Politecnico di Torino

Politecnico di Torino

Master's Degree in Biomedical Engineering

Real-Time Ultrasound Segmentation of the Inferior Vena Cava: A Study on Software Reliability and Measurement Repeatability

Supervisor:
Luca Mesin

Co-Supervisor:
Piero Policastro

Candidates:

Carlotta Avietti

Caterina Greco

A.Y. 2024/2025

Abstract

The inferior vena cava ultrasound (IVC) is a non-invasive method used to estimate the patient's volemic status and support in fluid therapy. A frequently adopted technique to assess fluid responsiveness is the passive leg raise (PLR) maneuver, which improves venous return through lower limbs. However, the reliability of ultrasound measurements depends on the operator's experience, acquisition conditions, vein motion and non-uniform pulsatility. In this context, the aim of this study was to evaluate the reliability and repeatability of IVC ultrasound measurements performed using VIPER real-time segmentation software, which provides the diameter, caval index (CI), cardiac caval index (CCI) and respiratory caval index (RCI). The analysis focused on evaluating intra and inter-operator variability between measurements obtained with and without real time feedback (RealTime vs NOF), and with the manual method compared to the semi-automatic approach without feedback (Manual VS NOF) for supine and PLR, so as to also compare the ability of the three methods to discriminate between the two conditions. Manual measurements were performed via a graphical user interface (GUI), developed in Python, which allowed comparison with the method traditionally used by clinicians. An experimental protocol was defined which consisted in the acquisition of IVC longitudinal section ultrasound from two nonexpert operators, with and without visual software support, in two different conditions (Supine and PLR) repeating the measurement twice (R1 and R2). The sample consisted of 26 healthy subjects (15 females and 11 males), aged 25 $\pm$ 1.37 years. Intra and inter-operator variability was assessed by calculating the Coefficient of Variation (CoV) and Intraclass Correlation Coefficient (ICC) on all measured parameters: mean diameter, CI, CCI, and RCI. The results showed for the RealTime case that all parameters present a lower dispersion than those obtained by post-processing of acquisitions without feedback. CI, RCI and CCI had slightly lower median values of CoV, while diameter's CoVs did not vary significantly. (CI's median CoV, Supine (RealTime): OA=0.15, OB=0.21, inter=0.16; (NOF): OA= 0.15, OB= 0.25, inter= 0.18;). Minor variability was obtained compared to the manual method by observing lower median values of CoV (Manual, Supine CI's median CoV OA=0.24, OB=0.34, inter=28). A linear mixed effects model was developed and showed

that the condition had a statistically significant effect on all parameters while the presence of real-time feedback on all but the diameter. To assess to what extent the two conditions could be distinguished using each of the three measurement methods, Cohen's d was calculated as a measure of effect size. The no-feedback (NOF) approach showed higher discriminative power compared to the manual method. It also outperformed the real-time modality. These unexpected results may be explained by a learning bias, as the operators gained experience while simultaneously becoming familiar with the software. These findings highlight the need to repeat the study with clinically experienced operators in order to isolate the true effect of real-time feedback from operator learning effects. Overall, it is evident that the software enables a faster and more standardized acquisition of measurements, with reduced intra- and inter-operator variability, in contrast to the manual method, which remains more time-consuming and operator dependent.

Table of Contents

List of Figures

Chapter 1

Principles of Anatomy and Physiology

1.1 Anatomy and physiology of the cardiovascular system

The cardiovascular system is the system responsible for delivering blood to various parts of the body. Its primary functions include the delivery of oxygen and nutrients to tissues and the removal of metabolic waste products. This system consists of three components: the heart, a muscular pump that forces blood around the body, blood vessels (arteries, veins, and capillaries), and blood.
The circulation is organized into two main circuits:

- Systemic circulation, which carries oxygenated blood from the left heart to peripheral tissues and returns deoxygenated blood to the right heart.
- Pulmonary circulation, which transports deoxygenated blood from the right heart to the lungs for oxygenation, then back to the left heart.

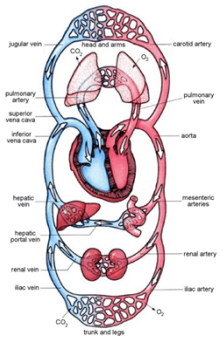


Figure 1.1: Schematic representation of double circulation in the human cardiovascular system.

This double circulation ensures the separation of oxygenated and deoxygenated blood and allows for efficient gas exchange and tissue perfusion. The proper function of the cardiovascular system depends on the balance between cardiac output, vascular resistance, and venous return. These elements are interdependent parameters that determine overall circulatory efficiency and ensure adequate perfusion pressure and tissue oxygenation, which are critical for homeostasis. Because it reflects the blood volume status and the function of the right heart, venous return is a crucial parameter in clinical settings.

1.1.1 Heart

The heart is the central organ of the cardiovascular system and its main function is to pump blood continuously throughout the body. [1] The heart is located in the mediastinum, the central compartment of the thoracic cavity, and encased in a double-layered serous sac called the pericardium. It is divided into four chambers: two upper chambers called atria and two lower chambers called ventricles. The right atrium receives deoxygenated blood from the systemic circulation via the superior and inferior venae cavae, while the left atrium receives oxygenated blood from the pulmonary veins. This blood then flows into the right and left ventricles, respectively, which contract to propel the blood into the pulmonary and systemic circulations. The walls of the ventricles are composed of cardiac muscle, or myocardium, which is thicker in the left ventricle due to the higher pressures required to circulate blood throughout the body. The right ventricle, in contrast, pumps blood only to the lungs and thus has a thinner wall. Four heart valves regulate blood flow and prevent backflow: the tricuspid, pulmonary, mitral, and aortic valves. These open and close in response to pressure gradients between chambers. Atria and ventricles are separated by atrioventricular valve, which open and close in response to pressure differences between the chambers, permitting unidirectional blood flow from the atria to the ventricles.

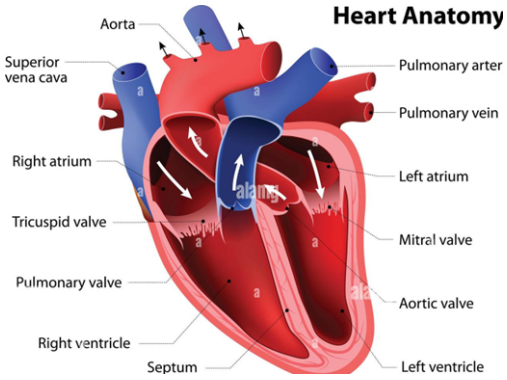


Figure 1.2: Representation of the heart.

The cardiac cycle consists of two phases: systole (ventricular contraction phase) and diastole (ventricular relaxation phase). During systole, the ventricles contract, ejecting blood into the aorta and pulmonary artery. During diastole, the ventricles relax and fill with blood from the atria. This cycle is coordinated by the heart's intrinsic conduction system and regulated by the autonomic nervous system, making cardiac contraction involuntary. Electrical activity starts from the sinoatrial (SA) node, spreads through the atria, and reaches the ventricles via the atrioventricular (AV) node, bundle of His, and Purkinje fibers, enabling near-simultaneous ventricular depolarization in about 30 milliseconds. Cardiac output, defined as the volume of blood the heart pumps per minute, is the product of stroke volume and heart rate. It is a critical parameter in assessing cardiovascular function and tissue perfusion. [2]

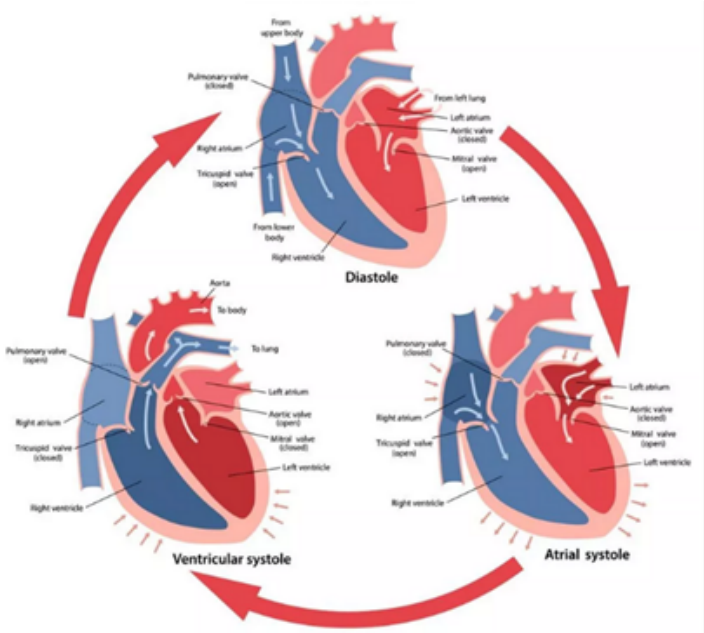


Figure 1.3: Cardiac Cycle.

1.1.2 Electrocardiogram (ECG)

The electrical activity of the heart is traced using an electrocardiogram (ECG) [3]. Each heartbeat is initiated by an electrical impulse generated by the SA node. This impulse causes the atria to contract (P wave), is delayed at the AV node, and then travels through the ventricles via the bundle of His and Purkinje fibers, leading to ventricular contraction (QRS complex). Ventricular repolarization is represented by the T wave. Each wave or complex in the ECG has a specific meaning and corresponds to a distinct phase of the cardiac cycle:

- P wave corresponds to atrial depolarization. It is the first visible sign on the ECG trace, and this is the moment when the atria contracts to push blood into the ventricles. The P wave indicates that the electrical impulse is generated by the sinoatrial (SA) node and that it propagates correctly through the atrial walls.
- QRS complex indicates ventricular depolarization, the phase during which the ventricles contract to eject blood into the pulmonary and systemic circulations. The QRS complex is essential for evaluating ventricular function.
- T wave represents ventricular repolarization, the phase in which the ventricles relax after contraction and prepare for the next heartbeat. The shape and orientation of the T wave provide insight into the health of the ventricular myocardium and the efficiency of repolarization.

In addition to the previously described waves, other elements of the ECG can also carry clinical significance. Analyzing the shape, duration, and timing of the waves can reveal arrhythmia, conduction disorders, myocardial infarction, electrolyte disturbances, and other cardiac pathologies. For instance, a prolonged QT interval may suggest a risk of ventricular tachyarrhythmias, while ST-segment elevation is often associated with acute myocardial infarction.

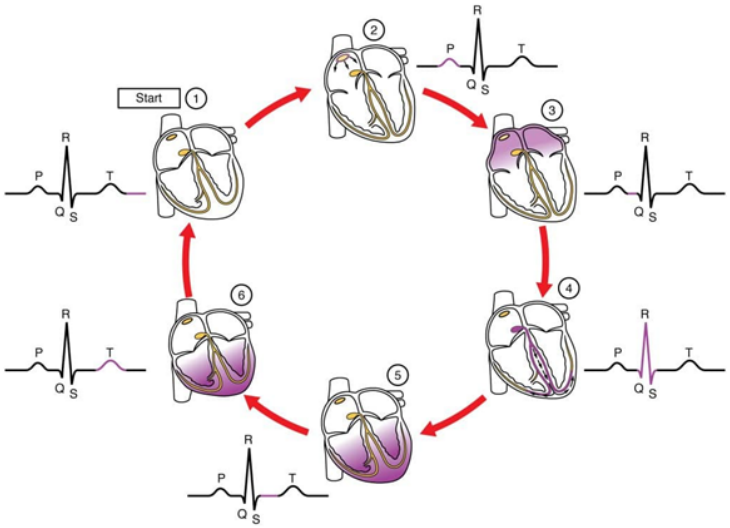


Figure 1.4: Representation of the cardiac cycle and corresponding ECG waveform.

1.1.3 Blood Vessels

Blood vessels play a fundamental role in sustaining cardiovascular function, acting as the conduits that distribute blood to and from all tissues and organs. [2] There are three main types of vessels:

- Arteries carry blood away from the heart, delivering oxygenated blood to the tissues (except for the pulmonary arteries). The largest artery, the aorta, branches into arterioles, which are divided into capillaries at the tissue level. Arteries function as pressure reservoirs, storing elastic energy during systoles and releasing it during diastole to maintain continuous blood flow.

- Capillaries are the smallest blood vessels and serve as the primary site of gas, nutrient, and waste exchange between blood and tissues.

- Veins return deoxygenated blood to the heart (except for the pulmonary veins, which carry oxygenated blood from the lungs). Veins act as volume reservoirs, accommodating fluctuations in blood volume and containing up to 60 of the total circulating blood.

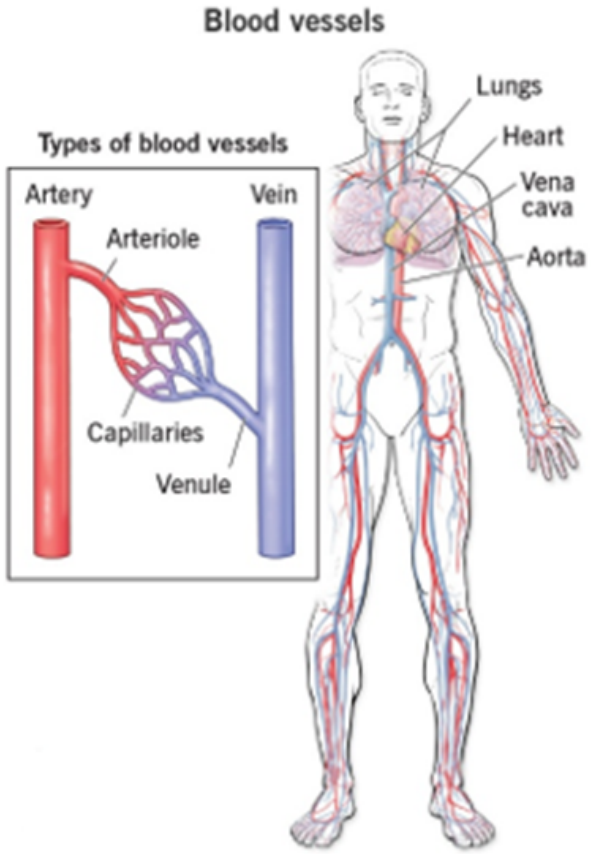


Figure 1.5: Anatomical overview of the blood vessels in the human body..

1.1.4 Inferior Vena Cava (IVC)

The vena cava is divided into Inferior Vena Cava (IVC) and Superior Vena Cava (SVC). The IVC is the largest vein in the body, responsible for returning deoxygenated blood from the limbs and abdominopelvic region to the heart. It is located at the posterior abdominal wall on the right side of the aorta. It is formed by the confluence of two large veins: the left and right common iliac veins, each of which results from the union of the internal and external iliac veins. [4] The iliac veins are responsible for collecting blood from the lower limbs, pelvis, external genitalia, and lower abdominal wall. After its formation at the level of the L4–L5 vertebrae, the IVC ascends through the abdomen, receiving venous blood from several tributaries, including the lumbar veins, renal veins, adrenal veins, hepatic veins, and gonadal veins. The hepatic veins are of particular importance: they transport filtered, deoxygenated blood from the liver directly into the terminal portion of the IVC, just before it enters the right atrium of the heart. Unlike the portal vein, which carries blood into the liver for processing, the hepatic veins carry it away from the liver. Then It passes through the diaphragm at the level of the T8 vertebra via the caval opening and empties directly into the posterior aspect of the right atrium.

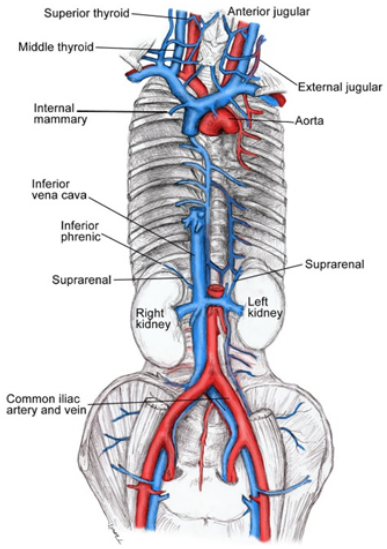


Figure 1.6: Anatomical course of the IVC and related vascular structures.

It plays a vital role in cardiovascular dynamics and is commonly assessed via ultrasound to estimate central venous pressure, particularly in critically ill patients. Infact, the size and dynamics of the IVC depend on the blood volume and right atrial pressure, which are important indicators of a patient's hydration and reflect possible pathological conditions. [5] The IVC is approximately 22 cm in length (18

of which run through the abdomen) and about 2.5cm in diameter (with a range of diameter from 1.8 cm to 3.2 cm). A larger diameter may suggest conditions such as congestive heart failure or hepatic cirrhosis, whereas a smaller diameter may be a sign of hypovolaemia or constrictive pericarditis. The diameter of the IVC increases when the total blood volume increases and decreases when the total blood volume decreases. [5] The diameter also varies during the respiratory cycle due to the changes in intrathoracic pressure during inspiration and expiration, this variation can be expressed as the caval index:

$$CI = \frac{\max(D) - \min(D)}{\max(D)} \tag{1.1}$$

Where max(D) and min(D) represent, respectively, the maximum diameter of the IVC during exhalation and the minimum diameter of the IVC during inspiration. Spontaneous respiratory activity and cardiac pulsations introduce two distinct oscillatory components into the dynamics of the IVC, each characterized by different frequencies that can be isolated through frequency- domain analysis. Indeed, by filtering the IVC diameter signal in different frequency ranges, it is possible to isolate these two components and estimate two derived indices: the Respiratory Collapsibility Index (RCI) and the Cardiac Collapsibility Index (CCI). They provide information that is partly independent and unrelated. To calculate theme, it is necessary to estimate the IVC diameter in each frame of the video and they represent an extension of the classic Caval Index (CI), offering a more detailed view of the respiratory and cardiac components in venous dynamics.

1.2 Anatomy and physiology of the respiratory system

The term 'respiratory system' is derived from the process of respiration, which is the breathing activity and gas exchange. The primary function is providing oxygen into the body, which is necessary for cellular metabolism, and removing carbon dioxide which is a by-product of metabolic activity. It consists of diverse structural components which are grouped under two main divisions: upper respiratory tract includes the nose , nasal cavity, mouth, pharynx, nasopharynx, and larynx, while the lower respiratory tract includes the diaphragm, intercostal muscles, trachea, bronchi, bronchioles, alveoli, and lungs (Figure 1.7). In addition to its main function, respiration, the respiratory system performs several other functions, including contributing to the regulation of acid-base balance in the blood, enabling vocalization, participating in defense against pathogens and foreign particles in the airways, providing a route for water and heat losses (via the expiration of air that was moistened and warmed during inspiration), enhancing venous return (through the respiratory pump), and activating certain plasma proteins as they pass through the pulmonary circulation.[6]

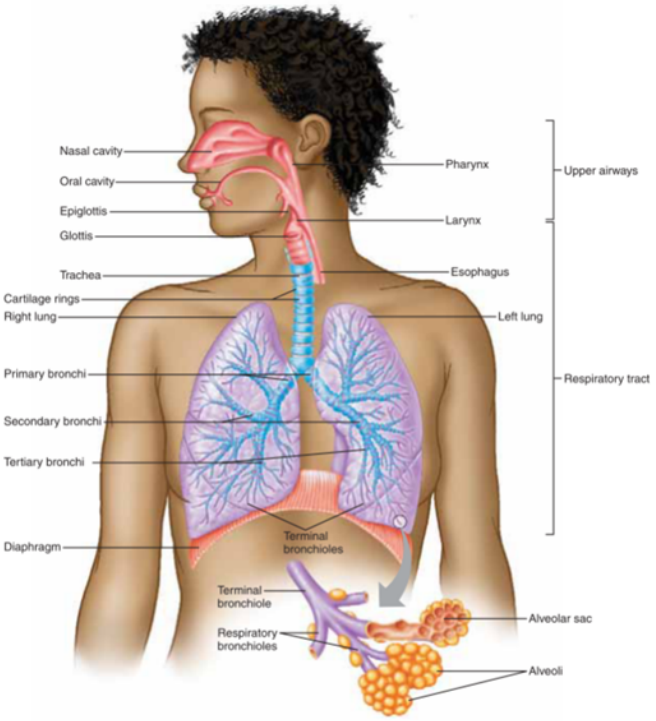


Figure 1.7: Anatomy of the upper airways and the respiratory tract.

1.2.1 Mechanics of Breathing

Breathing is the process by which air moves into and out of the lungs, allowing for gas exchange. This movement is driven by pressure gradients created by the respiratory muscles, which change the volume of the thoracic cavity and, consequently, the lungs. The process consists of two main phases: inspiration (air entering the lungs) and expiration (air leaving the lungs). Ventilation is governed by four main pressures:

- Atmospheric Pressure (P_{atm}): the pressure of the air outside the body, typically 760 mm Hg at sea level. All other lung pressures are measured relative to this;
- Intra-alveolar Pressure (P_{alv}): the pressure of air within the alveoli. The difference between P_{alv} and P_{atm} creates the pressure gradient that drives air movement. When P_{alv} is less than P_{atm}, air flows into the lungs (inspiration). When P_{alv} is greater than (P_{atm}), air flows out (expiration);
- Intrapleural Pressure (P_{ip}): the pressure within the intrapleural space, the thin fluid-filled compartment between the visceral and parietal pleurae of the lungs;
- Transpulmonary Pressure: the difference between intra-alveolar and intrapleural pressure (P_{alv} - P_{ip}). This pressure represents the distending force across the lungs, so an increase in transpulmonary pressure results in lung expansion.

The graph in (Figure 1.8) displays how these pressures change over the course of a respiratory cycle.
The primary muscles responsible for breathing include the diaphragm and the external intercostal muscles, both of which play key roles during inspiration. The diaphragm is a dome-shaped muscle that separates the thoracic cavity from the abdominal cavity. When it contracts, it flattens and moves downward, increasing the vertical dimension of the thoracic cavity and allowing the lungs to expand. When this muscle is more actively engaged during breathing, it is referred to as 'diaphragmatic breathing'. This is considered the natural and most efficient breathing pattern, often characterized by slower, deeper breaths, as it allows air to reach the lower, larger regions of the lungs. Meanwhile, the external intercostal muscles, located between the ribs, contract to elevate and rotate the ribs upward and outward, which further expands the chest wall and increases the overall volume of the thoracic cavity. This pattern of breathing is referred to as 'thoracic breathing', where air is primarily drawn into the upper regions of the lungs. Thoracic breathing tends to be more shallow and is commonly observed during periods of stress, anxiety, or poor posture.
In addition to these main inspiratory muscles, several accessory muscles assist during forceful or deep inspiration. These include the scalenes, sternocleidomastoids, and

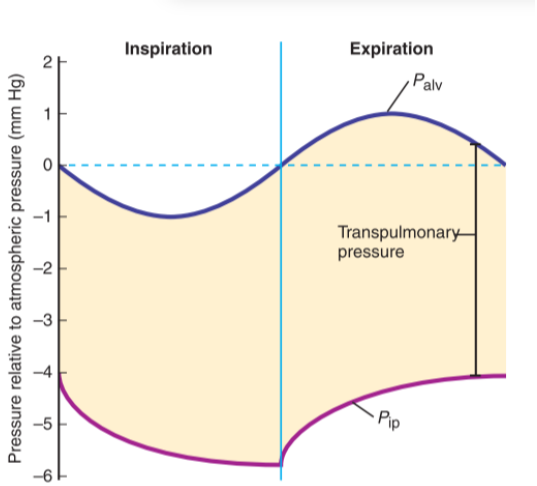


Figure 1.8: Changes in P_{alv}, P_{ip}, and transpulmonary pressure during a respiration cycle.

pectoralis minor, which help to further elevate the ribs and sternum to maximize lung expansion when greater airflow is needed.

On the other hand, expiration during quiet breathing is generally passive, relying on the elastic recoil of the lungs and chest wall. However, during active or forced expiration, other muscles come into play. The internal intercostal muscles contract to pull the ribs downward and inward, reducing the chest cavity volume, while the abdominal muscles contract to push the diaphragm upward, both actions facilitating the expulsion of air from the lungs. The Figure 1.9 shows a schematic of the muscles just mentioned. Changes in the pressure inside the thoracic cavity are the primary force behind the breathing mechanism. The thoracic cavity enlarges as the inspiratory muscles contract. Boyle's law states that the intra-alveolar pressure falls as the thoracic cavity's volume increases. Air flows inward as a result of this pressure drop because it produces a gradient between the outside atmospheric pressure and the lower pressure inside the lungs.

The procedure is usually the opposite during expiration. Because the lungs are elastic, they naturally recoil as the inspiratory muscles relax. The thoracic cavity's volume is reduced by this recoil, raising the intra-alveolar pressure above atmospheric pressure. Air is consequently forced out of the lungs.

