasive estimation of right atrial pressure using the pulsatility of inferior vena cava. Ultrasound Med Biol, in press, 2019a.

Mesin L, Pasquero P, Albani S, Porta M, Roatta S. Semi-automated tracking and continuous monitoring of inferior vena cava diameter in simulated and experimental ultrasound imaging. Ultrasound Med Biol, 2015;41:845–857.

Mesin L, Pasquero P, Roatta S. Tracking and monitoring of pulsatility of a portion of inferior vena cava from long axis ultrasound imaging. Ultrasound Med Biol, in press, 2019b.

Moreno F, Hagan A, Holmen J, Pryor T, Strickland R, Castle C. Non-invasive estimation of right atrial pressure using the pulsatility of inferior vena cava. Am J Cardiol, 2019;53:579–585.

Nakamura K, Tomida M, Ando T, Sen K, Inokuchi R, Kobayashi E, Nakajima S, Sakuma I, Yahagi N. Cardiac variation of inferior vena cava: new concept in the evaluation of intravascular blood volume. J Med Ultrasonics, 2013;40:205–209.

Pasquero P, Albani S, Sitia E, Taulaigo A, Borio L, Berchialla P, Castagno F, Porta M. Inferior vena cava diameters and collapsibility index reveal early volume depletion in a blood donor model. Crit Ultrasound J., 2015;7:17.

644 Resnick J, Cydulka R, Platz E, Jones R. Ultrasound does not detect early
645 blood loss in healthy volunteers donating blood. J Emer Med., 2011;41:270–
646 275.

647 Wallace D, Allison M, Stone M. Inferior vena cava percentage collapse during
648 respiration is affected by the sampling location: an ultrasound study in
649 healthy volunteers. Acad Emerg Med, 2010;17:96–99.

650 Weekes A, Lewis M, Kahler Z, Stader D, Quirke D, Norton H, Almond C,
651 Middleton D, Tayal V. The effect of weight-based volume loading on the
652 inferior vena cava in fasting subjects: a prospective randomized double-
653 blinded trial. Acad Emerg Med., 2012;19:901–907.

654 Yang L, Georgescu B, Zheng Y, Meer P, Comaniciu P. 3d ultrasound tracking
655 of the left ventricles using one-step forward prediction and data fusion of
656 collaborative trackers. Proc. IEEE Conf Comput Vis Pattern Recognit,
657 2008.

658 Yeung F, Levinson S, Fu D, Parker K. Feature-adaptive motion tracking of
659 ultrasound image sequences using a deformable mesh. IEEE Trans. Med.
660 Imaging, 1998;17:945–956.

661 Zhang Z, Xu X, Ye S, Xu L. Ultrasonographic measurement of the respiratory
662 variation in the inferior vena cava diameter is predictive of fluid respon-
663 siveness in critically ill patients: Systematic review and meta-analysis.
664 Ultrasound Med Biol, 2014;40:845–853.

Figure Captions

Figure 1: A) Selection of a rectangle including the IVC portion of interest in the first frame of the video-clip. B) Reference points (squares), leftmost and rightmost sections of interest (continuous lines) and points close to the vessel edges along the leftmost section (indicated by X). C) The algorithm computes 21 lines uniformly distributed between the extreme sections indicated in B) and estimates the profile of the vein along them (the estimated border points are indicated with circles). D) From the estimated border of the vessel, the midline is computed and interpolated with a parabola (dash-dot line); five equidistant points are selected on this parabola, starting from the confluence of the hepatic vein in the IVC and new lines perpendicular to it are considered as sections along which to compute the vein diameters (border points indicated with diamonds).

Figure 2: Experimental protocol. Each operator acquired three manual measurements (in M-mode) and then the video (in B-mode). The same procedure was followed twice for each of the three operators.