More specifically, the diaphragm and external intercostal muscles contract during inspiration, which is an active process. Their contraction causes the thoracic cavity to enlarge, which lowers the intra-alveolar pressure and permits air to enter the lungs. When breathing quietly and peacefully, expiration is typically passive; it happens when the inspiratory muscles relax and the chest wall and lungs return to their resting positions due to elastic recoil. As a result, air is forced out,

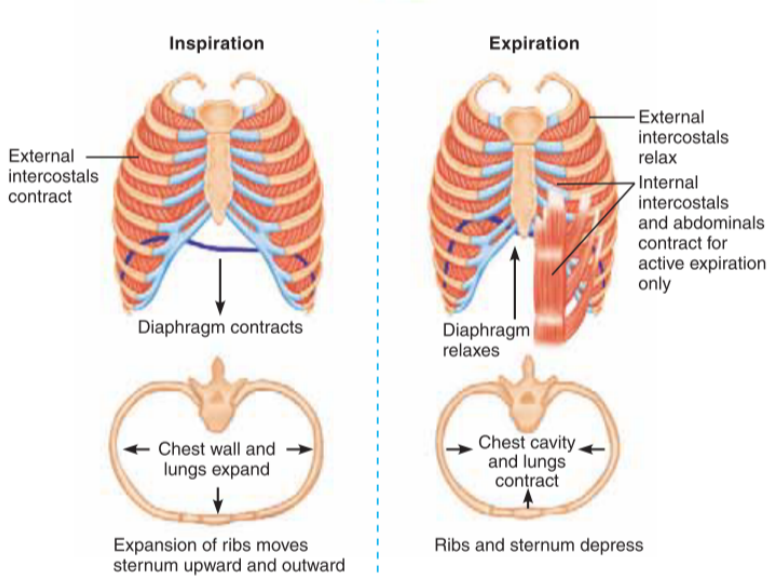


Figure 1.9: Locations of inspiratory and expiratory muscles.

intra-alveolar pressure rises, and thoracic volume decreases.
Expiration, however, turns into an active process during times of elevated demand, like exercise or forced breathing. The contraction of expiratory muscles in this situation causes the thoracic cavity's volume to decrease even more, which raises intra-alveolar pressure and speeds up the expulsion of air from the lungs.[6]

1.2.2 Effects of respiration on IVC

During respiration, the size of the IVC, particularly its abdominal section, varies in response to pressure variations in the thorax and abdomen.
In conditions of respiratory rest at the end of expiration, that is, when neither the thorax nor the abdomen are actively involved in respiration, the IVC is in its maximum diameter condition. This is the moment in which the internal and external pressures are in equilibrium, and therefore the transmural pressure is relatively high, allowing the vein to remain well distended.
During thoracic inspiration, however, a reduction in intrathoracic pressure occurs and thus promotes venous return. However, this effect can temporarily lower the pressure within the IVC, thereby reducing transmural pressure and causing a slight reduction in the diameter of the vein.
An even more pronounced effect is observed during diaphragmatic inspiration, which causes an increase in abdominal pressure that then compresses the IVC from the outside, further reducing transmural pressure. As a result, the vein narrows more noticeably than during chest inspiration.

The opposite situation is observed during positive pressure ventilation, such as that used in intubated patients or those on ventilator support. In this case, the pressure within the chest is artificially increased, preventing the return of blood to the heart. This leads to an increase in pressure within the IVC, which results in an increase in transmural pressure and therefore dilation of the vein. An explanatory scheme is shown in Figure 1.10.[7]

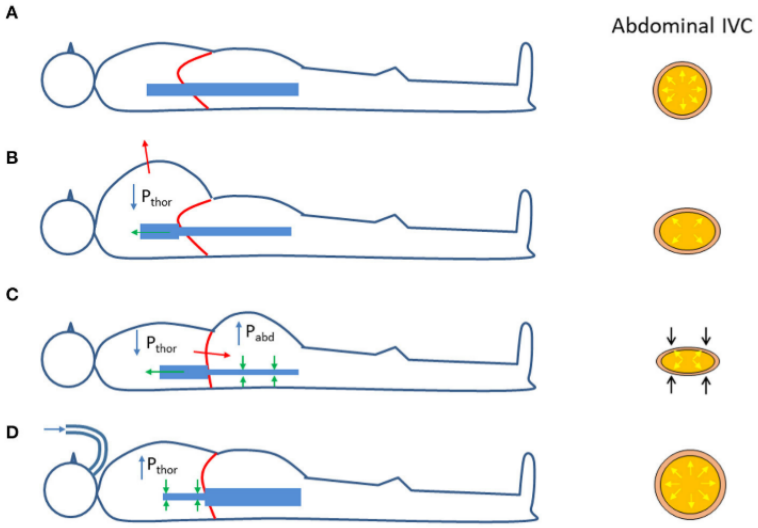


Figure 1.10: Effects of respiration on the size of the abdominal IVC; Pthor = Intrathoracic pressure; Pabd = abdominal pressure. (A) Functional residual capacity (end expiration). (B) Thoracic inspiration. (C) Abdominal (diaphragmatic) inspiration. (D) Positive pressure ventilation (inspiration).

1.3 Fluid therapy

Fluid therapy is a fundamental therapeutic intervention that involves the administration of fluids to maintain or restore the body's water and electrolyte balance. Intravenous fluid administration is a common practice in the emergency department (ED), intensive care unit (ICU), or operating room. Crystalloids (such as 0.9% saline and balanced solutions such as Hartmann's, Ringer's, and Plasma-Lyte) are the most commonly used fluids.
Colloids have been proposed to reduce tissue edema, but studies such as the SAFE trial [8] and the Cochrane review [9] have shown that colloids are associated with an increased risk of mortality and may increase the risk of kidney injury. Currently, data suggest that crystalloids are preferable to colloids for initial resuscitation in the ED.

The amount of fluid to administer is a complex and understudied issue in patients who are not in shock or critically ill. In patients in shock (hypovolemic or septic), fluid therapy aims to increase cardiac output and correct tissue hypoperfusion according to the Frank-Starling law. The use of simple physiological parameters (blood pressure, heart rate, urine output) to guide therapy is common, but may be insufficient, since shock may persist even with normal values of these parameters. A strategy often used in the clinic is the "goal-directed fluid therapy" (GDFT) for the administration of fluids (and sometimes vasoactive or inotropic drugs), which is based on the continuous monitoring of hemodynamic parameters such as cardiac output (CO) or caval index (CI), stroke volume (SV) or its variation (SVV), mean arterial pressure (MAP), central venous oxygen saturation ($ScvO_2$) and the response to "fluid challenge" maneuvers [10].
The principle behind the fluid challenge technique is that by administering a small amount of fluid over a short period of time, the clinician can assess whether the patient has a preload reserve that can be used to increase stroke volume with additional fluid. In patients who respond to volume expansion (defined as individuals whose SV or CO increases $\geq$ 10–15% from baseline values), administration of fluids is repeated until the patient no longer has a preload reserve (i.e., until increases in SV or CO are <10–15 % from baseline) or until other signs of shock (e.g., hypotension) resolve [10].
The Figure 1.10 shows an example of a Frank-Starling curve and its relationship with the variation of the IVC in patients undergoing invasive mechanical ventilation, in fact a dilation of the vein can be seen during the inspiration phases. This figure shows the expected increase in stroke volume after fluid administration, which depends on cardiac function and initial preload. For the same amount of fluid administered and for a similar initial preload, the resulting variation in stroke volume depends on cardiac function: (A) The Frank-Starling curve of a patient with normal cardiac function. (B) The Frank-Starling curve in a patient with reduced cardiac function. In case A the results of fluid administration depend only on the initial preload: if it is low (ascending phase of the curve), the stroke volume increases significantly therefore the patient is responsive and there is a significant variation in the diameter of the inferior vena cava with the application of positive pressure to the chest during inspiration in the ventilated patient. If it is high (flat phase of the curve), no significant increase in stroke volume is observed, therefore the patient is a non-responder, resulting in pulmonary overload, which corresponds to a poorly distended inferior vena cava. In case B, fluid administration, even with a low initial preload, can cause pulmonary fluid overload without a significant increase in stroke volume.
It is essential to avoid both underdosing and overdosing of fluids, and to individualize therapy based on clinical response and associated risks.

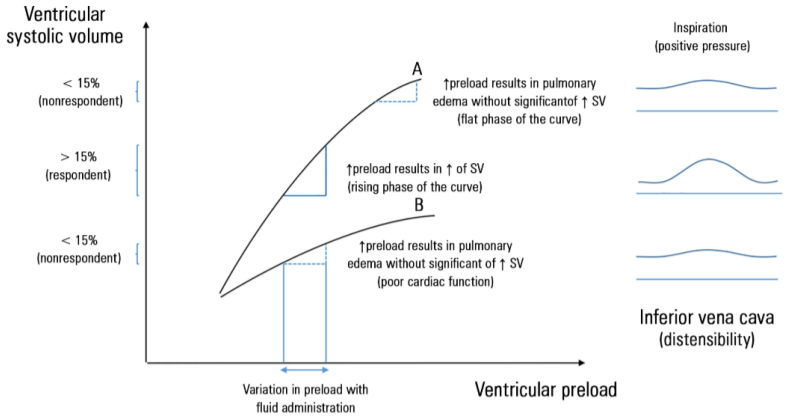


Figure 1.11: Frank-Starling curve and its relationship with IVC variation among patients under invasive mechanical ventilation. SV - systolic volume

1.3.1 PLR

PLR-induced changes in cardiac output (CO) very reliably predict the CO response to volume expansion in adults with acute circulatory failure. Patients who increase CO following a fluid load are defined as fluid responders. It has been shown in the literature that about half of haemodynamically unstable patients do not show a positive response to fluid therapy [11].
Therefore, it is important to predict whether or not fluid therapy will induce a response in the patient, because otherwise it can lead to potentially catastrophic effects, as there is no possibility of returning to a lower fluid state once overload occurs. It might lead to increased leftovers ventricular filling pressure with pulmonary and tissue edema, which is associated with increased mortality and invasive mechanical ventilation (IMV) time [12].
Among the tests and indices developed to predict fluid responsiveness, the passive leg raise (PLR) test has received considerable attention. PLR is performed by raising the patient's legs to 45° from two possible baseline positions 1.12 :

- the 45° semi-recumbent position [13], which induces further recruitment of the vast splanchnic reservoir. This position is more suitable for inducing a larger increase in cardiac preload and may therefore be preferred for predicting fluid responsiveness;

- the supine position, which results in a lower hemodynamic impact.

This maneuver causes a hydrostatic increase in mean systemic pressure, thus mimicking the effects of volume expansion [14]. Approximately 300 mL of venous

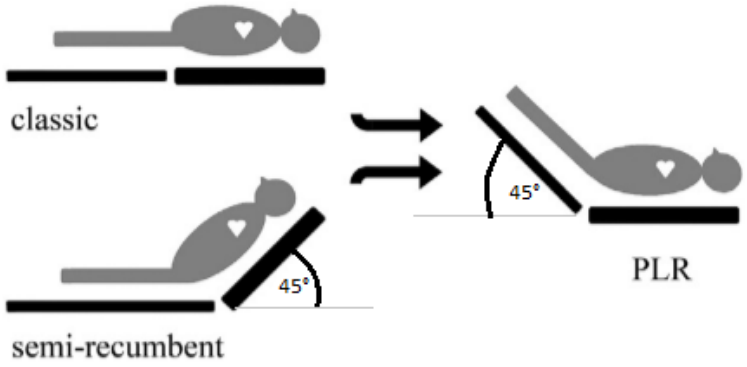


Figure 1.12: The two starting positions for passive leg raising (PLR): supine and semi-recumbent

blood is transferred from the lower body to the right atrium through the IVC, acting as a reversible 'fluid challenge'.
If both ventricles are preload dependent, that is, if their stroke volume increases with increasing venous return, then the resulting increase in central venous return will lead to an increase in CO.
One of the advantages of the PLR test is that it can be used in clinical situations where dynamic fluid responsiveness indices (as mentioned in the discussion on goal-directed therapy) are not reliable, such as in cases of cardiac arrhythmias, low lung compliance, or low tidal volume ventilation.
Another major advantage is that, since there is no actual fluid administration, this test can be considered a reversible or 'virtual' fluid load, thus avoiding the risk of fluid overload [15].

1.3.2 IVC assessment in fluid therapy and limitations

Among the most used approaches to assess the volemic state, there is the analysis of the IVC, whose diameter and its respiratory variation, expressed by the index of collapsibility (CI) , can provide indirect indications on the CVP (central venous pressure) and thus on the filling state of the patient. However, there are several limitations to the assessment of IVC. Ultrasound (US, ultrasound) is an increasingly used diagnostic tool in the clinical field due to its non-invasive nature, speed of execution and availability at the patient's bed (point-of-care ultrasound, POCUS)[16]. The ability to be informed quickly of the patient's volemic status is valuable, particularly in septic shock and suspected hypovolemia, where fluid resuscitation early in the process enhances tissue perfusion and prevents organ

dysfunction [5]. First, the measurement of IVC diameter may be influenced by a marked intra- and inter-operator variability, related to both the operator's experience and the technique used. In many cases, the assessment is also carried out visually [17], in the absence of an automated quantitative analysis, making the process subjective and potentially unrepeatable. In patients with mechanical respiration, the assessment is generally more reliable, since the intra-thoracic pressure variation is more predictable and standardized. In contrast, in patients with spontaneous respiration, a common situation in an emergency setting, the IVC may show less consistent variations, making measurements less reliable and more difficult to interpret [18] [19]. This dependence upon the operator and the poor standardization of procedures used highlight the need for equipment capable of automatically and more objectively processing ultrasound. The use of algorithms that are able to identify and separate IVC in ultrasound images automatically can significantly reduce measurement variability, improve repeatability, and provide more uniform and reliable information for patient diagnosis and monitoring [20]. In this context, the development of a fast, intuitive and integrable software in clinical protocols can represent a decisive step towards greater effectiveness of fluid therapy guided by ultrasound images.

Chapter 2

Ultrasonography

Ultrasonography is a diagnostic imaging technique that uses high-frequency sound waves to create images of the organs and internal structures of the body. The image is the result of the interaction of ultrasound with body tissues, which depends on the acoustic impedance of each tissue. In medical applications, the frequency range of sound waves is in the range of 2 to 10 MHz. [21] One of the advantages is that in ultrasound imaging non-ionizing radiation is used, which is not harmful to the human body and can be absorbed repeatedly: this allows repeated measurements over time for monitoring and applied to at-risk patients, such as pregnant women. Ultrasound has a high temporal resolution which enables it to follow rapidly changing phenomena.

2.1 Physical Principles of Ultrasonography

Sound is the result of the propagation of mechanical energy through matter in the form of a pressor wave, which alternately produces bands of compression and rarefaction of molecules within the conduction medium. The propagation speed of the ultrasound in the medium is regulated by this equation:

$$v = \lambda \cdot f \tag{2.1}$$

where v is the velocity of sound waves, λ is the wavelength, f is the frequency. The propagation of sound occurs through longitudinal pressure waves, along the direction of movement of particles. The rate of propagation is conditioned by the physical properties of the tissue, mainly by the resistance of the tissue to compression; it increases in the denser tissues and decreases in those with less density. In the parenchymatous tissues of the organism, the rates of propagation are similar to each other, so in diagnostic clinic an average speed of 1540 m/sec is considered. [21] The assumption that the rate of propagation is constant is essential for the ultrasound machine to estimate the distance or depth of a reflecting surface.

Another important physical property is acoustic impedance, which play a key role in the formation of echoes that can be used in ultrasound. It is measured in Rayl (1 Rayl = 1 kg/m2*s).

$$Z = \rho \cdot v \tag{2.2}$$

where v is the velocity of sound waves, rho is density of the medium.
Air, or structures containing a lot of air (such as the lung) have a very low acoustic impedance, while other tissues, except for bone, have a higher acoustic impedance. The production of ultrasounds takes place by piezoelectric effect [22], in which a crystal that is electrically excited to vibrate, placed in contact with the external surface of the medium, generating mechanical waves that propagate in the medium. When sound waves pass through biological tissues, part of the beam is transmitted, part is reflected, part is absorbed and part is diffused. These effects are described by Snell's law, which relates the angle of incidence and the angle of refraction with the speed of sound in the two media crossed. This law, in the context of ultrasound, helps to explain how ultrasonic waves bend when they encounter tissues with different acoustic densities or impedances (Z1 and Z2). It is defined as:

$$\frac{\sin \theta_i}{\sin \theta_t} = \frac{v_1}{v_2} \tag{2.3}$$

- θ_i: angle of incidence
- θ_t: angle of refraction
- v_1: speed of sound in the first medium
- v_2: speed of sound in the second medium

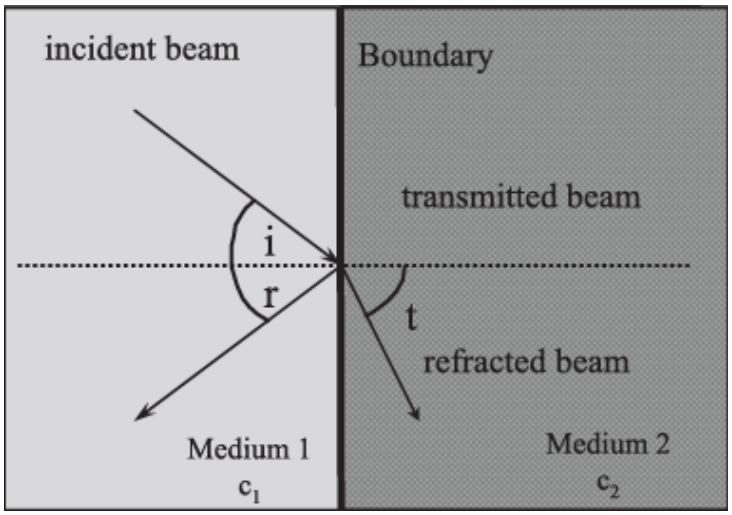


Figure 2.1: Snell's Law at the Interface Between Two Tissues.

Reflection is the phenomenon that consists in the origin of the wave returning to the transducer, forming an echo. It takes place with an angle that is equivalent to that incident of ultrasound. The amount of portion of the sound wave that is reflected back is described by the reflection coefficient:

$$R = \left(\frac{Z_2 - Z_1}{Z_2 + Z_1}\right)^2 \tag{2.4}$$

The part of the wave that is not reflected, is transmitted in the second medium and this is described by the transmission coefficient:

$$T = 1 - R \tag{2.5}$$

If the acoustic impedance of the two media is almost identical, the wave is transmitted totally($R \approx 0, T \approx 1$). In contrast, if the two means have very different acoustic impedances, the entire beam is almost completely reflected back ($R \approx 1$). This can happen when one of the two means is bone or air.
The intensity of the reflected wave is greater the more marked the impedance difference between the two interfaces and the more perpendicular the incident radius. The reflection is maximum when the wave hits the medium perpendicularly.
Refraction is the deflection that the ultrasonic beam undergoes after hitting the interface, passing through the structure. It occurs when ultrasound impacts at a different angle from the perpendicular on an interface between two media having different propagation speeds.
Another phenomenon that can occur is attenuation, the loss of acoustic energy which occurs then sound passes through tissues. [21] The main factors contributing to the attenuation are absorption, the conversion of acoustic energy into heat, reflection, and dispersion of echoes (scattering) of the ultrasonic beam. The attenuation follows the Lambert-Beer law:

$$A(z) = A_0 \cdot e^{-\alpha z} \tag{2.6}$$

- $A(z)$: amplitude of the wave at depth z
- A_0: initial amplitude of the wave at the surface
- α: attenuation coefficient [Np/cm or dB/cm]
- z: depth or propagation distance [cm]

The attenuation is directly proportional to the wave frequency and increases with the tissue depth; in practice, in a given tissue or medium, higher frequencies (7.5-10 MHz) provide a better resolution but are more easily attenuated than lower frequencies, making them suitable for the study of superficial organs. Whereas

lower frequencies (3- 3.5 MHz) are used for deeper organ analysis, because they penetrate more but with reduced resolution. [22] In addition, there is a need to compensate for the reduced amplitude of echoes coming from deep structures. This function is performed by an amplifier, with time-dependent gain, known as Time Gain Compensation (TGC).

2.2 Technical Components of Ultrasound Imaging

2.2.1 Ultrasound probes

The transducer (commonly referred to as the probe) plays the dual role of transmitting and receiving ultrasound through the use of piezoelectric crystals. The range of frequencies emitted by a particular probe depends on the characteristics and thickness of the crystals contained within it. [21]
The most commonly used probes are:

- Linear probes consist of multiple crystals, arranged on a line within a rod-shaped housing. Their main feature is that they offer high frequencies (from 6 to 18 MHz or higher), providing the best possible spatial resolution. They produce an image larger than the surface field and because they work at high frequencies and don't penetrate deeply, they are used for the study of superficial structures.

- Convex probes are linear probes with a convex curvilinear shape. Their design involves the formation of an image with a deeper field wider than linear ones and have different curvature radii, for example microconvex probes have a very small contact surface and are across a broad frequency range (from 5 to 12 MHz.), making them suitable for pediatric or intercostal imaging. Whereas those with a larger curvature are generally only available in medium and low frequencies (for example 1 to 8 MHz) and are commonly used in abdominal imaging of adult subjects.

- Phased array probes consist of piezoelectric crystals in a narrow rectangular or sector configuration. The beam is directed in different directions and can be focused on various depths. They have a frequency range of 2-10 MHz and are used in cardiac and intercostal imaging, in anatomically constrained regions.

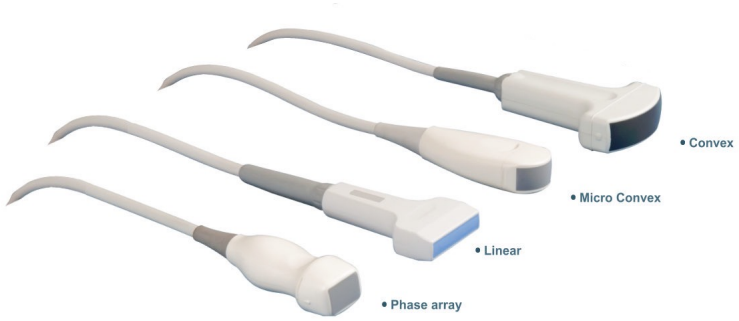


Figure 2.2: Common ultrasound transducer types. From left to right: phased array, linear, microconvex, and convex.

2.2.2 Echo Display Modes

There are several strategies for visualizing the ultrasound signal on the display. [21]

- A-mode (Amplitude mode): this is the simplest way to represent the ultrasound signal and using one-dimensional format. A single transducer is used to generate the signal, so it also allows the display of the single signal. This method is now rarely used because it provides limited information and is generally confined in very specific fields.

- B-mode (Brightness mode): this is the most used two-dimensional mode. Echoes are displayed as points of different brightness in gray scale corresponding to the intensity (amplitude) of each signal. The image obtained with B-mode is composed of multiple lines of piezoelectric crystals.

- M-mode (Motion mode): this mode is used to analyze moving structures. Echoes generated by a stationary transducer are recorded continuously over time and their changes are displayed along a time axis. The M-mode is obtained by combining B-mode sequences sampled at successive time points. It is a very important type of application in cardiology.

2.3 Artifacts in Ultrasound Imaging

Artifacts are errors that occur in image formation and are typically caused by physical phenomena which affect the ultrasound beam and alter the basic assumptions made by the imaging system. It is important for the sonographer to recognize

theme, because they represent inaccurate representations of anatomical structures and physiological function. [23]
There are several types of artifacts: some originate from the ultrasound equipment, others from the imaging technique, patient movement, or from the interaction between the ultrasound beam and tissues. Some of the most common artifacts include:

- **Reverberation artifact:** appears as multiple parallel and equidistant lines below a reflective surface. It is caused by sound waves that repeatedly bounces back and forth between two interfaces and then returns to the receiver. It may mask underlying structures.
- **Acoustic shadowing:** manifests as an area of low-amplitude or intensity echoes, which is located distal to highly attenuating structures. It appears as a dark area, limiting the evaluation of deep structures, but may also serve as a diagnostic indicator, for example in the detection of renal or biliary stones.
- **Posterior acoustic enhancement:** is seen as a localized increase in echo amplitude, which is located distal to highly attenuating structures. It appears as a dark area, limiting the evaluation of deep structures, but its presence can be a useful diagnostic indicator, for example in the detection of renal or biliary stones.
- **Ring-down artifact:** they appear as a series of echoes of closely spaced echoes. Ring-down artifacts occur when small crystals or gas bubbles resonate at the frequency of ultrasound and emit sound. These echoes are interpreted by the system as originating from deeper structures, often creating images resembling tunnels.
- **Comet tail artifact:** they also appear as a series of small echoes at close distance, discrete and luminous, which originate from a large difference in acoustic impedance between the metal and the fabric; it occurs in the presence of metallic objects or highly reflective surfaces, as in the case of surgical clips or needles.
- **Mirror image artifact:** The sound can bounce off a strong and smooth reflector, such as the diaphragm. The surface acts as a mirror and reflects the beam towards another tissue interface. The ultrasound system misinterprets the position of this second structure, generating a duplicate in an opposite location to the reflecting surface. This artifact is commonly seen in chest and abdominal ultrasounds.
- **Refraction artifact:** is a normal physical but sometimes can lead to artifacts; in some conditions, in fact, when the sound wave is transmitted to the second

tissue, it changes direction, causing an organ to be appear displaced or duplicated incorrectly. This type of artifact can make anatomical localization difficult.

- **Attenuation:** refers to the progressive loss of ultrasound beam intensity as it passes through tissue. Deeper tissues appear darker because the intensity of the transmitted beam is lower, which may reduce overall image quality.
- **Side lobe and grating lobe artifacts:** : result from the emission of secondary beams, which are not aligned with the main one. Even if they have a lower intensity, they can generate echoes that are misinterpreted by the system as coming from the main direction, producing incorrect spurious echoes in incorrect locations.
- **Speckle artifact:** : is a granular appearance seen in B-mode imaging, but if they are present in large quantities, they could reduce the sharpness of the contours and make measurements more difficult.

Chapter 3

State of the Art

3.1 Standard clinical method for IVC measurement

The evaluation of IVC using ultrasound is a fast, non-invasive and risk-free technique used in clinical practice to estimate right atrial pressure (RAP) and evaluate the patient's volemic status. [5] The patient is placed in a supine position and a convex or phased array probe is used. The probe should have a low frequency, usually 2 to 5 MHz, to allow better penetration and imaging of deeper structures. The probe is generally placed subcostally, using the liver as an acoustic window. [24] The IVC can be analyzed along the longitudinal and the transverse axis. However, the longitudinal axis is preferred for measurements of diameter and respiratory variability [25].

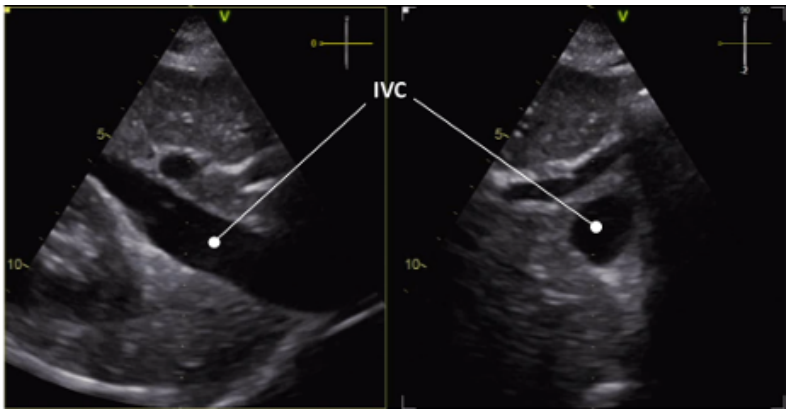


Figure 3.1: B-mode ultrasound views of the IVC: longitudinal section (left); transverse section (right).[5]

In a longitudinal view, the liver is in the upper region of the image and the IVC appears as a thin-walled, hypoechoic structure extending horizontally across the screen and entering the right atrium. Hepatic veins are observed that pass through the liver and enter the IVC. It is important not to confuse it with the aorta, which runs parallel to the IVC: the aorta has thick and hyperechogenic walls and presents a pulsatility, while the IVC has thinner and more compliant walls.

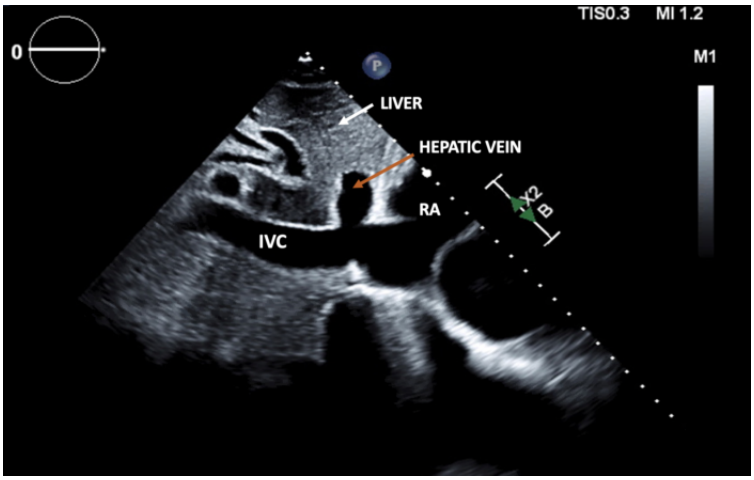


Figure 3.2: Longitudinal view of the IVC and right atrium (RA), with hepatic vein and liver.

Initially, B-mode ultrasound is used to identify the IVC in a longitudinal view. The M-mode is then used to record the variation in diameter over time. The diameter measurement is performed 1-2 cm caudal to the confluence of the hepatic veins, with the beam positioned perpendicularly to the vessel wall [19]. However, it is necessary a perfect alignment of the probe perpendicular to the long axis of the IVC, which is almost never possible because of the movements of the diaphragm due to the respiratory act and the heart, which also induces movements of the vein.

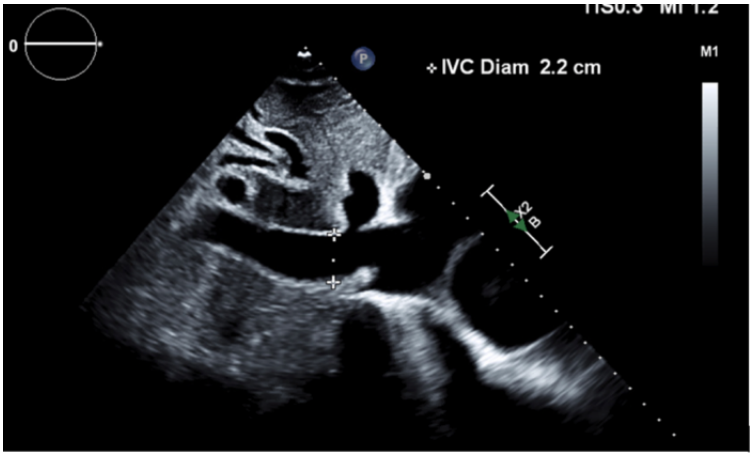


Figure 3.3: Measurement of the IVC diameter in B-mode 1-2cm below the hepatic vein confluence, perpendicular to the vessel wall.

The maximum diameter is detected at the end of the exhalation, when the intrathoracic pressure is higher, while the minimum diameter is recorded at the end of the inspiration, when the vein reaches its minimum width. These values are used to calculate the CI, as defined in Equation (2.1).

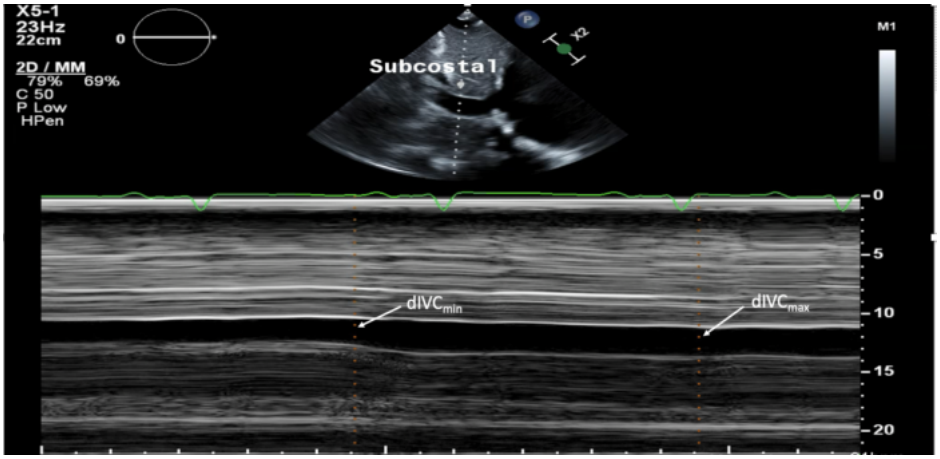


Figure 3.4: Respiratory variation in M-mode.

3.1.1 Limitations of manual measurement

Although IVC ultrasound is one of the most widely used methods for estimating the patient's body condition in critical settings, its application in manual measurements has several limitations that reduce clinical reliability.

- **Non-ideal section**: One of the main problems is the difficulty in ensuring that the scan is performed on the central and axial section of the IVC for longitudinal scans. In many cases the operator mistakenly takes a lateral and non-central section of the vein which can lead to underestimate the diameter, negatively affecting the accuracy of the measurement 3.5. Having an accurate real-time measurement of the diameter could help the operator more easily identify the center section, which is usually the one with the maximum diameter.

- **Restrictions of M-mode**: use of M-mode, often made for recording changes in IVC diameter with time, is inaccurate if not perfectly aligned along the longitudinal vein axis or transversal vein diameter.

- **Impact of breathing and heart movements**: M-mode uses a fixed axis, but the vein undergoes movements due to breathing and cardiac activity that introduce variability and lead to inaccurate diameter estimates. This study [26] proposed a solution using a non-fixed axis but solidly to the respiratory and cardiac movements. In addition, a study [27] highlighted the clinical value of the separate assessment of cardiac IVC collapsibility index. however, these advanced collapsibility indexes (RCI and CCI) can only be calculated from the analysis of the time series of diameters that can be obtained by automatic algorithms, not manually.

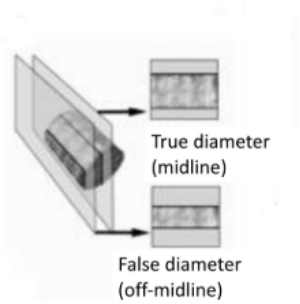


Figure 3.5: True vs false central section of the IVC.[28]

3.2 Repeatability of semi-automatic and real-time software

3.2.1 Theoretical Background on Measurement Error

Clinical practice involves measuring quantities for multiple purposes, such as facilitating diagnosis, monitoring the patient, predicting future outcomes, and serving as endpoints in randomized trials or studies. However, measurements are almost always subject to errors, resulting in a discrepancy between the observed value and the true value; as a result, studies investigating the measurement error frequently appear in scientific literature. As pointed out by Bartlett and Frost (2008), the impact of measurement error depends on the application context. For example, a certain degree of measurement error may be acceptable in a comparative study such as a clinical trial, while it may be unacceptable in individual patient management, such as screening or risk prediction. For this reason, it is essential to evaluate the variability associated with a measurement method by distinguishing some fundamental concepts such as repeatability, reproducibility, agreement and reliability. [29]

- **Repeatability** indicates the consistency of repeated measurements taken on the same subject, using the same instrument, by the same operator and under the same conditions. It represents the measurement error under ideal and controlled conditions.

- **Reproducibility** refers to the variation in measurements made on a subject under variable conditions; therefore, such as changes in the operator, equipment, or time of acquisition.

- **Agreement** measures how two different methods provide similar values on the same subject. It is particularly relevant when comparing different methods (e.g. manual vs automatic), where there will be differences both due to the inherent variability of each of the measurement methods and the presence of systematic biases.

- **Reliability** describes how much the measurement error affects the total observed variability and, therefore. It reflects the ability to distinguish between subjects based on the measurement's values and it depends not only on the accuracy of the method, but also on the heterogeneity of the population in which it is applied. It is often expressed through the Intraclass Correlation Coefficient (ICC).

3.2.2 Automated Measurement Software in Clinical Applications

In recent years, numerous software have been developed in the clinical field, with the aim of automating measurements by reducing operator-dependent variability. These systems are usually based on artificial intelligence techniques or automatic tracking algorithms, which have been useful in improving measurements repeatability and accuracy. The following are studies which have had as objectives those mentioned above. In cardiology, Leclerc et al. (2019) proposed a deep learning system for automatic segmentation of heart cavities on 2D ultrasound images with a CAMUS dataset (500 patients, two- and four-chamber acquisition). The study evaluated intra- and inter-operator variability compared to three experts, obtaining high ICC values (>0.98) and an average absolute error below 10 ml for the ejection fraction. [30] Similarly, Kim et al. (2022) evaluated three different deep learning algorithms on a large dataset containing two- and four-chamber images from 500 patients. They reported Very high Dice coefficient (0.91-0.95) and intra (>0.98) and inter-operator ICC (>0.97). [31]
In the vascular context, Lainé et al. (2022) developed a caroSegDeep algorithm capable of automatically segmenting the carotid wall from B-mode ultrasound sequences. It was trained on more than 2000 images taken by experts, demonstrating high accuracy and low with an average absolute error of less than 120 μm, which is lower than the typical inter-observer variability (180 μm). [32]
However, specific applications for the IVC remain limited. In this context, VIPER introduces an innovative method with potential improvement in terms of accuracy and speed for real-time automatic IVC analysis.

Chapter 4

Experimental Protocol

4.1 Aim and Scope of the study

The aim of this study is to evaluate the repeatability and reliability of the IVC diameter measurements and other derived parameters such as the Caval Index (CI), the Respiratory Caval Index (RCI), and Cardiac Caval Index (CCI), obtained using the VIPER software in real-time mode.
In particular, the objective is to evaluate how the visual feedback given by the software during acquisition may constitute useful support for obtaining more repeatable and stable measurements. Intra-operator and inter-operator variability were also assessed, and a comparison between the software-assisted method and the traditional manual method was performed.

4.2 Experimental Protocol

In order for the subjects to be in the most homogeneous conditions possible, they were asked to maintain good hydration on the previous day, consuming at least 2 liters of water.
The protocol had a total duration of 90 minutes and was structured in 4 different acquisition phases:

- Supine without feedback (SUPINE-NOF).
- Supine with feedback provided by VIPER software (SUPINE-VIPER)
- PLR (Passive Leg Raising) without feedback (PLR-NOF)
- PLR with feedback provided by VIPER software (PLR-VIPER);

4.2.1 Execution of the Protocol

The protocol was implemented following sequential steps:

1. Compilation of personal data: Each subject has compiled an Excel file, in which they were inserted: age, gender, height, weight, personal habits (e.g. sport, smoking) and the presence of any relevant pathologies.

2. Ultrasound parameters setting: the subject was positioned supine for 5 minutes, to have an even fluid distribution. During this time, an ultrasound was made in longitudinal section to set the echograph's parameters (frequency, depth, gain, TGC).

3. First acquisition session: no feedback supine (NOF). This phase involved the acquisition of ultrasound videos in the supine position, without visual feedback from the software. Operator A (OA) started to acquire the first acquisition along the longitudinal axis, followed by a pause of approximately 15 seconds for the change of the operator. Then, Operator B (OB), performed the same sequence (Repetition 1). Both operators made a second acquisition (Repetition 2). Each recording lasted between 10 to 15 seconds, at the end of which the video was saved in the corresponding patient folder. At the end of this first phase, the subject was made to stand for 2 minutes to avoid a slow change in the distribution of fluids in the body.

4. Second Acquisition Session: supine with feedback (VIPER). This phase involved the acquisition of video in a supine position with feedback from the Viper software. The patient was placed again in a supine position and waited 1 minute to allow a new fluid distribution before starting to acquire. As in the previous phase, operator A (OA) performed the acquisition, followed by operator B (OB). Both operators then repeated the process for the second acquisition (R2). In this phase, the operator who was doing the ultrasound, once satisfied with the quality of the ultrasound image, activated the VIPER software and after about 10-15 seconds, the software was stopped, and the excel file obtained by VIPER was saved in the folder corresponding to the patient and the condition. At the same time, the second operator saved the corresponding video. At the end of this phase, the subject stood up again for 2 minutes.

5. Third session: PLR without feedback (NOF) After the standing phase, the subject was placed in the PLR position, with the legs raised at a 45-degrees angle. It took 3 minutes before starting a new acquisition to pass the initial hemodynamic peak associated with increased venous return. This session was conducted without visual feedback, as in the first session. Also, in

this condition each operator performed two repetitions. At the end of each acquisition, the video has been saved in the corresponding folder.

6. Fourth session: PLR with feedback (VIPER) After another 2 minutes in standing position and 3 minutes of adaptation in PLR position, the last phase began. In this case, the acquisitions in PLR were carried out with the aid of VIPER software. The execution modalities were identical to the previous phase. Operator A (OA) performed two acquisitions, followed by operator B (OB), (R1 and R2). In each acquisition, once a satisfactory image has been obtained, the VIPER software was activated for about 10-15 seconds in real-time mode. When finished, the Excel file generated by VIPER was saved in the folder corresponding to the patient and the specific condition. Meanwhile, the second operator saved the video of the capture.

The two different conditions, supine and PLR, were included to see how different patient conditions and posture influence the measurements obtained by the software. During the PLR phase, the patient was asked to maintain a relaxed position to avoid contracting abdominal muscles. This was necessary to prevent possible alterations of the ultrasound image caused by involuntary muscle tensions.The sequence of the phases has been defined intentionally: the acquisitions without feedback were conducted first to avoid that the visualization with the software in real time could influence the behavior of the operators in the measurements. In addition, the operators alternated between the first and second repetition (R1 and R2) to minimize and reduce the risk of remembering exactly the anatomical point selected in the previous repetition, thus avoiding unintended influence on subsequent measurements. The scan's duration of 10-15 seconds was chosen to ensure that at least one or more complete respiratory cycles were acquired.

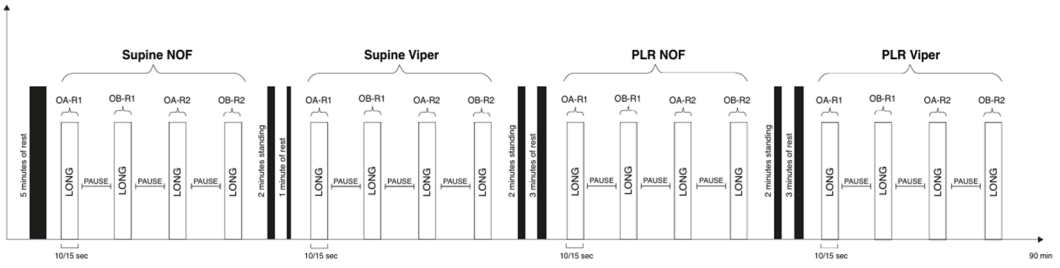


Figure 4.1: Experimental Protocol

For ultrasound imaging, a convex probe was positioned subcostally. For longitudinal acquisitions, the probe's orientation marker (notch) was chosen towards the patient's head (cranially). In this configuration, the right atrium appears on the left side of the screen. It is essential to check that the probe's orientation does not change, to avoid an inversion of the image. The main acoustic window used is represented by the liver. The hepatic vein was used as an anatomical reference point, so that each time the same portion of the IVC was selected.

4.2.2 Ultrasound Probe Positioning

For ultrasound imaging, a convex probe was positioned subcostally. For longitudinal acquisitions, the probe's orientation marker (notch) was chosen towards the patient's head (cranially). In this configuration, the right atrium appears on the left side of the screen. It is essential to check that the probe's orientation does not change, to avoid an inversion of the image. The main acoustic window used is represented by the liver. The hepatic vein was used as an anatomical reference point, so that each time the same portion of the IVC was selected.

Chapter 5

Methodology

5.1 Materials

The following tools were used to carry out the experiment:

- Medical bed: used to position the patient in the supine or PLR position during ultrasound scans.
- MicrUs EXT: It is a portable, compact and economical ultrasound system that transfers images to smartphones, tablets or PCs. It consists of a probe port (MicrUs EXT-1H) and the probe used was a convex model C5-2R60S-3. The probe has a frequency range of 2-5 MHz, a radius of 60mm and a scanning depth of 60-230mm. It is used for abdominal, obstetrics/gynecology (OB/GYN) and pediatric ultrasound. Before using the probe, a conductive gel for ultrasound was applied. [33]
- Echo Wave II software (X64, version 4.3.0): This is the ultrasound image acquisition software installed on the computer connected to the ultrasound device.
- VIPER software: Used for real-time tracking of the IVC, providing visual feedback to the operator.
- Two computers:
  - The first computer was connected directly to the MicrUs EXT ultrasound machine and allowed the visualization of the ultrasound image through the Echo Wave II software.
  - The second computer was connected to the first via a system consisting of a frame grabber and HDMI cable, allowing to view the ultrasound video

of the first computer. The VIPER software, used for IVC tracking, was installed on this second computer.

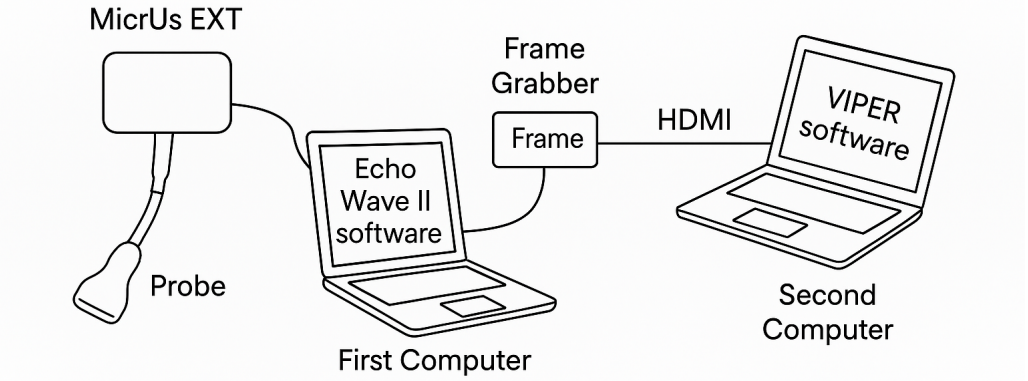


Figure 5.1: Experimental setup for ultrasound acquisition and real-time IVC tracking

5.2 Methods

5.2.1 Dataset

The study sample consisted of 26 healthy subjects: 15 females and 11 males, aged between 24 and 29 years (25 $\pm$ 1.37 years). Each subject was asked about their weight and height, smoking habits, physical activity (including type), and the presence of any known pathologies. Among them, 9 subjects reported they did not practice any sport, 16 declared they did not smoke, and none reported any known medical conditions.

5.2.2 Telemed

For the acquisition of ultrasound videos of the IVC, the EchoWave II software, produced by Telemed, was used, which allowed one to select the various ultrasound parameters and to record videos in B-Mode. The system was combined with a

convex-type probe (C5-2R60S-3), with frequency varying between 2 and 4 MHz, which is suitable for IVC exploration, allowing good penetration of deep tissues. All acquisitions were performed in B-Mode, as foreseen by the morphology and dynamics of the IVC. The software interface provides a quick adjustment of parameters(Figure 5.2).
To ensure an optimal image quality, at the beginnig of the protocol for each patient the parameters were set, depending on the anatomical characteristics of their vena cava. The scanning depth, expressed in millimetres (mm), adjusts the depth of the viewing field; it is increased to examine larger or deeper structures, while it is reduced to examine more superficial structures. For subjects with different body characteristics, different scan depth values were set ranging from 120 to 180 mm depending on the abdominal wall thickness and subcutaneous adipose tissue. This allowed better visualization for the different subjects. Then the focus depth was set between 64 and 80 mm, to concentrate the area of maximum sharpness on the longitudinal axis of the vein. The acoustic output power, which controls the strength of the emitted ultrasound signal; the power was maintained at -4 dB, to limit any unwanted reflections and artifacts. The gain parameter amplifies or reduces feedback signals and increases or decreases the amount of information on the image and it used to make clearer image or more excuse depending on the structure that is displayed. The gain values are between 10 and 100 % and it was adjusted between 85 % and 95% depending on the abdominal thickness and the depth of the IVC. The dynamic range, measured in decibels (dB) controls how the ultrasound intensity is converted to grayscale. It was set to a value of 72 dB, which allowed maintaining a good contrast between tissues with similar echogenicity, enhancing overall image clarity. Each video was saved in .mp4 format to be later processed in post processing using the VIPER software.

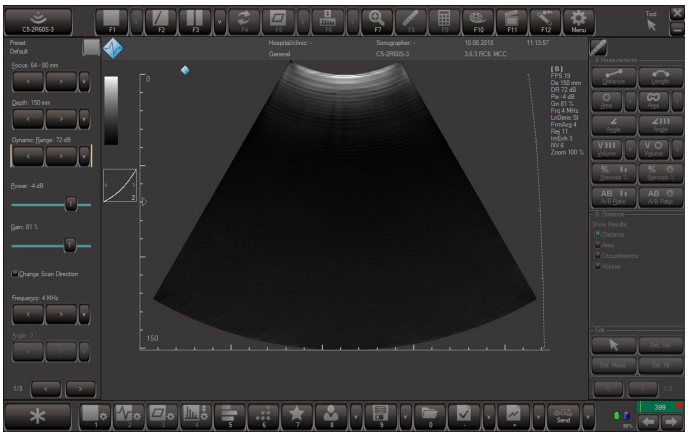


Figure 5.2: Interface of the EchoWave II

5.2.3 VIPER

Description of the Software and Tracking Functionality

The manual ultrasound measurement of the IVC is subject to several criticalities that compromise its reliability in clinical practice. In particular:

1. The IVC has an irregular geometry; its cross-section may vary along the longitudinal axis and is not circular in shape, making it difficult to obtain a representative estimate.

2. During the acquisition, the position of the vein may shift due to respiration, making it difficult to keep the same portion of vein visible in all frames.