Figure 3: A) Caval index (CI) estimated on the whole signal. The local maxima and minima of the respiratory component are found; then a window of 1 s duration centred on each of these points is explored to find the maxima or minima on the whole signal (indicated with circles). B) Respiratory caval index (RCI), computed on the breath component. This component is isolated with a low pass filter; then, maxima and minima (indicated with circles) are automatically found

and used for RCI calculation. C) Cardiac caval index (CCI) computed on the heartbeat component. The component is isolated with a high pass filter; then, its local maxima and minima (indicated with circles) are computed and used for CCI estimation.

Figure 4: A) Time course of IVC diameter at three different sections simultaneously monitored in a representative subject. B) Distribution of CoV of CI_{auto}, obtained considering the 6 measurements from all 10 subjects, separately for the five sections and compared with manual CI and CI_{global}.

Figure 5: Variation of the Caval Index (CI) when estimated by the automated method at different longitudinal positions, expressed as the distance from the confluence of the hepatic veins. A) Each trace corresponds to one subject (average of all sessions). B) Median, quartiles and range (outliers shown individually) of the coefficient of variation (CoV) of the CI across the 5 sections along the vein, for each subject.

Figure 6: Coefficient of variation (CoV) for each index (manual CI and automated estimation of CI, CCI and RCI) computed across different experimental sessions (median, quartiles and range; outliers shown individually). A), B) and C): CoV of the indexes (CI, CCI and RCI, respectively) extracted at different distances from the confluence of the hepatic vein into the IVC and, to the right, the CoV of manual and global estimations (averaging the CI across sections). D) Comparison of CoV of the manual and global CI.

Figure 7: Comparison between CoV of manual and automated Caval In-

dex (CI) values. Intra- and inter-operator variabilities are considered (showing the distribution of 10 values, one for each subject, in terms of median, quartiles and range, plus an outlier shown individually). The manual CI estimations are the mean of three CI measurements in M-mode (reflecting the choice of 3 respiration cycles). The automated CI estimations are given by the mean of all CI measurements obtained from each video-clip (CI_{global}, obtained averaging across 3 respiration cycles and 5 longitudinal sections).

Table 1: ANOVA table considering the CI obtained using either the standard approach (manual CI) or the automated one (CI_{global}); DOF - degrees of freedom, RC - respiration cycle.

Source	**DOF**	**Sum of squares**		**Mean squares**		**F**		**p-value**	
		manual	*global*	*manual*	*global*	*manual*	*global*	*manual*	*global*
Subject	9	4.03	2.30	0.45	0.25	29.01	30.01	$\approx 10^{-29}$	$\approx 10^{-29}$
Repetition	1	$6{\cdot}10^{-4}$	0.026	$6{\cdot}10^{-4}$	0.026	0.03	3.03	0.84	0.083
Operator	2	1.05	0.111	0.53	0.055	34.22	6.49	$\approx 10^{-13}$	0.002
RC	2	0.02	$3.5{\cdot}10^{-4}$	0.01	$1.7{\cdot}10^{-4}$	0.67	0.02	0.51	0.98
Error	165	2.54	1.40	0.015	0.008				
Total	179	7.66	3.84						

Table 2: Intraclass correlation coefficient (ICC), considering intra- and inter-operators estimates of different caval indexes (manual and automated CI, CCI and RCI, obtained averaging across different sections). Different operators are shown in order of increasing experience (FC less than 1 year, AR 2 years, PP more than 20 years of experience).

	ICC			
Operator	CI *standard*	CI_{global}	CCI_{global}	RCI_{global}
FC	48.9%	45.3%	61.2%	6.9%
AR	81.7%	46.8%	72.8%	41.0%
PP	77.6%	78.6%	89.5%	70.7%
Inter-operator	61.5%	70.4%	87.5%	49.9%

Table 3: Fisher ratio of estimates of different caval indexes (manual and automated CI, CCI and RCI, obtained averaging across different sections), considering intra- and inter-operator values.

	Fisher ratio			
Operator	CI *standard*	CI_{global}	CCI_{global}	RCI_{global}
FC	3.20	2.24	2.54	1.43
AR	31.52	2.11	48.83	3.02
PP	9.11	7.34	25.92	9.73
Inter-operator	2.06	8.21	23.52	2.56