3. There is poor repeatability, both intra- and inter-operator.

To overcome these limitations,the VIPER software was developed. The software version (Software version v2.3.2 pro) used in this study allows real-time IVC tracking, making it a particularly useful tool for fluid therapy assessment. Compared to previous versions, the diameter is calculated for each frame, thus obtaining a time series of diameters for each ultrasound video. The resulting diameter signal contains two main components:

- a low-frequency respiratory component (< 0.4 Hz)
- a high-frequency heart component (> 0.8 Hz).

Through low-pass and high-pass filters these components are separated: from the raw signal is obtained the CI, while from the filtered signals the RCI (Respiratory Collapsibility Index) and the CCI (Cardiac Collapsibility Index) are obtained respectively. 5.3 [34]

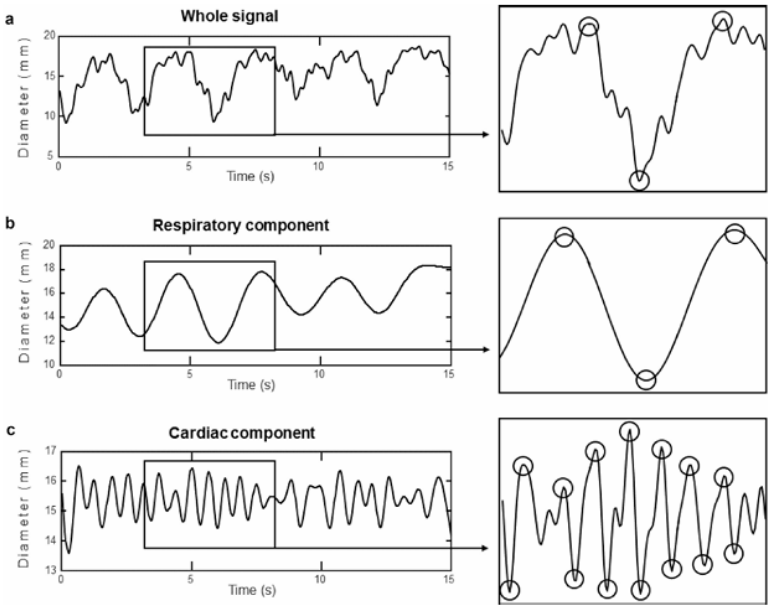


Figure 5.3: Decomposition of the IVC diameter signal into its main components. **(a)** Whole signal and extimation of CI. **(b)** Respiratory component extracted by filtering the low-frequency content. **(c)** Cardiac component extracted from the high-frequency part. [34]

The software applies edge detection methods to detect the upper and lower margins of the vein along a defined line. For each frame, the position of the vein is realigned with respect to a moving reference system so that it does not suffer from the effects of breathing or involuntary movements of the probe. This allows to keep the same region analyzed over time, thus allowing a reliable estimation of the diameter even in the presence of variations. At the selected section of the vein, the software constructs an interpolated mean line between the middle points of the upper and lower edges. Along this line 21 equidistant points are placed, from each of which a perpendicular is drawn that intersects the margins of the vein. The length of these perpendiculars represents the local diameter; the average of the 21 values gives the overall diameter of the IVC for each frame. [35]
Before the analysis, it is necessary to calibrate to convert pixel-based measurements into millimeters. This is performed manually by selecting a 50mm segment on the depth scale visible on the side of the ultrasound image. After the segment is selected, the user must click the button *Compute Conversion Factor* in the software interface. This conversion is important because the diameters initially calculated are expressed in pixels, that is units relative to the digital image, but to compare them with clinical data it is necessary to express them in millimeters.

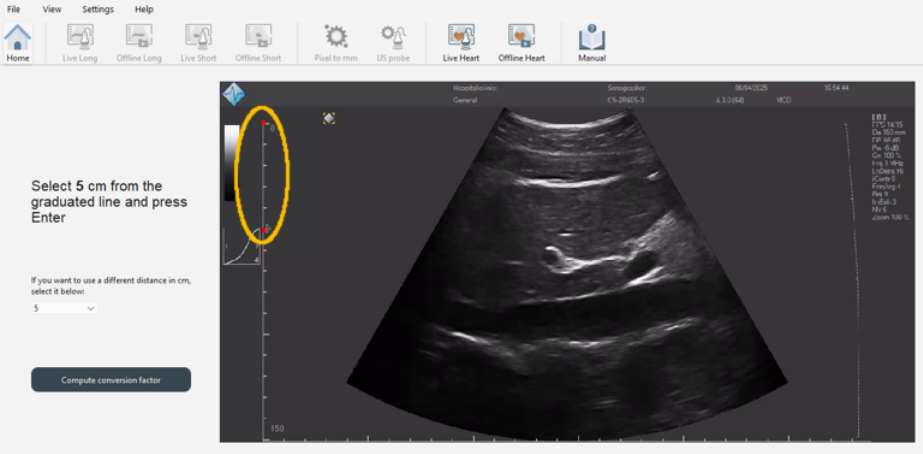


Figure 5.4: VIPER software interface during the calibration phase. The user selects a 5 cm segment from the graduated scale on the left and then clicks on the Compute Conversion Factor.

After calibration, the software shows the screen that invites the user to click on the section of the IVC to be analyze, so on the first frame of the video the user chooses the section of IVC that he wants to analyze, clicking with the cursor on the desired area. (Figure 5.5). The tracking is then automatically initiated. (Figure 5.6). During the analysis with the VIPER software, two graphs are shown in real time on the left of the interface: the first shows the evolution of the diameter and the average diameter of the IVC; the second represents the central venous pressure (CVP) through a red horizontal bar that fills up according to the estimated value. The three reference bands are: normal (< 5 mmHg), intermediate (5-10 mmHg) and high (10-20 mmHg). This indicator provides an immediate estimate of the patient's volemic status.
At the end of that, an excel file is generated containing the series of diameters and parameters with CI,RCI,CCI.

Figure 5.5: After the calibration step, the user click on the desired section of the IVC to be analyzed. The yellow dot, added for illustrative purposes, highlights the point selected by the user on the first frame of the video.

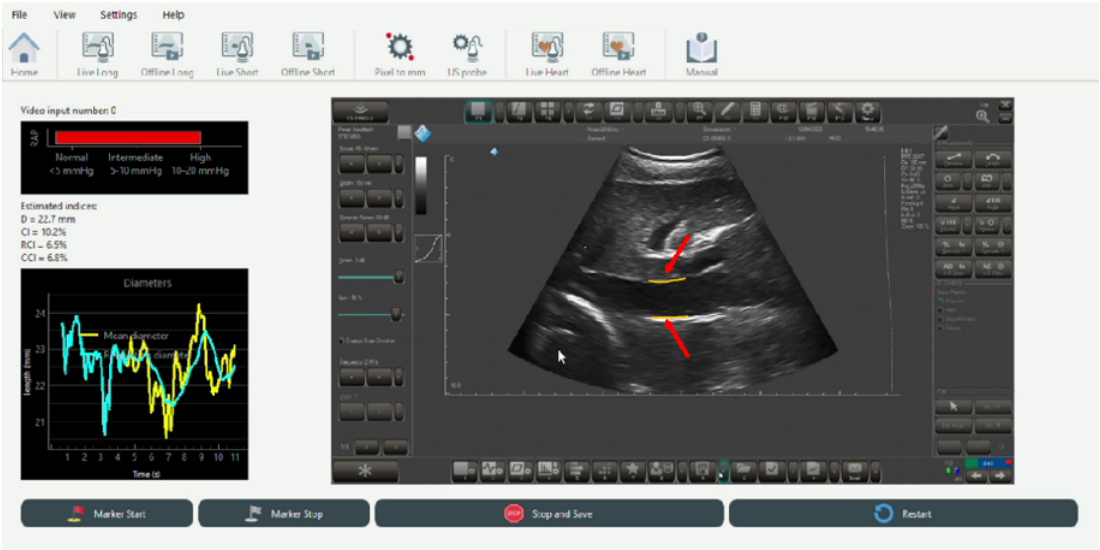


Figure 5.6: IVC tracking process: the yellow lines represent upper and lower borders of the IVC indicated by red arrows (manually added for clarity). The graphs on the left show the temporal evolution of the diameters extracted from the video and CVP.

User Interface and Control Buttons

At the top of the VIPER software interface, there is a toolbar that provides quick access to key functions. The following buttons are available:

- *Live Long/ Offline Long*: allow the acquisition in real time or the offline loading of ultrasound videos in the longitudinal section.
- *Live Short / Offline Short*: allow the acquisition in real time or the offline loading of ultrasound videos in the cross section.
- *Pixel to mm*: Activates the calibration tool, which is necessary to convert measurements expressed in pixels into millimeters.
- *US probe*: Access the ultrasound probe settings.
- *Live Heart/ Offline Heart*: mode dedicated to cardiac ultrasound acquisition, respectively, in real time or from recorded files.
- *Manual*: opens the software user's manual, containing instructions on how to use the various functions.

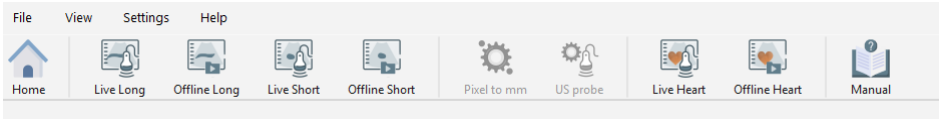


Figure 5.7: Toolbar at the top of the VIPER software interface.

During the analysis, the user can also use the buttons at the bottom of the interface to control the process. In particular:

- *Marker Start* is used to insert a start mark,
- *Marker Stop* to indicate the end of use of the marker,
- *Stop and Save* allows you to stop processing and save the excel file containing the results and the corresponding tracking video,
- *Restart* resets the tracking and allows to start again.

Figure 5.8: Buttons available during the analysis with VIPER.

5.2.4 Python GUI for manual measurements

During the experimental phase, the measurement of the IVC diameter was initially carried out using the Telemed software environment connected to the ultrasound system used. However, this mode had some operational limitations that hindered the efficiency of the acquisition protocol. In particular:

- The software allowed manual measurements only on static images obtained with the freeze command, making it impossible to measure directly inside a sliding cineloop;
- It was not possible to load previously saved cineloops for subsequent measurements;
- There was no efficient way to save the measurements, but the software allowed saving in the form of a screenshot, so the values had to be extracted manually, a long and prone to errors process. In addition, a tool for the automatic calculation of the CI was not included, but this also had to be done separately.

These restrictions meant that, for each measurement, it was necessary to stop the acquisition, freeze the image and manually carry out the measurement, resulting in a slower protocol. To overcome these problems, a simple software tool was developed in Python with a graphical interface made using the tkinter module [36]. The main objective was to reproduce the manual measurement method autonomously, allowing the operator to perform measurements after capturing the video. This made it possible to make the acquisition phase more fluid, separating the moment of recording from that of analysis, while still maintaining the visual control of the operator on the choice of frames to be measured.

Figure 5.9: GUI for Manual Measurements of the IVC

The GUI presents a main window (Figure: 5.9) structured in a simple and functional way, including the following buttons and tools:

- Select video: opens a dialog window for uploading an ultrasound video saved on your computer. The uploaded file is displayed frame by frame in the central GUI window;

- Calibration: allows the user to define the scale factor by clicking two known points on the image (e.g., referring to an ultrasound grid or marker). The software calculates the mm/pixel conversion factor, which will be used to convert all subsequent measurements;

- Frame slider: after calibration, the slider allows you to interactively scroll through the frames of the cineloop to find the most appropriate moment to make the measurement;

- Selection of measuring points: by left click, the user can select two points corresponding to the anterior and posterior walls of the vena cava in the displayed frame. After the selection of the first point, a segment is drawn which the operator must make as perpendicular to the vessel lumen during the measurement , which visually guides the selection of the second point;

- Valid and save: calculates the metrics of interest (maximum, minimum and mean diameter, collapsibility index - CI - and time taken for measurement) and saves them in .json format. The data is associated with the analyzed file and stored for subsequent statistical analysis.

- Reset: allows to reset the calibration and to repeat the process from scratch, useful in case of error in the selection of scale points;

- Close: closes the current video window, allowing you to select a new file for analysis.

The GUI also includes informative labels showing the current frame number, conversion factor, calculated diameters (mm) and corresponding frame, CI and time taken to measure (Figure: 5.10).

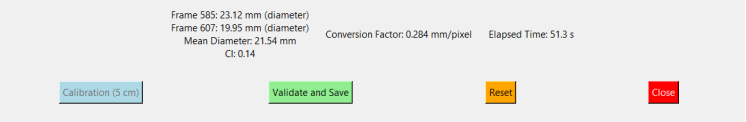


Figure 5.10: GUI: Buttons and Labels

5.2.5 Intra and Inter-operator variability

The evaluation of intra and inter-operator variability is essential to determine the repeatability and reliability of ultrasound measurements, both manual and automated. Intra-operator variability refers to how different are the measurements performed on the same subject by the same operator on different acquisitions, whereas inter-operator variability refers to the discrepancy between measurements performed by two different operators under the same conditions.
An evaluation was performed to verify that the software was reliable and repeatable in relation to the selection of the initial point for segmentation. For this purpose, each video was processed five times using the software in post-processing by each operator. This allowed us to assess whether the same operator constantly tends to select the same starting point, or whether this choice is also consistent and similar between different operators.
For each measurement modality, specific comparisons were carried out to assess intra and inter-operator variability:

1. **Post-processing with VIPER (NOF) vs Manual Method**
   This comparison aims to assess whether the use of software in general, even in the absence of real-time feedback, allows for greater repeatability and consistency in measurements compared to the manual method.
   Manual measurements, performed by each operator on appropriately selected frames, were compared to measurements obtained from the post-processing of videos acquired without feedback and processed using VIPER. One of the five post-processing iterations was randomly selected for each case. This analysis was performed on the average diameter and the CI.

2. **Post-processing with VIPER (NOF) vs VIPER used in real-time (REAL)**
   This comparison analyzed how the use of the real-time VIPER interface leads to improvements in terms of measurement variability.
   Measurements acquired with feedback were compared to randomly selected post-processed measurements from videos acquired without feedback. This comparison was conducted on the following quantitative parameters, extracted from the Excel output generated by VIPER: average diameter, caval index (CI), cardiac caval index (CCI), and respiratory caval index (RCI).

For each variable and comparison, the following indices were evaluated:

- **Coefficient of Variation (CoV)**:
  The coefficient of variation (CoV) is a statistical measure that represents the ratio between the standard deviation and the average of a data set. It is

calculated as follows:

$$\text{CoV} = \frac{\sigma}{\mu} \times 100$$

where:

- σ is the standard deviation of the measurements;
- μ is the mean value of the same measurements.

It is provided to assess the relative variability of data points in relation to the average. It is useful to compare the degree of variation between different data sets. In the repeatability analysis, a lower CoV means less variability and greater consistency between measurements, reduced absolute variability relative to the mean.

Boxplots were used to represent the different CoV values and to compare them. These plots are graphical representations of the distribution of a data set, showing the central trend, variability and potential outliers. These graphs allowed to visualize of the overall distribution of the CoV values within each group (NOF, Real Time, Manual).

In each boxplot:

- the **median**, shown as a center line within the box;
- the **lower and upper quartiles**, representing the 25th and 75th percentiles, respectively;
- the **whiskers**, which extend to the most extreme values not considered outliers;
- the **outliers**, which are plotted individually using the symbol '+'.

These graphs allow an immediate visual comparison between the various acquisition groups.

- **Intraclass Correlation Coefficient (ICC)**:
  Intraclass Correlation is a statistical measure used to assess the reliability or consistency of measurements performed by different operators measuring the same quantity. [37]
  In this study, the ICC(2,1) model was adopted which corresponds to the 'A-1' option of the Matlab function [38], referring to the classification proposed by Shrout and Fleiss (1979)[39] [37].
  This choice was made based on the following considerations:
  - The operators were considered randomly sampled from a large population.
  - A single measurement per subject per operator was evaluated (single measurement).

– The absolute agreement criterion was adopted, meaning that it requires that the measurements coincide in absolute value, and not just in consistency.

The ICC (2,1) was evaluated with the following formula:

$$\text{ICC}(2,1) = \frac{MS_R - MS_E}{MS_R + (k-1) \cdot MS_E + \frac{k \cdot (MS_C - MS_E)}{n}}$$

– MS_R is the mean square between rows, representing the variability between subjects;
– MS_E is the residual mean square (error);
– MS_C is the mean square between columns, representing the variability within subjects;
– k is the number of raters
– n is the number of subjects.

5.2.6 Statistical analysis

Linear Mixed Effects Model

To demonstrate the impact of condition and feedback factors on the measurement, a mixed-effects linear model approach was employed. This approach was particularly appropriate for the experimental design, which involved repeated measurements of the same subjects and required an accurate modeling of the correlation among non-independent observations. Mixed-effect linear models extend the classic linear models by introducing, in addition to fixed effects, also random effects. Fixed effects are parameters associated with experimental factors of general interest, repeatable, and known (e.g., experimental condition, presence or absence of feedback), whereas random effects represent sources of variability associated with randomly selected experimental units from a larger population (e.g., subjects, operators, measurement repetitions). In this context, random effects make it possible to model intra-subject and inter-subject variability and take into account the hierarchical and dependent structure of the data. In particular, each subject represents an experimental unit that can show a different response to the same fixed factors, and this is captured by the model, including a random term for the subject and, as nested terms, the operator and repetition. Mixed effect models assume:

- Linearity of the relationship between independent variables and response.
- Normality of the residuals, that is, the errors are distributed according to a normal distribution with mean zero and constant variance.

- Independence between random effects and residues, and between different random effects.
- Homogeneous variance of residues (homoscedasticity). [40]

In fact, as will be explained later in the results, linear models, especially those with mixed effects, exhibit good robustness even in the presence of slight violations of the normality assumption, particularly when the number of observations is sufficiently large [41]. The parameters were estimated using the REML method, which is based on maximizing restricted likelihood (a modified version of likelihood) and is preferable to maximum classical likelihood (ML) in the presence of random effects, because it provides less skewed estimates of variance.[40] The model adopted for general analysis can be expressed in the following analytical form:

$$Y_{fcijk} = \beta_0 + \beta_1 \cdot F_f + \beta_2 \cdot C_c + \beta_3 \cdot (F \cdot C)_{fc} + u_i + u_{ij} + u_{ijk} + \varepsilon_{fcijk} \tag{5.1}$$

- Y_{fcijk}: Response variable (dependent variable)
- β_0: intercept (overall mean)
- $\beta_1 \cdot F_f$: Fixed effect of factor Feedback f
- $\beta_2 \cdot C_c$: Fixed effect of factor Condition c
- $\beta_3 \cdot (F \cdot C)_{fc}$: Fixed interaction effect between Feedback and Condition
- u_i: random effect of factor Subject i
- u_{ij}: random effect of factor Operator j nested within i
- u_{ijk}: random effect of factor Repetition k nested within j
- ε_{fcijk}: Residual error term (within-group error)

This analysis allowed for a formal test of whether the condition, feedback, and their interaction had a statistically significant effect on the measurements of diameter, IC, RCI, and CCI, taking into account inter-subject variability.
The p-values for fixed effects have been calculated using an estimation of degrees of freedom based on the Satterthwaite approximation [42]. This method allows a more accurate assessment of statistical significance in the presence of nested random effects and moderate sample sizes. The Satterthwaite method provides a rough estimate of the degrees of freedom for each fixed effect contrast, taking into account the estimated variance of the random effects and their contribution to the overall model uncertainty. For each estimated coefficient, an expression of the variance of the estimate is constructed, and degrees of freedom are obtained as:

$$\mathrm{df} = \frac{(\mathrm{Var}[\hat{x}])^2}{\mathrm{Var}\left[\mathrm{Var}[\hat{x}]\right]} \tag{5.2}$$

where $\mathrm{Var}[\hat{x}]$ is the estimated variance of the coefficient, and $\mathrm{Var}(\mathrm{Var}[\hat{x}])$ is the variance of that estimate, which depends on the variance components of the random effects.
Following the results of this first linear mixed-effects model, we performed post-hoc analyses on the factors involved to explore the differences between specific conditions in more detail. In particular, contrast tests were conducted between specific levels of the Condition factor within different levels of Feedback, and vice versa. The contrasts were implemented using the Matlab `coefTest` function [43], which allows the testing of linear combinations of the coefficients estimated by the model. Such comparisons were conducted not only for those with a p-value less than 0.05, but also for those in the range of 0.05 to 0.08, to investigate potentially relevant trends that could be significant in a study with higher statistical power. This approach can be justified in the exploratory phase, as suggested by [44] , especially when the objective is to better understand the behaviour of experimental conditions and generate hypotheses for future investigations. Indeed, no multiplicity corrections have been applied for this purpose.

Cohen's d

Next to compare the different modes of acquisition (manual, with real-time feedback, without feedback) in terms of discriminability between two conditions (PLR and Supine), a similar model was adopted but focused only on the Condition variable, whose analytical writing is:

$$Y_{cijk} = \beta_0 + \beta_1 \cdot C_c + u_i + u_{ij} + u_{ijk} + \varepsilon_{cijk} \tag{5.3}$$

- Y_{cijk}: Response variable (dependent variable)
- β_0: Intercept (overall mean)
- $\beta_1 \cdot C_c$: Fixed effect of factor Condition c
- u_i: random effect of factor Subject i
- u_{ij}: random effect of factor Operator j nested within i
- u_{ijk}: random effect of factor Repetition k nested within j
- ε_{cijk}: Residual error term (within-group error)

The results obtained from this model were used to calculate the Cohen's d dimension effect, in order to compare the discriminative capacity between PLR and supine in the different measurement approaches (Manual, Real-Time feedback, NO feedback), by assessing which mode of measurement was more sensitive to the change in condition. The model has been applied to all parameters for Real-time and no-feedback cases, while it has only been applied to diameter and CI for the manual case. In the context of paired measurements, where each subject is observed under both conditions, Cohen's d has been calculated using the following relation [45]:

$$d = \frac{t}{\sqrt{n}} \tag{5.4}$$

where t is the t-statistic associated with the fixed effect of the Condition (PLR vs Supine), as estimated by the mixed-effects model, and n is the number of subjects. This formula is analytically equivalent to the paired Cohen's d, defined as the mean difference divided by the standard deviation of the differences between subjects, and it is suitable when the statistic t is obtained from a model which takes into account repeated measurements and random effects.
Cohen (1988) proposed that Cohen's d values can be interpreted as follows:

- $d \simeq 0.20$: *small effect*
- $d \simeq 0.50$: *medium effect*
- $d \simeq 0.80$: *large effect*

However, Cohen himself cautioned that these thresholds are relative to the context, the relevant discipline, and the experimental design.

Bland-Altmann Analysis

The Bland-Altman method is a statistical technique used to assess the agreement between two measurement methods. It is based on the construction of a Cartesian diagram, called the Bland-Atman Plot, where the differences between the two measurements are plotted on the y-axis (ordinates), and their mean is plotted on the x-axis (abscissas).[46] This approach allows you to visualize the degree of agreement or disagreement between two techniques, as well as to identify any systematic biases, outliers (outlier) and patterns of disagreement. In particular, The plot includes a bias line, representing the average differences between the two methods, and the concordance limits (LoA), defined as bias 1.96 DS, which contain about 95% of the differences, assuming a normal distribution of errors. [47] Points falling within the limits of agreement indicate that the two methods provide congruent results, while the points outside represent cases of discrepancy. In the case of perfect agreement, the bias would be zero, and the points align along the x-axis at a difference value of zero.

Chapter 6

Results and discussion

6.1 Mixed effects analysis

The initial analysis of the data used a linear model with mixed effects in order to verify whether the presence of feedback affected the measurements collected. The conditions for use of the model were met, with the sole exception that the residues were not found to be normal. Although the Lilliefors test did not support the normality hypothesis, the observation of the Q-Q plot showed only slight deviations in the tails 6.1). It was therefore considered that this breach was not substantial and did not affect the validity of the model.

The reported results in tables 6.1, 6.2, 6.3, 6.4 are derived from a marginal F-test type III (marginal ANOVA), calculated on the linear mixed effect model using the Satterthwaite method for approximation of degrees of freedom. The F statistic tests whether each effect significantly contributes to explaining variance in the dependent variable, accounting for the model structure. DF1 and DF2 represent the degrees of freedom associated with the fixed term and with the estimation of residual variance, respectively, based on the Satterthwaite approximation.

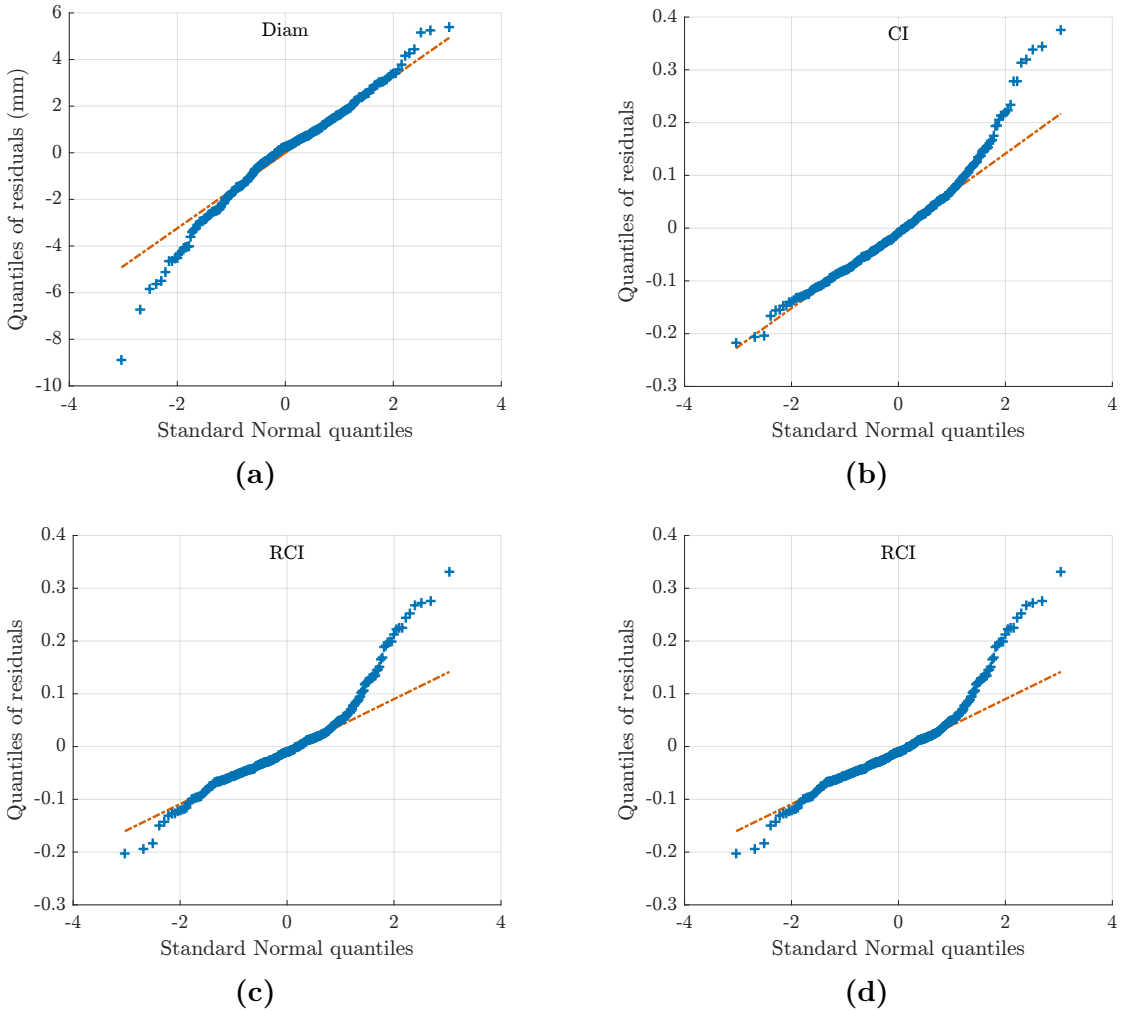


Figure 6.1: Normal Q-Q plots of residuals for each variable: (a) Diam, (b) CI, (c) RCI, and (d) CCI. The plots evaluate the assumption of normality in the linear mixed-effects model.

Table 6.1: Results of the linear mixed-effect model for fixed factors Condition and Feedback and their interaction on **Diameter**

Term	**F Stat**	**DF1**	**DF2**	***p*-value**
Condition	195.13	1	387	**<0.001**
Feedback	0.265	1	387	0.607
Condition × Feedback	3.185	1	387	0.075

Table 6.2: Results of the linear mixed-effect model for fixed factors Condition and Feedback and their interaction on **CI**

Term	**F Stat**	**DF1**	**DF2**	***p*-value**
Condition	110.32	1	361	**<0.001**
Feedback	17.86	1	361	**<0.001**
Condition × Feedback	8.26	1	361	**<0.01**

Table 6.3: Results of the linear mixed-effect model for fixed factors Condition and Feedback and their interaction on **RCI**

Term	**F Stat**	**DF1**	**DF2**	***p*-value**
Condition	58.144	1	361	**<0.001**
Feedback	5.925	1	361	**<0.05**
Condition × Feedback	3.62	1	361	0.057

Table 6.4: Results of the linear mixed-effect model for fixed factors Condition and Feedback and their interaction on **CCI**

Term	**F Stat**	**DF1**	**DF2**	***p*-value**
Condition	120.1	1	387	**<0.001**
Feedback	51.2	1	387	**<0.001**
Condition × Feedback	16.86	1	387	**<0.001**

For all the parameters considered, a significant difference was associated with the condition factor, as expected, given the physiological importance of the PLR maneuver which leads to an increase in blood volume compared to the supine position. The feedback factor was significant for CI, RCI and CCI parameters except for the diameter. The interaction between condition and feedback was significant for CI and CCI, with p values well below the significance threshold, 0.05. For the diameter and RCI, the p-values related to the interaction were 0.075 and 0.057 (Table 6.1, Table 6.3), respectively, which are close to the significance threshold. Given the exploratory nature of the analysis, it was nevertheless considered appropriate to proceed with post hoc tests also for the latter, in order to explore in detail the specific combinations of the experimental conditions: The post hoc test results for each parameter are shown in the following tables.

Table 6.5: Post hoc comparisons of Diameter across Feedback modalities and postural conditions.

Comparison		**F Stat**	***p*-value**
RT vs No Feedback	- PLR	0.265	0.606
RT vs No Feedback	- Supine	9.234	<0.01
Supine vs PLR	- No Feedback	195.13	<0.001
Supine vs PLR	- RT	131	<0.001

Table 6.6: Post hoc comparisons of CI across Feedback modalities and postural conditions.

Comparison		**F Stat**	***p*-value**
RT vs No Feedback	- PLR	17.859	<0.001
RT vs No Feedback	- Supine	68.733	<0.001
Supine vs PLR	- No Feedback	110.322	<0.001
Supine vs PLR	- RT	41.458	<0.001

Table 6.7: Post hoc comparisons of RCI across Feedback modalities and postural conditions.

Comparison		**F Stat**	***p*-value**
RT vs No Feedback	- PLR	5.92	<0.05
RT vs No Feedback	- Supine	26.28	<0.001
Supine vs PLR	- No Feedback	58.14	<0.001
Supine vs PLR	- RT	24.33	<0.001

Table 6.8: Post hoc comparisons of CCI across Feedback modalities and postural conditions.

Comparison		**F Stat**	***p*-value**
RT vs No Feedback	- PLR	51.2	<0.001
RT vs No Feedback	- Supine	168.04	<0.001
Supine vs PLR	- No Feedback	120.1	<0.001
Supine vs PLR	- RT	26.53	<0.001

All post hoc tests showed significant differences between the compared conditions, except for the comparison between Feedback RealTime and NOF in PLR for the diameter parameter (Table 6.5), as already seen in the global analysis. Unexpectedly, however, the same comparison, but in supine position showed a significant difference in diameter. This result suggests that the presence of feedback could bring a benefit in diameter measurement only in the supine condition. A possible explanation can be given by the fact that in PLR position, the vein is dilated and the venous walls more echogenic. This could result in more accurate measurements, reducing the scope for improvement through feedback. Overall, the analysis provided useful indications and motivated the deepening of subsequent analyses, focusing on the variability of the measures with the objective of comparing the acquisitions obtained in the presence and absence of feedback.

6.2 Intra and Inter-Operator Variability

6.2.1 Comparison between Real-Time and Post-processing with VIPER (NOF) Modes

To evaluate the impact of using the VIPER software in Real-time mode versus post-processing mode without real-time feedback (NOF), a thorough analysis was carried out by evaluating the variability indices: coefficient of variation (CoV), Intraclass Correlation Coefficient (ICC), standard deviation (SD) and 95% confidence intervals. The comparison was performed separately for the supine and PLR positions to analyze how the patient's position can affect the reliability of the measurement and to determine whether the software maintains its effectiveness under physiologically variable conditions. The analysis was conducted to evaluate both intra-operator variability (i.e., between two acquisitions by the same operator) and inter-operator variability (i.e.between the two different operators).

- **Mean diameter**: In the supine position, boxplots for the Real-Time mode appear more compact than those for the NOF mode with narrower interquartile ranges and more symmetrical distributions, suggesting lower measurement dispersion than observed in the post-processing mode (NOF) Figure 6.2. Although the medians are similar in both methods. ICC values and standard deviations (Table 6.9) also confirm this trend, showing better repeatability in Real-Time mode, for intra- and inter-operator level. For operator OB in particular, ICC is lower in the NOF mode, while for OA it remains comparable. Also Inter-operator variability, there are slightly more favorable values for real-time, with lower medians. In the PLR position, the real-time method shows lower CoV values compared to the post-processing mode, with more compact boxplots and the medians are confirmed lower in real-time for both operators. ICCs remain high even in PLR and are generally greater in Real-Time, with exception of inter-variability6.11.

- **Caval Index (CI)**: In the supine position, boxplots for the CI show greater variability than the diameter, in both modes (as can be expected, since it is an index derived from two measurements, its relative uncertainty is generally greater than that of a single diameter measurement [17]). However, it remains broader in the case of NOF. The Real-Time mode shows boxplots with a more compact distribution, with reduced dispersion. CoV medians are lower in Real-Time for both OA and OB, indicating greater intra-operator stability Figure 6.3. This result is supported by ICC values, which, although they are lower than the diameter, are still higher for the Real Time case for OA and for the inter-operator, and comparable for OB 6.10. In PLR position, boxplots remain more compact in Real-Time, with lower medians than the NOF, ICC

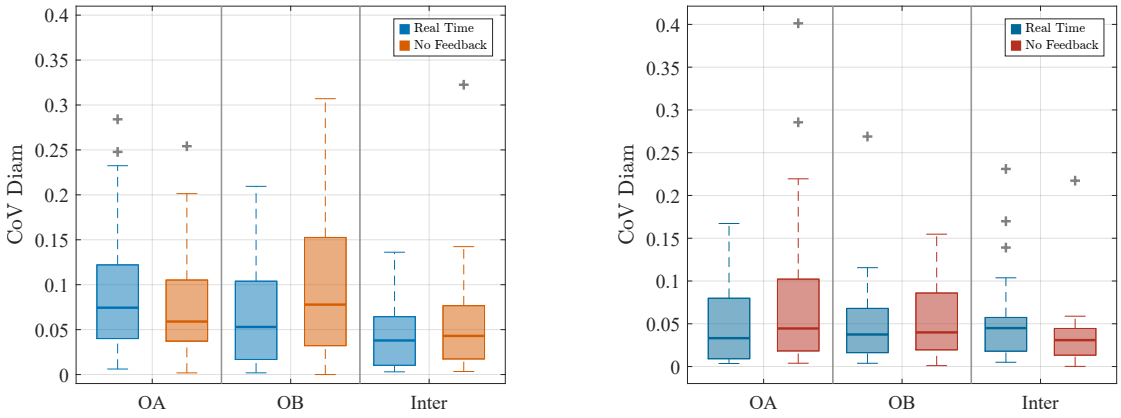
values are actually higher NOF.

- **Respiratory Caval Index (RCI)** : The boxplots have a higher dispersion both in the supine position and in PLR, compared to the diameter and the CI. In the supine position, real time mode results in lower variability for OB and for the inter variability Figure 6.4. The ICC's for RCI are lower than those for diameter and CI but remain higher in the RealTime case for OA and for the inter-operator 6.10. In the PLR position, on the other hand, all the boxplots are more compact in the Real time case, with slightly lower medians compared to the NOF. However, ICC values for OA and the inter-operator comparison are lower in Real-Time, while OB's ICC is higher in this mode.

- **Cardiac Caval Index (CCI)**: In supine position CCI boxplots show a low variability and are characterized by a higher dispersion in NOF mode. The Real-Time mode highlights tighter boxplots, lower medians, and more regular distributions, indicating greater repeatability, Figure 6.5. ICCs for Real-Time are higher in the case of the inter-operator, but lower for the OA. ICCs are instead comparable for the OB. In the PLR position, on the other hand, the variability increases slightly for both methods, but Real-Time again proves to be more reliable, with more compact boxplots and consistently lower medians. The ICC remain at high values and are slightly higher for OB operator, while are lower for OA and the inter-operator.

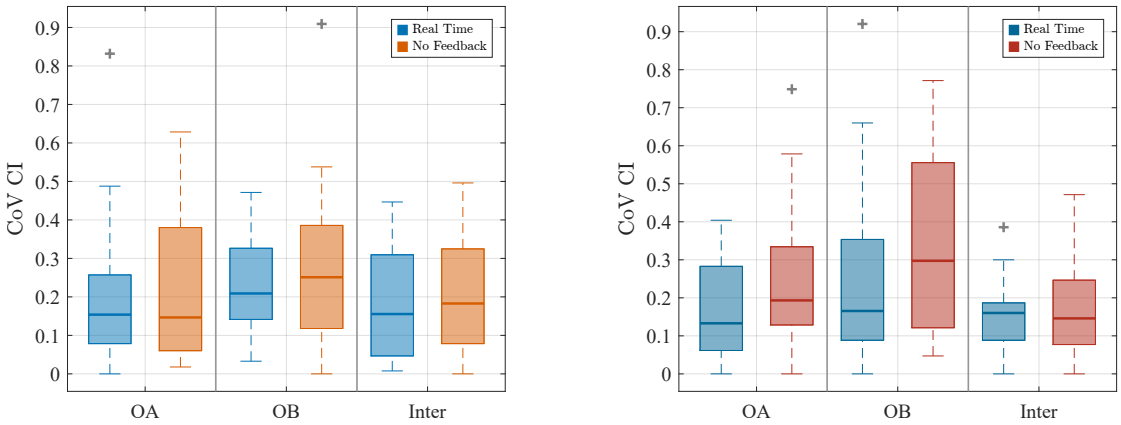
Overall, the analysis showed that the use of VIPER software in RealTime mode generally improved repeatability and consistency of measurements at both intra-operator and inter-operator. Most parameters show more compact boxplots, a lower dispersion in RealTime mode, with lower medians, which indicate less variability. Similarly, the ICC are often larger in Real time. However, there are cases where this is not observed, and the differences are less marked than expected.
One possible explanation is that, during the experimental phase, operators have gradually become familiar with the software and learned which features of the videos produce the best results. As a result, even in non-feedback acquisitions, operators were able to select favorable shots and sections more consciously, improving the quality of recordings already at the time of acquisition. To isolate the effectiveness of real-time feedback, it would be advisable to repeat the study with experts in ultrasound making measurements before knowing how the software works, to reduce the influence of learning during the experiment. In addition, it is necessary to consider that in our experimental setup, RealTime mode required the operator to manage the ultrasound probe and the VIPER software interface on the computer (e.g., to manually initiate the tracking). This can introduce small involuntary movements that may reduce the image quality and consequently limit the effectiveness of real-time feedback. This is a limitation related to the

experimental setup adopted, which could be overcome by integrating the software directly into the ultrasound interface, improving the overall accuracy of the process.



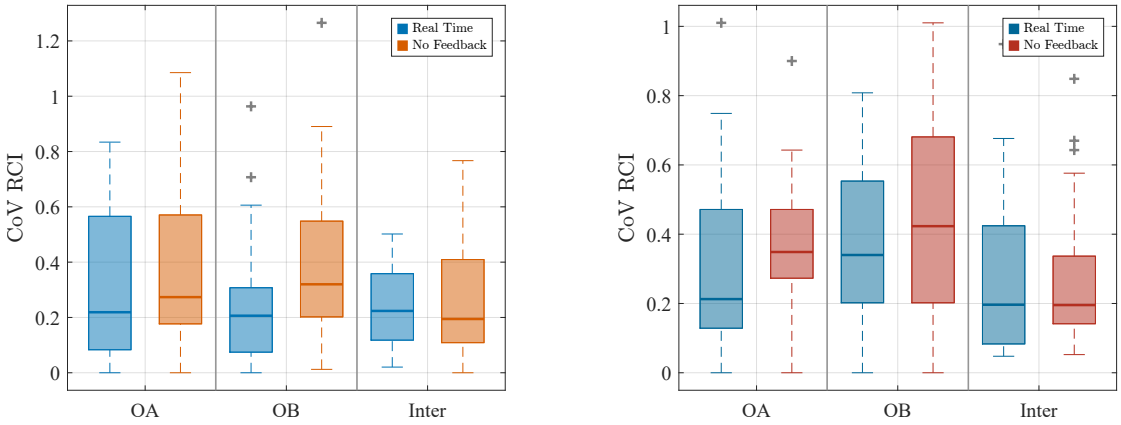
(a) SUPINE **(b)** PLR

Figure 6.2: Boxplot of the Coefficient of Variation (CoV) for diameter measurements in (a) supine and (b) PLR positions, comparing the Real-Time and the NOF (no-feedback) method for each operator and for the inter-operator comparison.



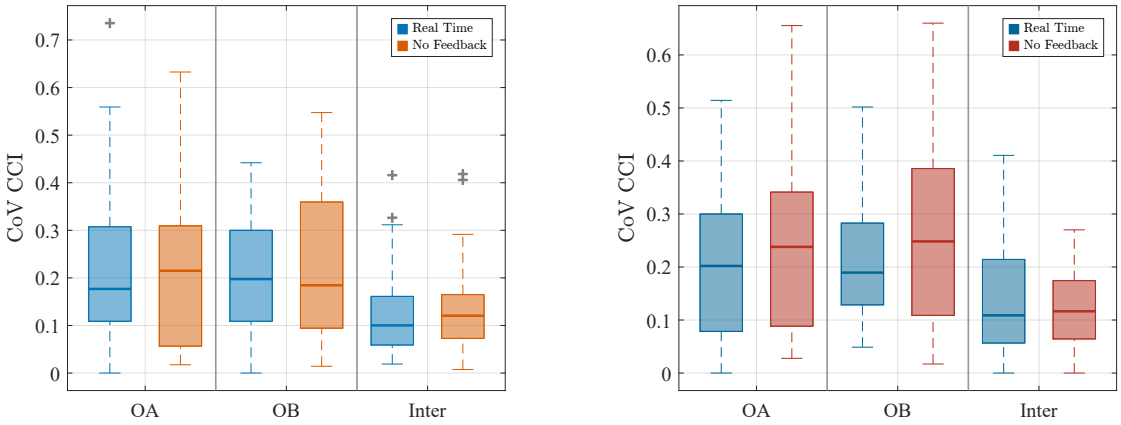
(a) SUPINE **(b)** PLR

Figure 6.3: Boxplot of the Coefficient of Variation (CoV) for CI measurements in (a) supine and (b) PLR positions, comparing the Real-Time and the NOF (no-feedback) method for each operator and for the inter-operator comparison.



(a) SUPINE **(b)** PLR

Figure 6.4: Boxplot of the Coefficient of Variation (CoV) for RCI measurements in (a) supine and (b) PLR positions, comparing the Real-Time and the NOF (no-feedback) method for each operator and for the inter-operator comparison.



(a) SUPINE **(b)** PLR

Figure 6.5: Boxplot of the Coefficient of Variation (CoV) for CCI measurements in (a) supine and (b) PLR positions, comparing the Real-Time and the NOF (no-feedback) method for each operator and for the inter-operator comparison.

Table 6.9: Intraclass Correlation Coefficient (ICC), standard deviation (STD) and 95% confidence intervals (CI) for each parameter and operator in Real-Time mode, in SUPINE position.

Parameter	**ICC**	**STD**	**95% CI**
Diam_OA	0.806	1.504	[0.403 – 0.926]
Diam_OB	0.693	2.109	[0.424 – 0.849]
CI_OA	0.660	0.057	[0.378 – 0.831]
CI_OB	0.400	0.072	[0.012 – 0.680]
RCI_OA	0.488	0.065	[0.139 – 0.732]
RCI_OB	0.277	0.072	[–0.097 – 0.590]
CCI_OA	0.486	0.059	[0.121 – 0.733]
CCI_OB	0.605	0.038	[0.118 – 0.829]
Diam_INTER	**0.926**	**1.0439**	**[0.844 – 0.966]**
CI_INTER	**0.676**	**0.0481**	**[0.396 – 0.841]**
RCI_INTER	**0.745**	**0.0375**	**[0.509 – 0.877]**
CCI_INTER	**0.838**	**0.0276**	**[0.671 – 0.924]**

Table 6.10: Intraclass Correlation Coefficient (ICC), standard deviation (STD) and 95% confidence intervals (CI) for each parameter and operator in NOF mode, in SUPINE position.

Parameter	**ICC**	**STD**	**95% CI**
Diam_OA	0.810	1.573	[0.625 – 0.910]
Diam_OB	0.666	1.881	[0.381 – 0.835]
CI_OA	0.558	0.105	[0.223 – 0.775]
CI_OB	0.405	0.098	[0.052 – 0.675]
RCI_OA	0.417	0.095	[0.035 – 0.690]
RCI_OB	0.534	0.078	[0.205 – 0.758]
CCI_OA	0.697	0.098	[0.427 – 0.852]
CCI_OB	0.607	0.082	[0.304 – 0.801]
Diam_INTER	**0.871**	**1.154**	**[0.735 – 0.940]**
CI_INTER	**0.557**	**0.080**	**[0.233 – 0.772]**
RCI_INTER	**0.465**	**0.075**	**[0.108 – 0.717]**
CCI_INTER	**0.804**	**0.061**	**[0.606 – 0.907]**

Table 6.11: Intraclass Correlation Coefficient (ICC), standard deviation (STD) and 95% confidence intervals (CI) for each parameter and operator in Real-Time mode, including inter-operator comparisons, in PLR position.

Parameter	**ICC**	**STD**	**95% CI**
Diam_OA	0.896	1.160	[0.782 – 0.952]
Diam_OB	0.868	1.340	[0.730 – 0.938]
CI_OA	0.724	0.036	[0.478 – 0.865]
CI_OB	0.283	0.074	[–0.122 – 0.603]
RCI_OA	0.680	0.031	[0.406 – 0.842]
RCI_OB	0.508	0.052	[0.163 – 0.743]
CCI_OA	0.640	0.037	[0.334 – 0.822]
CCI_OB	0.770	0.030	[0.556 – 0.890]
Diam_INTER	**0.869**	**1.290**	**[0.728 – 0.939]**
CI_INTER	**0.786**	**0.031**	**[0.579 – 0.898]**
RCI_INTER	**0.418**	**0.044**	**[0.052 – 0.688]**
CCI_INTER	**0.854**	**0.22**	**[0.705 – 0.932]**

Table 6.12: Intraclass Correlation Coefficient (ICC), standard deviation (STD) and 95% confidence intervals (CI) for each parameter and operator in the NOF (no-feedback) post-processing condition, including inter-operator comparisons, in PLR position.

Parameter	**ICC**	**STD**	**95% CI**
Diam_OA	0.731	1.876	[0.465 – 0.872]
Diam_OB	0.863	1.208	[0.720 – 0.936]
CI_OA	0.737	0.058	[0.490 – 0.873]
CI_OB	0.415	0.098	[0.031 – 0.690]
RCI_OA	0.755	0.038	[0.457 – 0.890]
RCI_OB	0.269	0.088	[–0.138 – 0.593]
CCI_OA	0.744	0.067	[0.511 – 0.876]
CCI_OB	0.751	0.058	[0.517 – 0.880]
Diam_INTER	**0.932**	**0.891**	**[0.855 – 0.969]**
CI_INTER	**0.803**	**0.049**	**[0.605 – 0.906]**
RCI_INTER	**0.600**	**0.051**	**[0.281– 0.799]**
CCI_INTER	**0.914**	**0.003**	**[0.818– 0.960]**

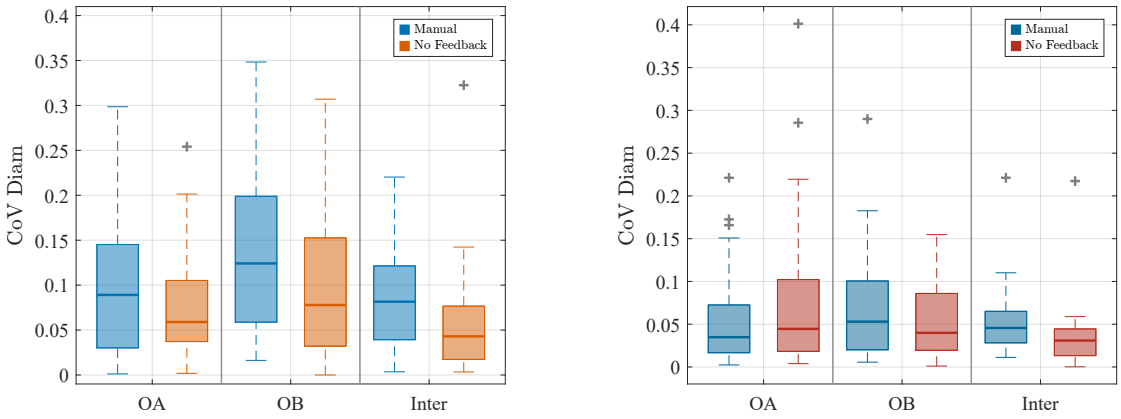
6.2.2 Comparison between Post-processing with VIPER (NOF) mode and Manual Method

This analysis was proposed to verify whether the use of the VIPER software without the visual support provided by real-time feedback, allows for more repeatable and consistent measurements compared to the traditional manual approach. The comparison was carried out on average diameter and CI and was conducted for supine and PLR positions.

- **Mean diameter**: in the supine position, the results show clear differences between the two methods. In both operators, the boxplots of the manual method show a higher dispersion and higher medians in the measurements. In contrast, the measurements obtained with VIPER in post-processing (NOF) are more compact for both inter and intra operator, characterized by lower medians. In particular, inter-operator variability is significantly reduced when using the software, as evidenced by narrower boxplots and fewer outliers. In the manual method boxplots, whiskers generally appear longer, indicating a greater dispersion of extreme values than in the automatic method, while in the NOF method, whiskers appear more contained and symmetrical. These results are confirmed by the ICC values, which result lower in the manual method for each operator and for the inter-operator comparison, as well as in the standard deviations, reduced in the NOF condition. For the PLR position results were more heterogeneous. For the OA operator, the manual method's CoV boxlots are more compact and with shorter whiskers than the VIPER method, suggesting better repeatability with the traditional approach in this case. In contrast, for the OB operator and in the inter-operator comparison, the manual boxplots are broader and characterized by extended whiskers, whereas the NOF boxplots are narrower and symmetrical, with whiskers contained, indicating superior repeatability with the software.

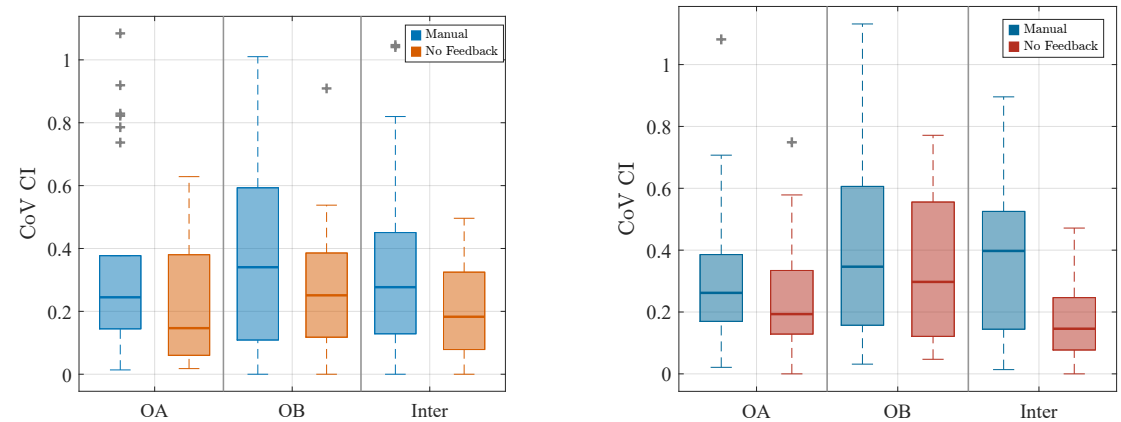
- **Caval Index (CI)** : there are significant differences between manual and software-based measurements. In the supine position, the boxplots obtained with the manual method are characterized by a higher dispersion, wider interquartile intervals, a higher number of outliers, and extended whiskers. In contrast, the NOF boxplots are more compact, characterized by lower medians and reduced whiskers, and fewer outliers. This suggests that, even for CI, automatic post-processing with VIPER helps to standardize the measurements and reducing intra- and inter-operator variability. The quantitative results the graphical analysis, showing lower ICC in the manual method compared to the software method. Although ICCs for CI were obtained compared to those of diameter, the trend is still positive and consistent with the hypothesis that the software use reduces operator-dependent variability. Standard deviations

also tend to be lower with the software. In the PLR position, CI boxplots are characterized by lower medians in the case with the software compared to the manual approach, with more narrow boxplots, except for the OA operator. ICC values for OB are similar between the two methods, while for OA, they are slightly lower when using the software.



(a) SUPINE **(b)** PLR

Figure 6.6: Boxplot of the Coefficient of Variation (CoV) for diameter measurements in (a) supine and (b) PLR positions, comparing the manual and the NOF (no-feedback) method , for each operator and for the inter-operator comparison.



(a) SUPINE **(b)** PLR

Figure 6.7: Boxplot of the Coefficient of Variation (CoV) for CI measurements in (a) supine and (b) PLR positions, comparing the manual and the NOF (no-feedback) method , for each operator and for the inter-operator comparison.

Table 6.13: Intraclass Correlation Coefficient (ICC), standard deviation (STD) and 95% confidence intervals (CI) for the diameter and CI parameters in the Manual method, including inter-operator comparisons, in SUPINE position.

Parameter	**ICC**	**STD**	**95% CI**
Diam_OA	0.682	2.015	[0.403 – 0.844]
Diam_OB	0.512	2.512	[0.170 – 0.746]
CI_OA	0.486	0.120	[0.136 – 0.730]
CI_OB	0.263	0.115	[–0.129 – 0.585]
inter_diam	**0.752**	**3.073**	**[0.521 – 0.880]**
inter_CI	**0.299**	**0.104**	**[–0.046 – 0.596]**

Table 6.14: Intraclass Correlation Coefficient (ICC), standard deviation (STD) and 95% confidence intervals (CI) for the diameter and CI parameters in the Manual method, including inter-operator comparisons, in PLR position.

Parameter	**ICC**	**STD**	**95% CI**
Diam_OA	0.824	1.516	[0.650 – 0.917]
Diam_OB	0.824	1.697	[0.649 – 0.917]
CI_OA	0.804	0.053	[0.612 – 0.907]
CI_OB	0.412	0.065	[0.028 – 0.687]
inter_diam	**0.884**	**3.557**	**[0.760 – 0.946]**
inter_CI	**0.528**	**0.083**	**[0.195 – 0.754]**

To assess the efficiency of the VIPER software, the time required to perform manual measurement was evaluated and compared with those obtained by VIPER post-processing. For manual method, the time was recorded from the moment immediately after the image calibration, until the moment of saving the measurement, excluding the ultrasound video acquisition, which had been performed previously. The average time for manual measurements was 33.55 seconds. In the case of automatic analysis with VIPER, the results are available directly at the end of the acquisition, without the need for any additional manual intervention. The entire process, which includes acquisition, tracking and saving, was estimated to take about 15 seconds allowing for the capture of at least three respiratory cycles. The comparison between the two methods, therefore, shows an average saving of 18.55 s (a 55.3% reduction). Beyond time efficiency, the real-time use of VIPER enables immediate access to data, thus simplifying the clinical process, eliminating post-acquisition processing analysis. This is particularly valuable in clinical contexts where a rapid response is essential, such as in the volemic state or the care of critically ill patients.
In conclusion, both for the CI and for the average diameter, the use of software in post processing has helped to obtain more stable and reliable measurements. In the PLR condition, however, some unexpected results emerged- such as with lower ICC in the case of software. This could be due to the fact that in position PLR, the IVC is physiologically more distended and clearly visible making the acquisition of quality ultrasound videos easier even in manual mode. Therefore, the intrinsic quality of the image in PLR is already optimal, allowing to obtain high repeatability even without using the software.

6.3 Cohen's d effect size

The size of Cohen's effect (Cohen's d) was calculated from the estimated coefficients of the linear model, considering only the fixed effect of the Condition (Supine vs PLR) in order to quantify the sensitivity of each approach to hemodynamic variation induced by posture change. The results show that all measurement modes enable the significant distinction between the Supine condition and the one in PLR, with consistent systematic variations among the analyzed parameters.

Table 6.15: Estimated difference between SUPINE and PLR conditions and Cohen's *d* for each method and parameter.

Method	Parameter	Estimated Difference	Cohen's *d*
No Feedback	Diam	-3.787 mm	-2.898
	CI	0.1314	1.783
	RCI	0.079	1.386
	CCI	0.128	1.940
Real-Time	Diam	-3.103 mm	-2.315
	CI	0.081	1.750
	RCI	0.051	1.174
	CCI	0.060	1.807
Manual	Diam	-4.058 mm	-2.498
	CI	0.108	1.401

Note. Differences are calculated as SUPINE − PLR. Negative values indicate larger measurements in the PLR condition.

- **IVC diameter**: the model was set with the Supine - PLR comparison , therefore negative estimates indicate an increase in diameter in PLR, consistent with the physiology of venous return, as can be seen in detail in the table 6.15. Although the manual measurement shows the largest estimated difference, the no-feedback mode has the highest effect size. This indicates that, relative to the variability of the data, the no-feedback mode is slightly more sensitive in detecting the difference. At the same time, the manual measurement shows a larger average difference but with a higher variability.

- **Collapsibility indices (CI, RCI, CCI):** The effect of the condition is always positive, confirming that these indices decrease in PLR (i.e., less collapsibility and greater venous filling). Sensitivity to postural variation varies between modes: CI presents maximum discriminability in the mode without feedback (d = 1,783), followed by that with feedback (d = 1,75) and lastly by manual

measurement (d = 1,401). RCI and CCI (calculated only in the mode with and without feedback) have slightly better values for no feedback mode.

Overall, all approaches show good discriminative capacity between supine and PLR, with wide-ranging effects on all parameters. Differences between modes suggest that:

- Manual measurement is the mode with the highest absolute diameter variation (the highest estimated difference) and also it exhibits higher variability than the no-feedback mode.
- The no-feedback mode shows the highest standardized effect size for diameter and collapse indices, suggesting greater internal consistency and sensitivity to postural variations.
- The real-time feedback mode offers good and balanced results on all parameters, with a slight reduction in sensitivity compared to the no-feedback mode.

However, this result was not expected: more discriminative effectiveness of the real-time mode was to be foreseen, simply because of the availability of visual support in real time. One possible explanation has to do with a biasing of the operators, who, having gained their initial knowledge of the software, could have mastered the technique even without feedback, reducing the difference between the two conditions. This learning effect would have cancelled out the theoretical advantage achieved through real-time feedback, making no feedback mode unexpectedly superior at discriminating PLR-induced variation.

6.4 Agreement Between Measurement: Bland–Altman Analysis

To evaluate the agreement between measurement methods, a Bland-Altman analysis was performed comparing the NOF method with the Real time method, separately for supine condition and PLR for both the diameter and CI parameters.

- **Comparison between Real-Time and NOF methods for the diameter**

  In the supine condition, the bias is 0.82 mm, contained within acceptable values. The limits of agreement (LoA) are relatively wide, ranging from -5.31 mm to +6.96 mm. Most of the points fall within these limits and are symmetrically distributed around the bias line, indicating a contained disagreement between the two methods. In the PLR condition, the bias is 0.14 mm, the same as in the Manual-NOF comparison. The LoA are comparable (-4.25 mm to +4.53 mm) to the manual comparison. It can also be observed that in the PLR condition most of the points are concentrated in the range between -2 mm and +2 mm. These changes in diameter are generally considered not to be physiologically relevant.

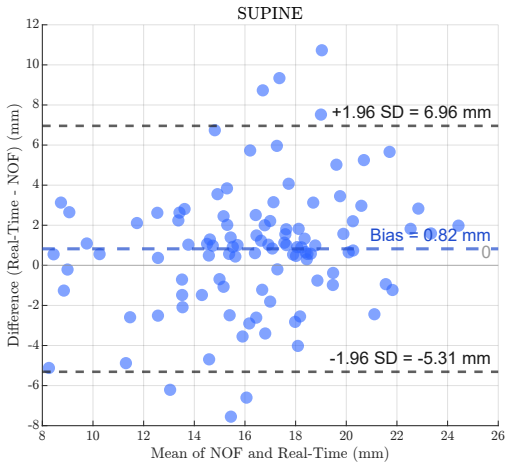


Figure 6.8: Diameter: Bland–Altman plot for NOF and Real Time modes in the SUPINE condition.

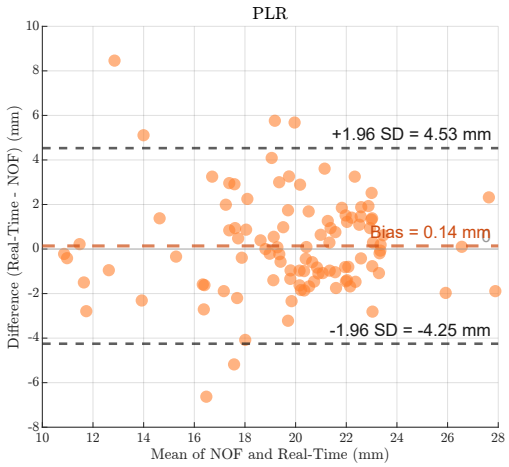


Figure 6.9: Diameter: Bland–Altman plot for NOF and Real Time modes in the PLR condition.

- **Comparison between NOF and MANUAL method for the diameter** In supine condition, the observed bias is 0.41 mm, slightly lower than the one obtained between NOF and Real, indicating a good agreement between the two methods. The LoA are between -3.89 mm and +4.71 mm, with a relatively low dispersion. Most of the points fall within these limits, suggesting reliable agreement between manual and NOF measurements. Even in the PLR condition, the bias is very low, equal to 0.14 mm, with LoA between -3.88 mm and +4.16 mm.

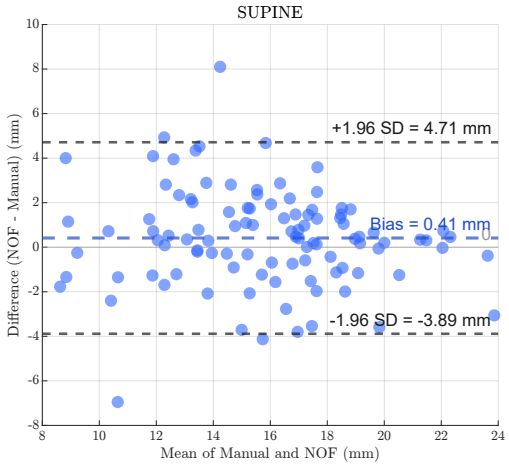


Figure 6.10: Diameter: Bland–Altman plot for NOF and MANUAL methods in the SUPINE condition.

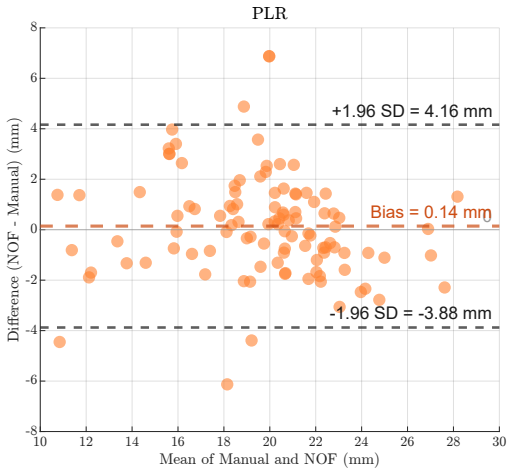


Figure 6.11: Diameter: Bland–Altman analysis for NOF and MANUAL methods in the PLR condition.

- **Comparison between Real-Time and NOF methods for the CI**

  In the supine condition, the bias is -0.1, which compared to The limits of agreement (LoA) is not so negligible. There are scattered points even beyond the agreement limits, although a more consistent cloud around zero is evident. The situation is better in the case of PLR with narrower agreement limits and almost all points within the range from 0.1 to -0.1, variations for which the value of CI has the same physiological significance.

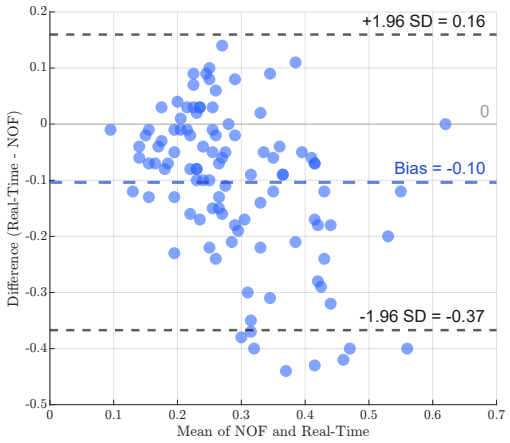


Figure 6.12: CI: Bland–Altman plot for NOF and Real Time modes in the SUPINE condition.

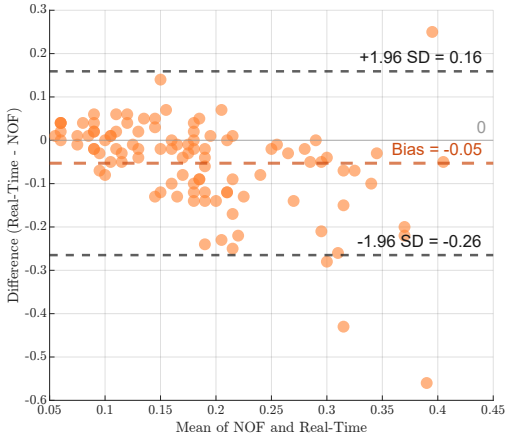


Figure 6.13: CI: Bland–Altman plot for NOF and Real Time modes in the PLR condition.

- **Comparison between Manual and NOF methods for the CI**

  In the comparison between Manual and NOF the biases are negligible, but especially for the supine condition there is a greater dispersion of the points.Also in this case the situation improves slightly for the condition PLR with most of the points within the range (0.1, -0.1).

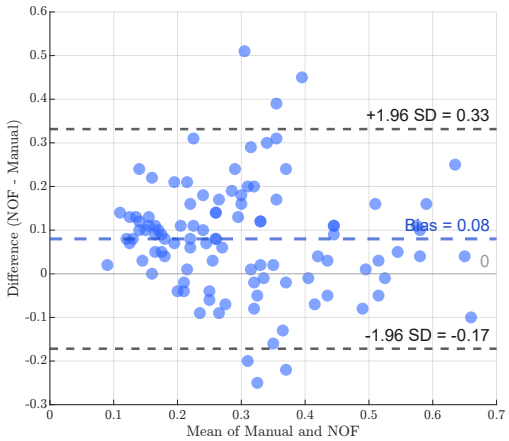


Figure 6.14: CI: Bland–Altman plot for NOF and MANUAL methods in the SUPINE condition.

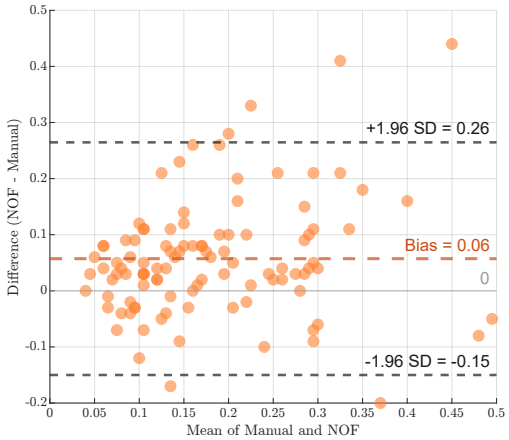


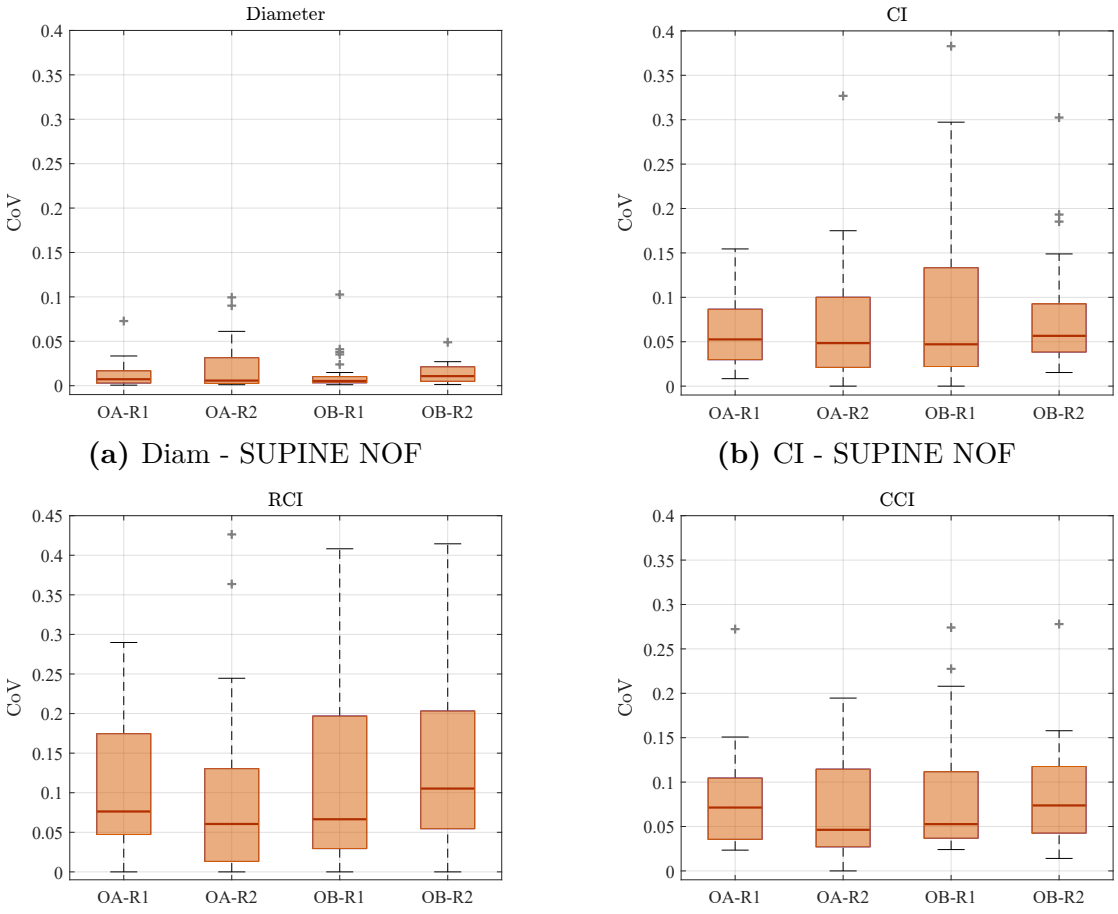
Figure 6.15: CI: Bland–Altman analysis for NOF and MANUAL methods in the PLR condition.

In conclusion, the Bland-Altman analysis confirmed good agreement between the methods considered, with low bias values and a distribution mostly within concordance limits, although the Real-Time method presents a slight variability more than the manual method for the diameter parameter. Probably this is due to the fact that while for the Manual-NOF the compared measurements were made on the same video, for the RealTime-NOF comparison, instead, the measurements were made on corresponding videos, but not exactly the same. In fact the real-time measurements were made on video where there was the real-time rendering of the vein, while those in NOF, from the post processing of videos obtained without the feedback in real-time. Nevertheless, for the CI parameter, the estimates are more robust in real time than the manual method. This highlights the variability and unreliability of manual measurements.

6.5 Reliability Assessment of Initial Segmentation Point Selection Across Operators

An analysis was conducted to verify the reliability of the software based on the initial point chosen for segmentation. Each ultrasound video was processed five times by each operator using the VIPER software in post-processing. The main objective was to assess whether each operator was consistent in selecting approximately the same starting point between across multiple trials, and to determine whether this selection was also similar between the two operators. This analysis was carried out under four different conditions: supine without Real -Time software feedback, supine with Real -Time software feedback, PLR without Real -Time software feedback, PLR with Real -Time software feedback. Coefficients of variation (CoV) and intraclass coefficients (ICC) were evaluated and on the whole, very positive results were observed in terms of reliability and repeatability of the software based on the choice of the initialization point.

1. **Supine - NOF**
   The variability is extremely low: both the graphical boxplot analysis 6.23 and the elevated ICC values 6.16 confirm excellent intra-operator repeatability and good inter-operator consistency. The measurements were tightly clustered and highly stable, with compact and symmetric boxplots for both the average diameter and derived indices (CI, RCI, CCI). ICC values exceeded 0.90 for most parameters. The average diameter shows greater stability, while the other indices are slightly more sensitive to variations between operators, although they remain at satisfactory levels.



(a) Diam - SUPINE NOF

(b) CI - SUPINE NOF

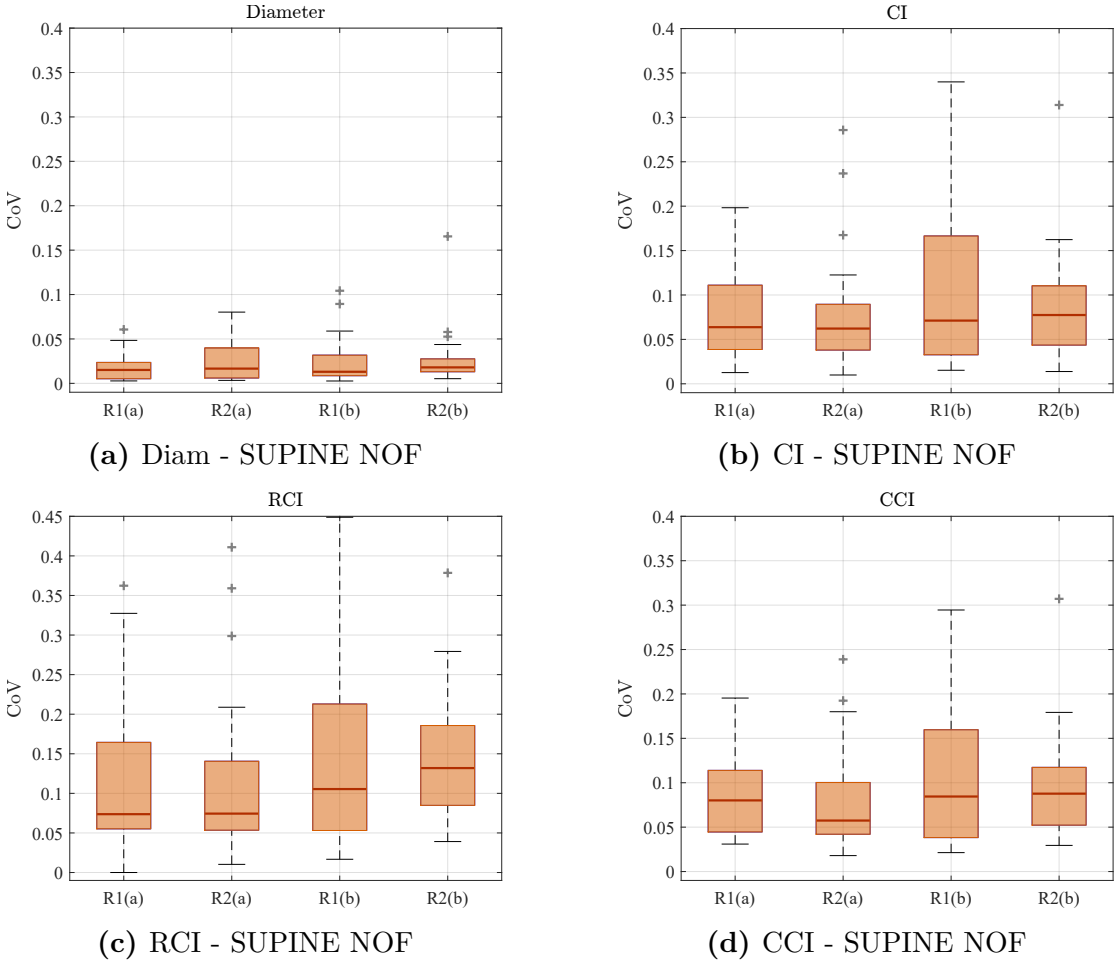
(c) RCI - SUPINE NOF

(d) CCI - SUPINE NOF

Figure 6.16: Boxplots of the Coefficient of Variation (CoV) computed from five post-processing repetitions (5pp) for operators OA and OB **supine NOF**. These plots represent the **INTRA-operator variability** for each parameter

Table 6.16: Intra-operator Intraclass Correlation Coefficient (ICC) computed on five post-processing repetitions (5pp) for SUPINE NOF.

Repetition	ICC_Diam	ICC_CI	ICC_RCI	ICC_CCI
OA_R1	0.996	0.976	0.933	0.962
OA_R2	0.979	0.966	0.955	0.975
OB_R1	0.989	0.884	0.878	0.924
OB_R2	0.992	0.935	0.912	0.946



(a) Diam - SUPINE NOF

(b) CI - SUPINE NOF

(c) RCI - SUPINE NOF

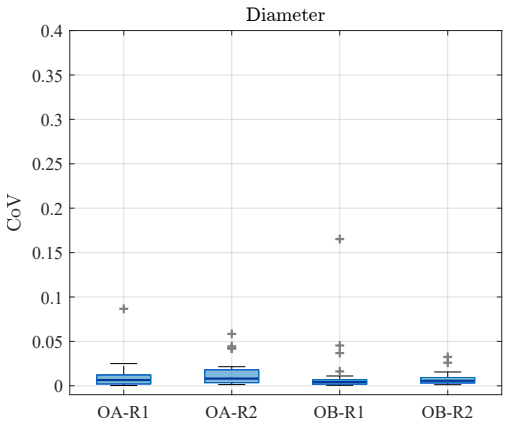
(d) CCI - SUPINE NOF

Figure 6.17: Boxplots of the Coefficient of Variation (CoV) computed from five post-processing repetitions (5pp) for operators OA and OB **supine NOF**. These plots represent the **INTER-operator variability** for each parameter

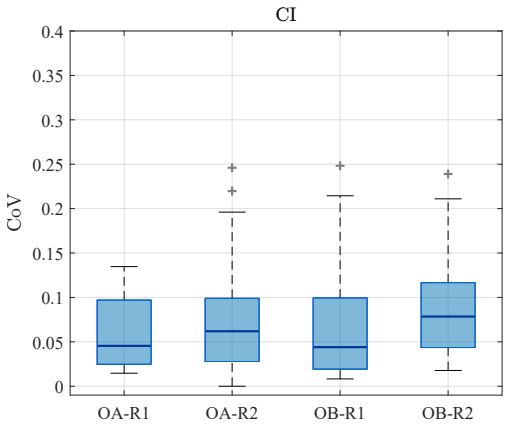
Table 6.17: Inter-operator Intraclass Correlation Coefficient (ICC) computed on five post-processing repetitions (5pp) for each parameter and repetition in the **supine NOF** condition.

Repetition	**ICC_Diam**	**ICC_CI**	**ICC_RCI**	**ICC_CCI**
OA_R1	0.977	0.919	0.835	0.881
OA_R2	0.934	0.898	0.883	0.909
OB_R1	0.950	0.615	0.631	0.729
OB_R2	0.954	0.810	0.756	0.853

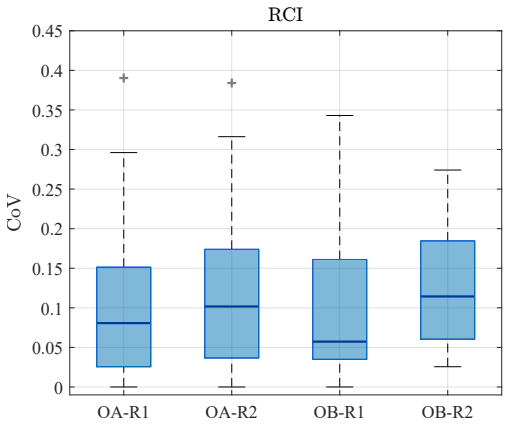
2. **Supine - RealTime**
   The results confirm excellent intra-operator repeatability and good inter-operator consistency. The boxplots are extremely compact and symmetrical for all parameters, indicating a low dispersion of measurements. ICC values for intra-operator variability are very high, with values in most cases slightly higher than the mode without feedback. The mean diameter is again confirmed as the most stable parameter, while the derived indices show a slightly higher sensitivity to differences between operators, still maintaining high absolute ICC values, indicative of strong reliability.



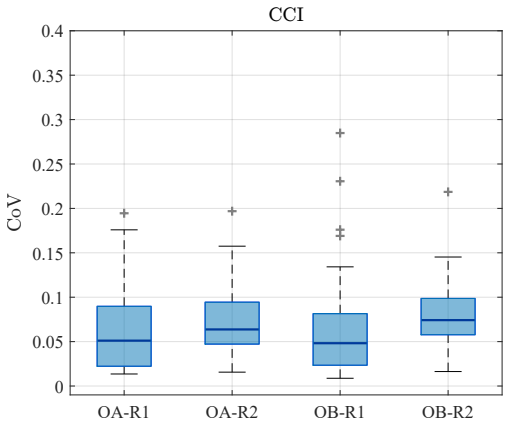
(a) Diam - SUPINE Real-Time



(b) CI - SUPINE Real-Time



(c) RCI - SUPINE Real-Time

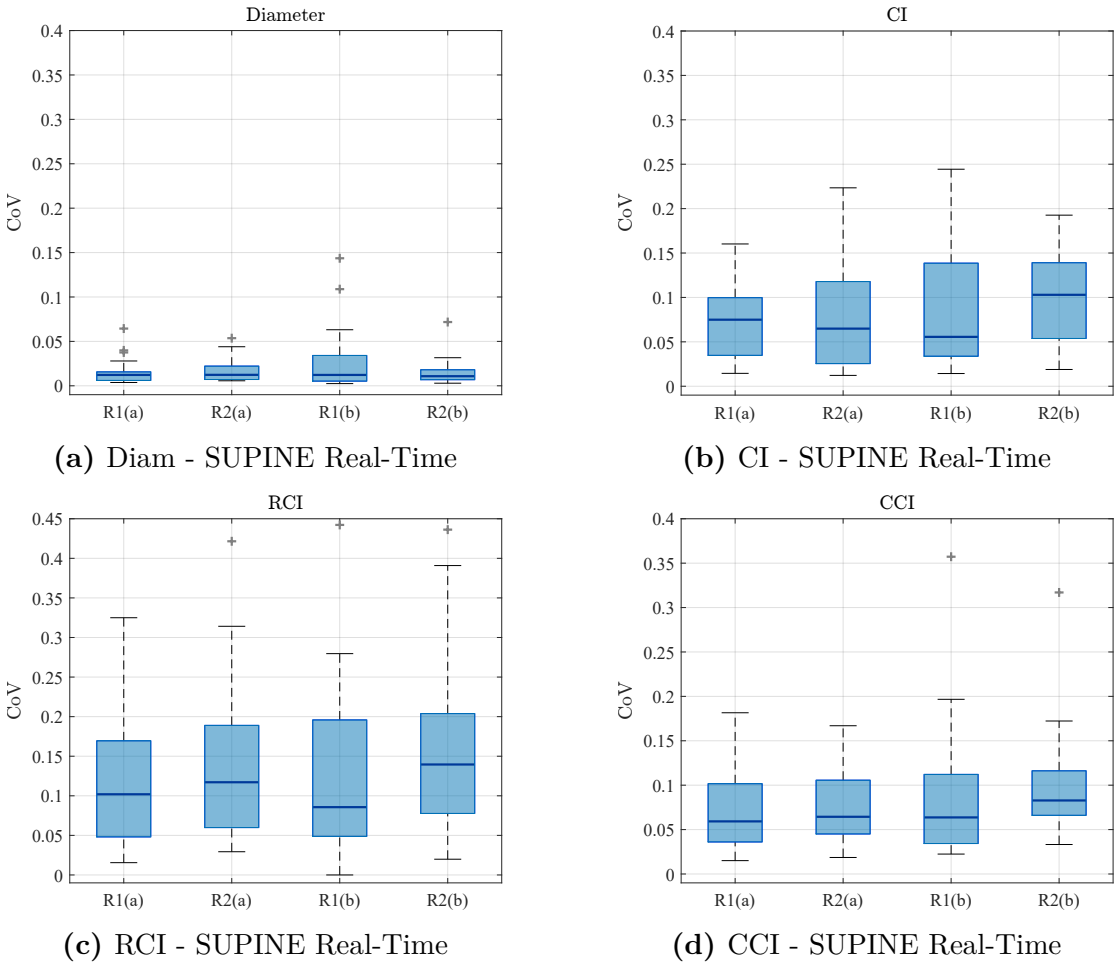


(d) CCI - SUPINE Real-Time

Figure 6.18: Boxplots of the Coefficient of Variation (CoV) computed from five post-processing repetitions (5pp) for operators OA and OB **supine VIPER**. These plots represent the **INTRA-operator variability** for each parameter

Table 6.18: Intra-operator Intraclass Correlation Coefficient (ICC) computed on five post-processing repetitions (5pp) for each parameter and repetition in the **SUPINO VIPER** condition.

Repetition	**ICC_Diam**	**ICC_CI**	**ICC_RCI**	**ICC_CCI**
OA_R1	0.993	0.965	0.969	0.970
OA_R2	0.991	0.952	0.941	0.951
OB_R1	0.991	0.936	0.924	0.929
OB_R2	0.997	0.950	0.965	0.973



(a) Diam - SUPINE Real-Time

(b) CI - SUPINE Real-Time

(c) RCI - SUPINE Real-Time

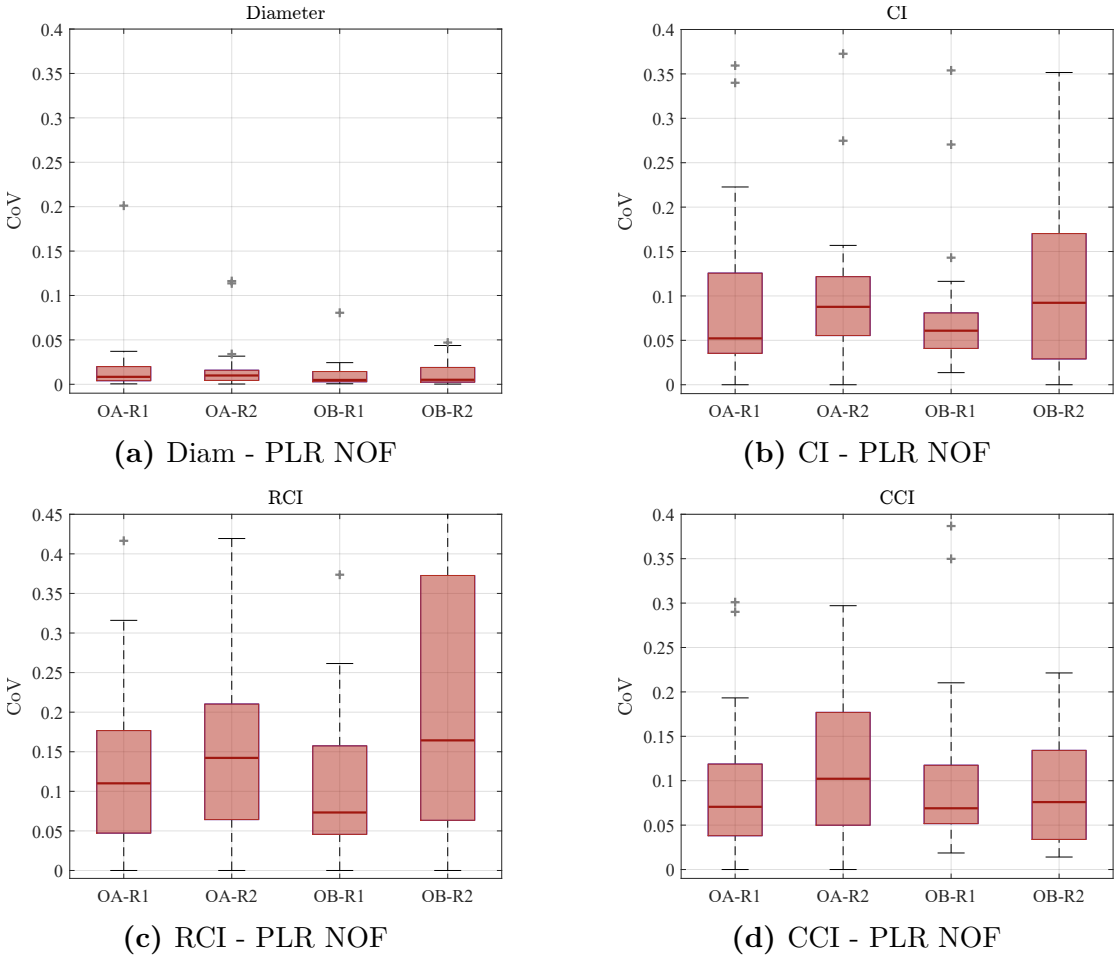
(d) CCI - SUPINE Real-Time

Figure 6.19: Boxplots of the Coefficient of Variation (CoV) computed from five post-processing repetitions (5pp) for operators OA and OB **supine VIPER**. These plots represent the **INTER-operator variability** for each parameter

Table 6.19: Inter-operator Intraclass Correlation Coefficient (ICC) computed on five post-processing repetitions (5pp) for each parameter and repetition in the **SUPINE VIPER** condition.

Repetition	**ICC_Diam**	**ICC_CI**	**ICC_RCI**	**ICC_CCI**
OA_R1	0.967	0.897	0.898	0.908
OA_R2	0.972	0.875	0.840	0.877
OB_R1	0.956	0.781	0.781	0.821
OB_R2	0.972	0.810	0.873	0.911

3. **PLR - NOF**
   The results show excellent intra-operator repeatability with high ICC values for all parameters, especially for diameter, which is confirmed as the most stable parameter. The associated boxplots show low dispersion and good symmetry. As for inter-operator variability, there is a slight decrease in consistency between operators, in particular for the RCI parameter, whose ICC values compared to the other indices. Nevertheless, overall performance remains satisfactory even between operators.



(a) Diam - PLR NOF

(b) CI - PLR NOF

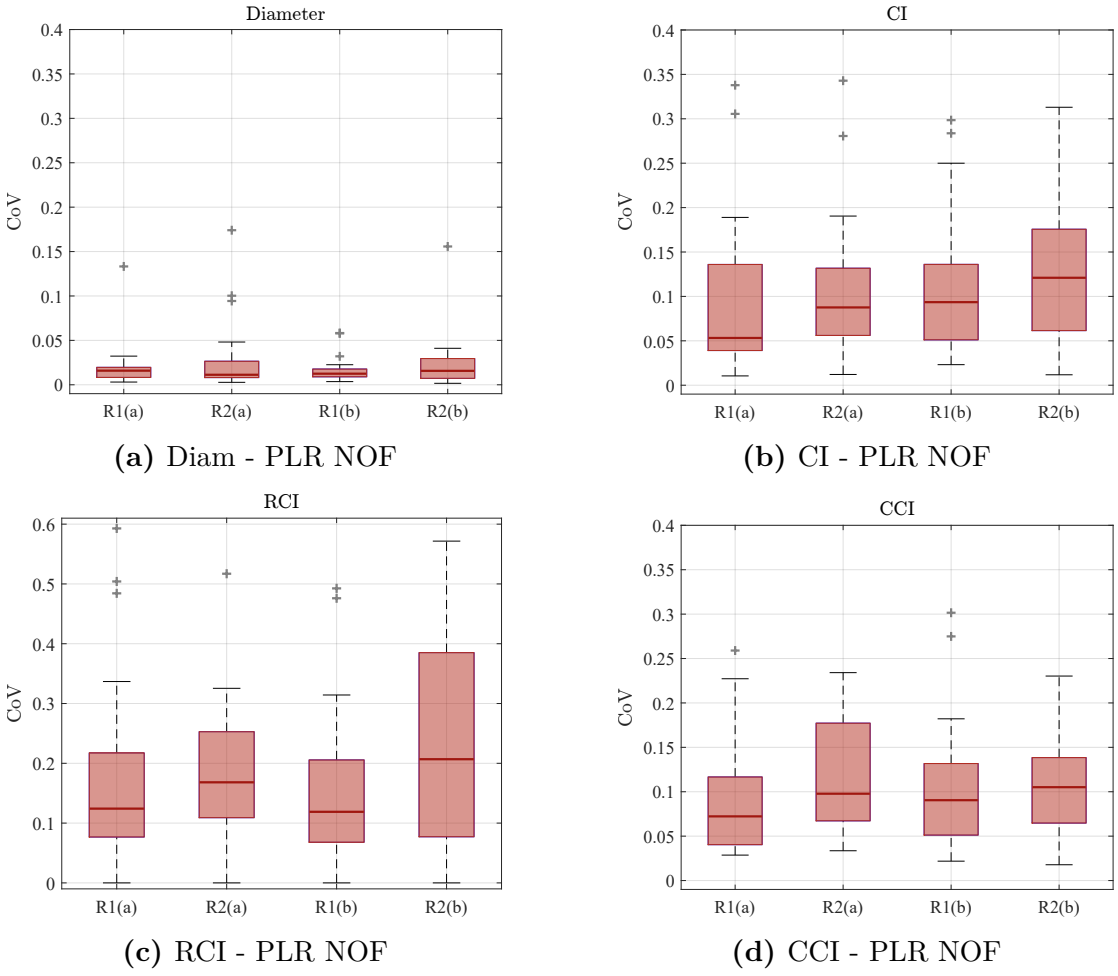
(c) RCI - PLR NOF

(d) CCI - PLR NOF

Figure 6.20: Boxplots of the Coefficient of Variation (CoV) computed from five post-processing repetitions (5pp) for operators OA and OB **PLR NOF**. These plots represent the **INTRA-operator variability** for each parameter

Table 6.20: Intra-operator Intraclass Correlation Coefficient (ICC) computed on five post-processing repetitions (5pp) for each parameter and repetition in the **PLR NOF** condition.

Repetition	ICC_Diam	ICC_CI	ICC_RCI	ICC_CCI
OA_R1	0.953	0.901	0.825	0.944
OA_R2	0.972	0.962	0.896	0.947
OB_R1	0.987	0.973	0.961	0.952
OB_R2	0.991	0.867	0.761	0.952



(a) Diam - PLR NOF

(b) CI - PLR NOF

(c) RCI - PLR NOF

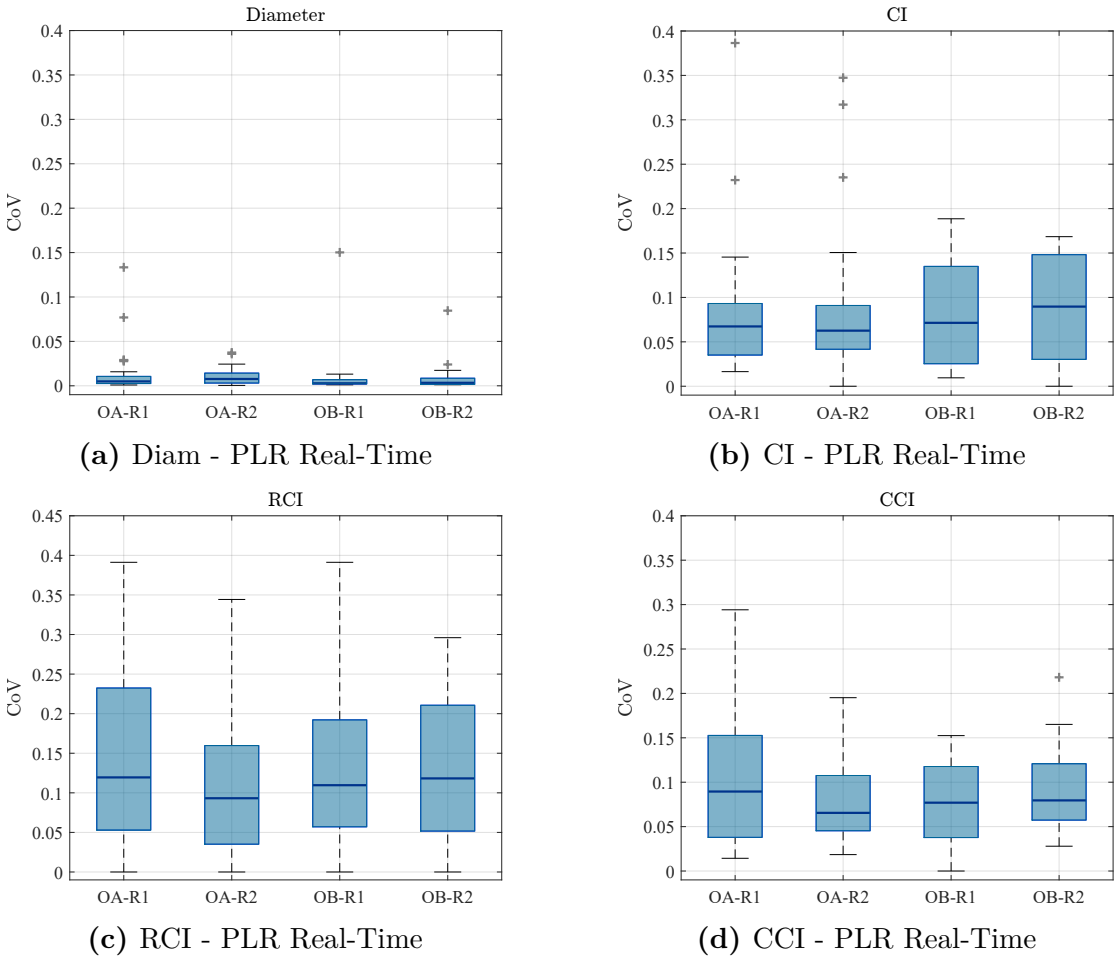
(d) CCI - PLR NOF

Figure 6.21: Boxplots of the Coefficient of Variation (CoV) computed from five post-processing repetitions (5pp) for operators OA and OB **PLR NOF**. These plots represent the **INTER-operator variability** for each parameter

Table 6.21: Inter-operator Intraclass Correlation Coefficient (ICC) computed on five post-processing repetitions (5pp) for each parameter and repetition in the **PLR NOF** condition.

Repetition	**ICC_Diam**	**ICC_CI**	**ICC_RCI**	**ICC_CCI**
OA_R1	0.871	0.736	0.479	0.864
OA_R2	0.898	0.891	0.696	0.838
OB_R1	0.950	0.900	0.842	0.870
OB_R2	0.927	0.659	0.373	0.861

4. **PLR - RealTime**
   Overall, the analysis confirmed good results in system reliability, especially at intra-operator level. The ICC values are high for all measurements, especially the diameter, and the boxplots show a narrow and regular distribution, indicating a limited dispersion of data within the same operator. At the inter-operator level, however, a higher degree of variability is observed, especially for the CI and RCI, which have lower ICC values. While still within acceptable limits, there is a slight decrease in consistency between operators compared to the other experimental conditions.



(a) Diam - PLR Real-Time

(b) CI - PLR Real-Time

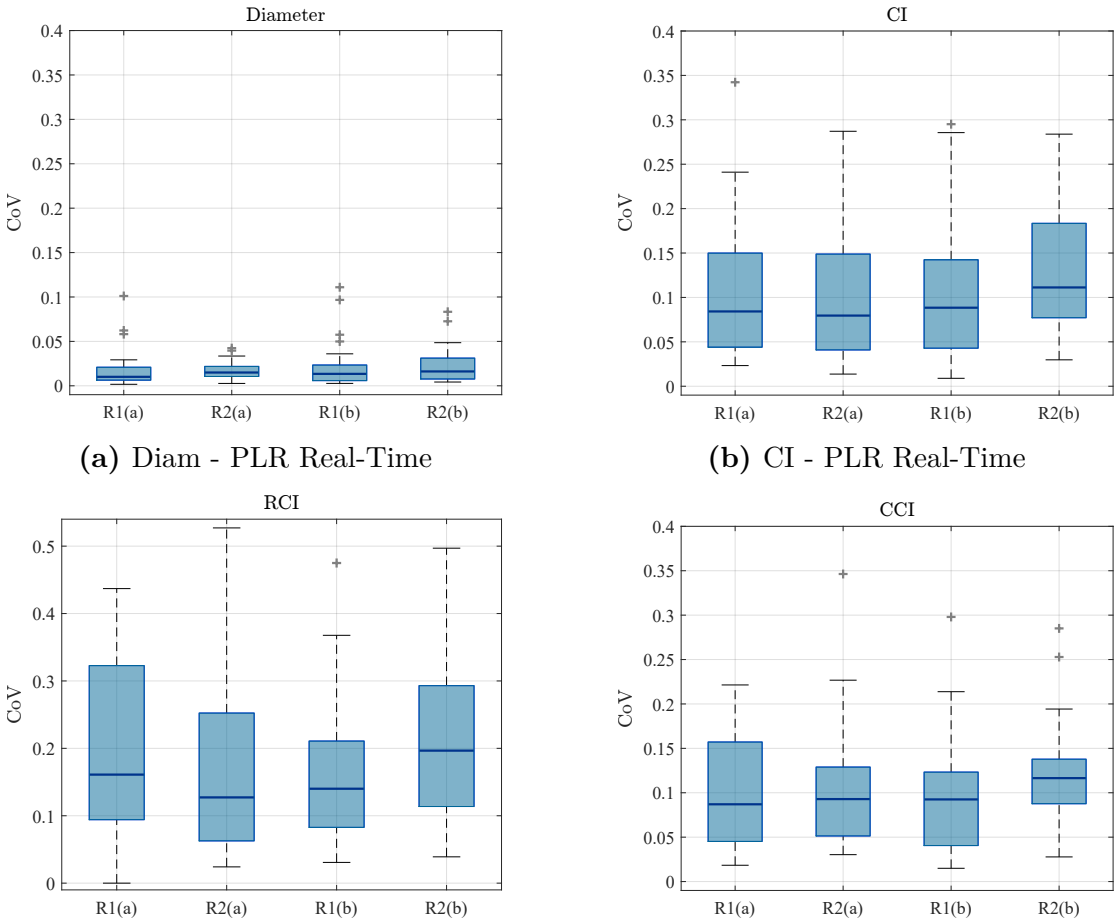
(c) RCI - PLR Real-Time

(d) CCI - PLR Real-Time

Figure 6.22: Boxplots of the Coefficient of Variation (CoV) computed from five post-processing repetitions (5pp) for operators OA and OB **PLR VIPER**. These plots represent the **INTRA-operator variability** for each parameter

Table 6.22: Intra-operator Intraclass Correlation Coefficient (ICC) computed on five post-processing repetitions (5pp) for each parameter and repetition in the **PLR Real-Time** condition.

Repetition	**ICC_Diam**	**ICC_CI**	**ICC_RCI**	**ICC_CCI**
OA_R1	0.988	0.849	0.663	0.919
OA_R2	0.994	0.901	0.766	0.963
OB_R1	0.968	0.969	0.968	0.982
OB_R2	0.995	0.898	0.781	0.964



(a) Diam - PLR Real-Time

(b) CI - PLR Real-Time

(c) RCI - PLR Real-Time

(d) CCI - PLR Real-Time

Figure 6.23: Boxplots of the Coefficient of Variation (CoV) computed from five post-processing repetitions (5pp) for operators OA and OB **PLR VIPER**. These plots represent the **INTER-operator variability** for each parameter

Table 6.23: Inter-operator Intraclass Correlation Coefficient (ICC) computed on five post-processing repetitions (5pp) for each parameter and repetition in the **PLR VIPER** condition.

Repetition	**ICC_Diam**	**ICC_CI**	**ICC_RCI**	**ICC_CCI**
OA_R1	0.94828	0.59614	0.33360	0.73459
OA_R2	0.95411	0.69132	0.42144	0.83523
OB_R1	0.84952	0.84059	0.85501	0.84022
OB_R2	0.95818	0.67677	0.32507	0.86229

Chapter 7

Conclusions and Future Developments

7.1 Conclusions

The objective of this study was to evaluate the reliability and repeatability of the ultrasound measurements of the IVC obtained using the VIPER software, comparing different operating modes (manual, post-processing without feedback, and real-time).
The statistical analysis, based on linear mixed effects models, showed a significant effect associated with the condition factor, given the physiological importance of the PLR maneuver, that increases blood volume compared to the supine position. The feedback factor was significant for all parameters, except for diameters. These results motivated the deepening of the subsequent analyses.
Intra- and inter-operator variability was evaluated by comparing three modes: real-time software use, post-processing without feedback (NOF) and the manual method. In particular, the Coefficient of Variation (CoV), the Intra Class Coefficient (ICC) and standard deviation were evaluated. The comparison between real-time and postprocessing modes showed that the use of real-time software did not lead to substantial improvements in measurement repeatability. The comparison between the manual method and the NOF mode showed that more stable and repeatable measurements were obtained with the use of the software.
The analysis of the standardized effect (Cohen's d) showed that all modes demonstrated a good discriminative capacity between the supine and PLR positions. The manual measurement showed the highest absolute variability in diameter measurements, whereas the no-feedback mode yielded the highest Cohen's d values for diameter and CI, indicating higher internal consistency and greater ability to discriminate postural variations. The real-time mode also offers good results,

although with a reduction in sensitivity compared to the no-feedback mode. In general, improvements in the repeatability of measurements were higher in the supine position compared to the PLR. This result can be understood considering some practical aspects of performing the ultrasound examination. In the PLR condition, in fact, the acquisition is generally simpler: the venous walls are more echogenic, the vein is more visible and, from a physiological point of view, moves less due to lower collapsibility indexes. These features help to increase image stability, simplifying data acquisition and reducing variability between measurements.
The expected results from this study were that the introduction of real-time feedback could contribute significantly to improving ultrasound quality and repeatability of parameter estimates. However, this effect was not evident in the study.
Lower quality video acquired without feedback might have been expected, in line with the results of previous studies [34], where unguided acquisitions often led to images that software could not process. In contrast, in this study, the success rate of video processing was 100%, even for acquisitions made without the support of real-time feedback. This result can be attributed to a learning bias explained later in the limits of the study. Although in some cases box-plots show greater compactness and less variability in the measurements obtained with feedback, these improvements seem modest. This suggests that the effectiveness of feedback, although potentially useful, was not evident in the specific context of this study, probably because the quality of the acquisitions was already high even in its absence.
A specific analysis was carried out to verify the reliability of the software in relation to the selection of the initial segmentation point, to assess whether each operator was consistent in choosing the starting point between the different repetitions and whether this choice was also similar between the two operators. The results showed excellent intra- and inter-operator repeatability in all the conditions.
Finally, another significant advantage of using VIPER is the significant reduction in time required to obtain measurements. While the manual method took an average of 33.55 seconds, VIPER was able to complete the entire process - from acquisition to saving - in approximately 15 seconds, with a 55.3% time reduction. Moreover, in the case of automatic analysis with VIPER, the results are available immediately at the end of the acquisition, without the need for any additional manual intervention. This efficiency is particularly important in the clinical field, especially in situations where a rapid response is required, such as evaluating volemic status or managing critical patients.
Other advantages that emerge from the use of the software are first, the possibility to visualize in real-time the venous diameter measurement. This resolves one of the main technical limitations of IVC longitudinal scans, namely the difficulty in correctly identifying the central and axial section of the vein. The instantaneous display of the diameter allows you to quickly correct the framing, making it easier to identify the central section.

Secondly, the software provides not only the mean diameter value and CI, but also two additional indices - RCI (Respiratory Collapsibility Index) and CCI (Cardiac Collapsibility Index) - which have been recognized in the literature as potentially useful from the clinical point of view [27]. These parameters allow a more complete evaluation of the patient's hemodynamic status.
Finally, the software provides graphical outputs as a trace of vessel diameter over time, marking also respiratory and cardiac oscillations. These dynamic plots provide the operator with an immediate intuitive sense of diameter trends, thereby enriching the interpretation of ultrasound data and supporting more considered and enlightened clinical analysis.

7.2 Limits of the study

Although the results demonstrate good performance of the VIPER software, the study highlighted some important limitations to be considered. First, during the experimental phase, operators have gradually become familiar with the software interface and with the features needed to obtain better video for analysis. This learning bias may have led operators to select better-quality videos more consciously and systematically, even in acquisitions made without real-time feedback (NOF), thus improving image quality regardless of mode. Another critical issue concerns the experimental setup: in real-time mode, the operator simultaneously manages both the ultrasound probe and the software interface on a second computer (e.g., to start and stop tracking). This double activity may have introduced involuntary movements, reducing the stability and quality of the acquisition. Conversely, in NOF mode, the operator could focus exclusively on video acquisition, without interacting with real-time tracking software, which might explain the better performance obtained in this mode. To isolate the effectiveness of real-time feedback, it would be advisable to repeat the study with experts in ultrasound and non-familiar with the software interface, to reduce the influence of learning during the experiment. In addition, direct integration of the software into the ultrasound interface, thus avoiding the need to interact with a second computer, could further improve the acquisitions quality and make the process more intuitive. Finally, the study was carried out on a sample of 26 subjects under controlled conditions with a limited number of operators. Therefore, future studies should extend the analysis to a larger and more heterogeneous sample in real clinical settings, to assess the system's effectiveness in clinical scenarios with increased variability and operational complexity.

Bibliography

[1] Medical News Today. «Cardiovascular System: Structure and Function». In: (2024). URL: https://www.medicalnewstoday.com/articles/cardiovascular-system (cit. on p. 2).

[2] Encyclopædia Britannica. «Human Cardiovascular System». In: (2024). URL: https://www.britannica.com/science/human-cardiovascular-system (cit. on pp. 3, 4).

[3] TeachMePhysiology. «The ECG». In: (2024). URL: https://teachmephysiology.com/cardiovascular-system/cardiac-cycle-2/the-ecg/ (cit. on p. 3).

[4] Humanitas. *Vena cava inferiore*. URL: https://www.humanitas.net/wiki/inferior-vena-cava/ (cit. on p. 6).

[5] Piero Policastro and Luca Mesin. «Processing Ultrasound Scans of the Inferior Vena Cava: Techniques and Applications». In: *Diagnostics* (2023). URL: https://www.mdpi.com/2075-4418/13/2/329 (cit. on pp. 6, 7, 16, 24).

[6] Cindy L.. Stanfield. *Principles of human physiology*. Pearson, 2017, p. 782. ISBN: 9780134169804 (cit. on pp. 8, 11).

[7] Luca Mesin et al. «Assessment of Phasic Changes of Vascular Size by Automated Edge Tracking-State of the Art and Clinical Perspectives». In: *Frontiers in Cardiovascular Medicine* 8 (2022). ISSN: 2297055X. DOI: 10.3389/fcvm.2021.775635 (cit. on p. 12).

[8] Simon Finfer, Rinaldo Bellomo, Nicole Boyce, Craig French, John Myburgh, and Robert Norton. «A comparison of albumin and saline for fluid resuscitation in the intensive care unit». In: *The New England Journal of Medicine* 350.22 (2004), pp. 2247–2256. DOI: 10.1056/NEJMoa040232 (cit. on p. 12).

[9] Pablo Perel, Ian Roberts, and Katherine Ker. «Colloids versus crystalloids for fluid resuscitation in critically ill patients». In: *Cochrane Database of Systematic Reviews* 2013.2 (2013), p. CD000567. DOI: 10.1002/14651858.CD000567.pub6 (cit. on p. 12).

[10] Francisco José Teixeira-Neto and Alexander Valverde. «Clinical Application of the Fluid Challenge Approach in Goal-Directed Fluid Therapy: What Can We Learn From Human Studies?» In: *Frontiers in Veterinary Science* 8 (Aug. 2021), p. 701377. ISSN: 2297-1769. DOI: 10.3389/fvets.2021.701377 (cit. on p. 13).

[11] Paul E. Marik, Xavier Monnet, and Jean-Louis Teboul. «Hemodynamic parameters to guide fluid therapy». In: *Annals of Intensive Care* 1.1 (2011), p. 1. DOI: 10.1186/2110-5820-1-1 (cit. on p. 14).

[12] Diana J Kelm, Jared T Perrin, Rodrigo Cartin-Ceba, Ognjen Gajic, Louis Schenck, and Cassie C Kennedy. «Fluid overload in patients with severe sepsis and septic shock treated with early goal-directed therapy is associated with increased acute need for fluid-related medical interventions and hospital death». In: *Shock* 43.1 (2015), pp. 68–73. DOI: 10.1097/SHK.0000000000000268 (cit. on p. 14).

[13] Julien Jabot, Jean Louis Teboul, Christian Richard, and Xavier Monnet. «Passive leg raising for predicting fluid responsiveness: Importance of the postural change». In: *Intensive Care Medicine* 35 (1 Jan. 2009), pp. 85–90. ISSN: 03424642. DOI: 10.1007/s00134-008-1293-3 (cit. on p. 14).

[14] Xavier Monnet and Jean Louis Teboul. *Passive leg raising: Five rules, not a drop of fluid!* Jan. 2015. DOI: 10.1186/s13054-014-0708-5 (cit. on p. 14).

[15] Xavier Monnet, Paul Marik, and Jean Louis Teboul. «Passive leg raising for predicting fluid responsiveness: a systematic review and meta-analysis». In: *Intensive Care Medicine* 42 (12 Dec. 2016), pp. 1935–1947. ISSN: 14321238. DOI: 10.1007/s00134-015-4134-1 (cit. on p. 15).

[16] Sofia Melgarejo, Andrew Schaub, and Vicki E Noble. «Point of Care Ultrasound: An Overview». In: *American College of Cardiology* (Oct. 2017). Expert Analysis. URL: https://www.acc.org/latest-in-cardiology/articles/2017/10/31/12/42/point-of-care-ultrasound-an-overview (cit. on p. 15).

[17] J. Matthew Fields, Paul A. Lee, Katherine Y. Jenq, Dustin G. Mark, Nova L. Panebianco, and Anthony J. Dean. «The interrater reliability of inferior vena cava ultrasound by bedside clinician sonographers in emergency department patients». In: *Academic Emergency Medicine* 18 (1 Jan. 2011), pp. 98–101. ISSN: 10696563. DOI: 10.1111/j.1553-2712.2010.00952.x (cit. on pp. 16, 55).

[18] Zhongheng Zhang, Xiao Xu, Sheng Ye, and Lei Xu. *Ultrasonographic measurement of the respiratory variation in the inferior vena cava diameter is predictive of fluid responsiveness in critically ill patients: Systematic review and meta-analysis*. 2014. DOI: 10.1016/j.ultrasmedbio.2013.12.010 (cit. on p. 16).

[19] Sofia Furtado and Luís Reis. «Inferior vena cava evaluation in fluid therapy decision making in intensive care: Practical implications». In: *Revista Brasileira de Terapia Intensiva* 31 (2 2019), pp. 240–247. ISSN: 19824335. DOI: 10.5935/0103-507X.20190039 (cit. on pp. 16, 25).

[20] Tim Harris, Timothy J. Coats, and Mohammed H. Elwan. «Fluid therapy in the emergency department: An expert practice review». In: *Emergency Medicine Journal* 35 (8 Aug. 2018), pp. 511–515. ISSN: 14720213. DOI: 10.1136/emermed-2017-207245 (cit. on p. 16).

[21] John S. Mattoon and Clifford R. Berry. *Principi fondamentali di ecografia diagnostica*. 2023 (cit. on pp. 17, 19–21).

[22] Mario Meola, E. Buoncristiani, A. Cupisti, V. Cozza, and G. Bartotti. «Principi fisici dell'ultrasonografia e criteri topografici in ecografia renale». In: *Giornale di Clinica Nefrologica e Dialisi* 4 (1992). DOI: 10.33393/gcnd.1992.1975. URL: https://journals.aboutscience.eu/index.php/gcnd/article/view/1975 (cit. on pp. 18, 20).

[23] World Health Organization. *Manual of Diagnostic Ultrasound*. 2nd ed. Vol. 1. Geneva: World Health Organization, 2011 (cit. on p. 22).

[24] OpenAnesthesia, American Society of Anesthesiologists. *Inferior Vena Cava Point-of-Care Ultrasound*. 2024. URL: https://www.openanesthesia.org/keywords/inferior-vena-cava-point-of-care-ultrasound/ (cit. on p. 24).

[25] Nalan Gökçe Çelebi Yamanoglu, Adnan Yamanoglu, Ismet Parlak, Pelin Pinar, Ali Tosun, Burak Erkuran, Gizem Aydinok, and Fatih Torlak. «The role of inferior vena cava diameter in volume status monitoring: The best sonographic measurement method?» In: *American Journal of Emergency Medicine* 33.3 (2015), pp. 433–438. ISSN: 1532-8171. DOI: 10.1016/j.ajem.2014.12.014 (cit. on p. 24).

[26] Luca Mesin, Paolo Pasquero, Stefano Albani, Massimo Porta, and Silvestro Roatta. «Semi-automated tracking and continuous monitoring of inferior vena cava diameter in simulated and experimental ultrasound imaging». In: *Ultrasound in Medicine and Biology* 41 (3 Mar. 2015), pp. 845–857. ISSN: 1879291X. DOI: 10.1016/j.ultrasmedbio.2014.09.031 (cit. on p. 27).

[27] Tomohiro Sonoo et al. «Prospective analysis of cardiac collapsibility of inferior vena cava using ultrasonography». In: *Journal of Critical Care* 30 (5 Oct. 2015), pp. 945–948. ISSN: 15578615. DOI: 10.1016/j.jcrc.2015.04.124 (cit. on pp. 27, 85).

[28] Stefano Albani et al. «Inferior Vena Cava Edge Tracking Echocardiography: A Promising Tool with Applications in Multiple Clinical Settings». In: *Diagnostics* (2022). URL: https://www.mdpi.com/2075-4418/12/2/427 (cit. on p. 27).

[29] J. W. Bartlett and C. Frost. «Reliability, repeatability and reproducibility: analysis of measurement errors in continuous variables». In: *Ultrasound in Obstetrics & Gynecology* 31.4 (2008), pp. 466–475. DOI: 10.1002/uog.5256 (cit. on p. 28).

[30] Sarah Leclerc et al. «Deep Learning for Segmentation using an Open Large-Scale Dataset in 2D Echocardiography». In: *IEEE Transactions on Medical Imaging* (2019) (cit. on p. 29).

[31] Sekeun Kim, Hyung-Bok Park, Jaeik Jeon, Reza Arsanjani, Ran Heo, Sang-Eun Lee, Inki Moon, Sun Kook Yoo, and Hyuk-Jae Chang. «Fully automated quantification of cardiac chamber and function assessment in 2D echocardiography: clinical feasibility of deep learning-based algorithms». In: *The International Journal of Cardiovascular Imaging* (2022) (cit. on p. 29).

[32] Jean-Philippe Lainé, Joachim Beorchia, Sébastien Montagne, and Laurent D. Cohen. «Segmenting the carotid-artery wall in ultrasound image sequences with a dual-resolution U-net». In: *Computer Methods and Programs in Biomedicine* (2022) (cit. on p. 29).

[33] Telemed Medical Systems. *MicrUs EXT - Sistema Ecografico Portatile*. 2024. URL: https://www.telemedultrasound.com/wp-content/uploads/2024/06/MicrUs-Brochure-it.pdf (cit. on p. 34).

[34] Luca Mesin, Tatiana Giovinazzo, Simone D'Alessandro, Silvestro Roatta, Alessandro Raviolo, Flavia Chiacchiarini, Massimo Porta, and Paolo Pasquero. «Improved Repeatability of the Estimation of Pulsatility of Inferior Vena Cava». In: *Ultrasound in Medicine and Biology* 45 (10 Oct. 2019), pp. 2830–2843. ISSN: 1879291X. DOI: 10.1016/j.ultrasmedbio.2019.06.002 (cit. on pp. 37, 38, 84).

[35] Luca Mesin, Paolo Pasquero, and Silvestro Roatta. «Tracking and Monitoring Pulsatility of a Portion of Inferior Vena Cava from Ultrasound Imaging in Long Axis». In: *Ultrasound in Medicine and Biology* 45 (5 May 2019), pp. 1338–1343. ISSN: 1879291X. DOI: 10.1016/j.ultrasmedbio.2018.10.024 (cit. on p. 38).

[36] Python Software Foundation. *Tkinter — Python interface to Tcl/Tk*. 2024. URL: https://docs.python.org/3/library/tkinter.html (cit. on p. 42).

[37] Patrick E Shrout and Joseph L Fleiss. «Intraclass Correlations : Uses in Assessing Rater Reliability». In: *Psychological Bulletin* 86 (2 1979), pp. 420–428 (cit. on p. 45).

[38] Arash Salarian. *Function name "ICC"*. Software distributed under the BSD license. 2016 (cit. on p. 45).

[39] Terry K. Koo and Mae Y. Li. «A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research». In: *Journal of Chiropractic Medicine* 15 (2 June 2016), pp. 155–163. ISSN: 15563707. DOI: 10.1016/j.jcm.2016.02.012 (cit. on p. 45).

[40] L. Douglas Case and Walter T. Ambrosius. «Power and sample size.» In: *Methods in molecular biology (Clifton, N.J.)* 404 (2007), pp. 377–408. ISSN: 10643745. DOI: 10.1007/978-1-59745-530-5_19 (cit. on p. 47).

[41] Holger Schielzeth et al. «Robustness of linear mixed-effects models to violations of distributional assumptions». In: *Methods in Ecology and Evolution* 11 (9 Sept. 2020), pp. 1141–1152. ISSN: 2041210X. DOI: 10.1111/2041-210X.13434 (cit. on p. 47).

[42] Yasuo Iwaki and Tadao Inamura. «Study of the effective freedom degree in Welch-Satterthwaite formula: To applied blood chemical analysis». In: *SICE-ICASE International Joint Conference*. Applied Chaos Research Office, Kyoto, Japan. Tel: +81-75-922-9064. Bexco, Busan, Korea: SICE, 2006, Oct. 18–21 (cit. on p. 47).

[43] MathWorks. *coeftest (LinearMixedModel)*. The MathWorks, Inc. 2024. URL: https://it.mathworks.com/help/stats/linearmixedmodel.coeftest.html (cit. on p. 48).

[44] Mark Rubin. «Do p Values Lose Their Meaning in Exploratory Analyses? It Depends How You Define the Familywise Error Rate». In: *Review of General Psychology* 21.3 (2017), pp. 269–275. DOI: 10.1037/gpr0000123. URL: https://doi.org/10.1037/gpr0000123 (cit. on p. 48).

[45] Jacob Cohen. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Routledge, 1988. DOI: 10.4324/9780203771587. URL: https://doi.org/10.4324/9780203771587 (cit. on p. 49).

[46] J. Martin Bland and Douglas G. Altman. «Statistical Methods for Assessing Agreement Between Two Methods of Clinical Measurement». In: *The Lancet* (1986) (cit. on p. 49).

[47] Francesco Franco and Anteo Di Napoli. «Valutazione della concordanza tra misurazioni di caratteri di tipo quantitativo: il metodo di Bland-Altman». In: *Epidemiologia & Prevenzione* (2007) (cit. on p. 49).