

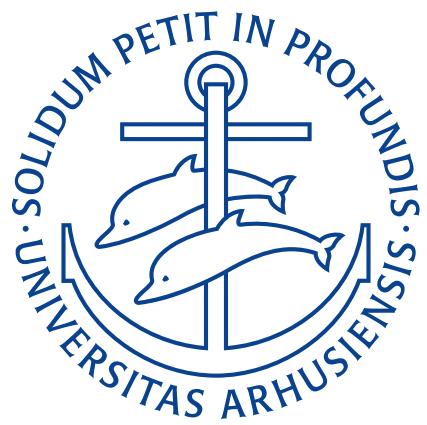


AARHUS
UNIVERSITY

Fluid Responsiveness Prediction During Surgery

- Physiological and Methodological Limitations and Considerations

PhD dissertation
Johannes Enevoldsen



Health
Aarhus University
Department of Clinical Medicine
2022

For my girls

Acknowledgements

Thanks to my supervisor, Simon Tilma Vistisen, colleagues, family and friends, as well as Aarhus University and Holger & Ruth Hesse's mindefond for founding this work.

Also, thanks to all the kind strangers on the internet who selflessly helped me with various statistical and technical issues. Special thanks to John George Karippacheril, developer of [VSCapture](#), who helped me customize his software to work with department ventilators; essential for acquiring the data used in study 2 and study 3. Also, special thanks to Gavin L. Simpson, who after helping me with several issues posted on <[stackoverflow.com](#)> and Twitter, agreed to co-write the paper resulting from Study 2.

Johannes Enevoldsen
Aarhus University
Fall, 2022

Abstract

Here is about one page of what this dissertation is about.

Dansk Resumé

Her er et resume af afhandlingen.

Contents

List of Figures	vi
List of Abbreviations	vii
Papers	1
1 Introduction	2
2 Background	4
2.1 History	4
2.2 Why give IV fluids?	5
2.3 The physiology and pathophysiology of a fluid bolus	7
2.4 How much fluid should we give and when?	13
Appendices	
A Paper 1	17
B Paper 2	26
C Paper 3	40
References	41

List of Figures

2.1 Steps from fluid administration to benefit	5
--	---

List of Abbreviations

AUROC	Area under the receiver operating characteristic curve.
CO	Cardiac output.
SV	Stroke volume.
SV	Stroke volume.
CI	Compatibility interval, e.g 95% CI [19.3; 25.2]. For Bayesian estimates, it is a <i>credible interval</i> ; for frequentist estimates, it is a <i>confidence interval</i> . In both cases it represents the interval of parameter values compatible of the observed data.

Papers

This PhD dissertation is based upon the following three papers:

- **Paper 1** Existing fluid responsiveness studies using the mini-fluid challenge may be misleading: Methodological considerations and simulations. *Published in Acta Anaesthesiologica Scandinavica, 2021. DOI: 10.1111/aas.13965.*
- **Paper 2** Using generalized additive models to decompose time series and waveforms, and dissect heart–lung interaction physiology. *Published in Journal of Clinical Monitoring and Computing, 2022. DOI: 10.1007/s10877-022-00873-7.*
- **Paper 3**

Medicine is a science of uncertainty and an art of probability.

— **Sir William Osler** (1849–1919).

1

Introduction

Fluid therapy is a ubiquitous medical intervention; both in the perioperative setting and for hospitalised patients in general. The aim of fluid therapy is to restore the patient's circulating blood volume to the *optimum*, normal, level. Hence, the terms “fluid *resuscitation*” and “fluid *replacement* therapy” are commonly used.

Intraoperative fluid management is mainly relevant in acute or long-duration surgery. In acute surgery, the preoperative fluid status is generally unknown, and it is often reasonable to assume that the patient arrives at the operating room dehydrated. In long-duration surgery, continuous loss of fluid through bleeding, perspiration and urination necessitates fluid replacement through the operation. Since these patients are anaesthetised, intravenous (IV) is the administration route of choice.

Like every treatment, IV fluid should only be given to patients who will benefit from the fluid. This is the setup for a prediction problem: can we, ahead of time, predict whether a patient will benefit from an intravenous fluid administration? The first task is to define what we mean with *benefit*, and how we can measure it. This is not trivial, and it will be discussed further in section 2.2. Luckily, a necessary (but not sufficient) condition for benefitting from a fluid administration is that the fluid causes an increase in cardiac output (CO). This can be measured, and allows us to formulate a simpler prediction problem:

1. Introduction

Can we predict whether a patient's CO will increase from a fluid administration?

There is an entire subfield of anaesthesia and intensive care research dealing with this problem. This PhD dissertation describes a small, but hopefully meaningful addition to the field.

In this dissertation, I present the available tools for fluid responsiveness prediction, with focus on the intraoperative setting. The dissertation covers 3 papers that tackle specific limitations to the most common tools.

The terms *fluid challenge* and *bolus* are used interchangeably. *Stroke volume* (SV) and *cardiac output* (CO) are often interchangeable and the term that best fits the context is used. Ventilation and respiration is also used interchangeably and refer to the act of breathing (either mechanically or spontaneously).

The most wonderful and satisfactory effect is the immediate consequence of the injection [of fluid]. To produce the effect referred to, a large quantity must be injected—from five to ten pounds in an adult—and repeated at longer or shorter intervals, as the state of the pulse, and other symptoms, may indicate.

— **Robert Lewins, M.D.**, 1832 (Injection of Saline Solutions Into the Veins).

2

Background

2.1 History

Intravenous fluid therapy first became popular in the cholera epidemic around 1830, when Thomas Latta “threw” several litres of saline into the veins of severely dehydrated cholera patients, and Robert Lewins reported enthusiastically on the “most wonderful and satisfactory effect” [7,8]. After the end of the epidemic, the treatment was mostly abandoned. Possibly because the early clinical reports from the epidemic mainly presented temporary effects and mostly in morbidly dehydrated patients, and also because the concerns raised by sceptics were probably highly relevant: the fluid was both unsterile and hypotonic [3].

Interest in fluid resuscitation reemerged nearly 50 years later, in 1879, when Kronecker and Sander demonstrated the importance of volume (as opposed to red blood cells) in the treatment of haemorrhage: They bleed down two dogs until bleeding stopped from lack of cardiac activity (approx. 50 % of the blood volume). Then, they reported how resuscitation with an equivalent volume of a saline solution would recover the animals’ cardiac activity [4,11]. This was followed by a number of reports of successful IV fluid resuscitations in humans [4].

Since the end of the nineteenth century, IV fluid administration has been a staple in the treatment of the acutely ill and during surgery. With better equipment

2. Background

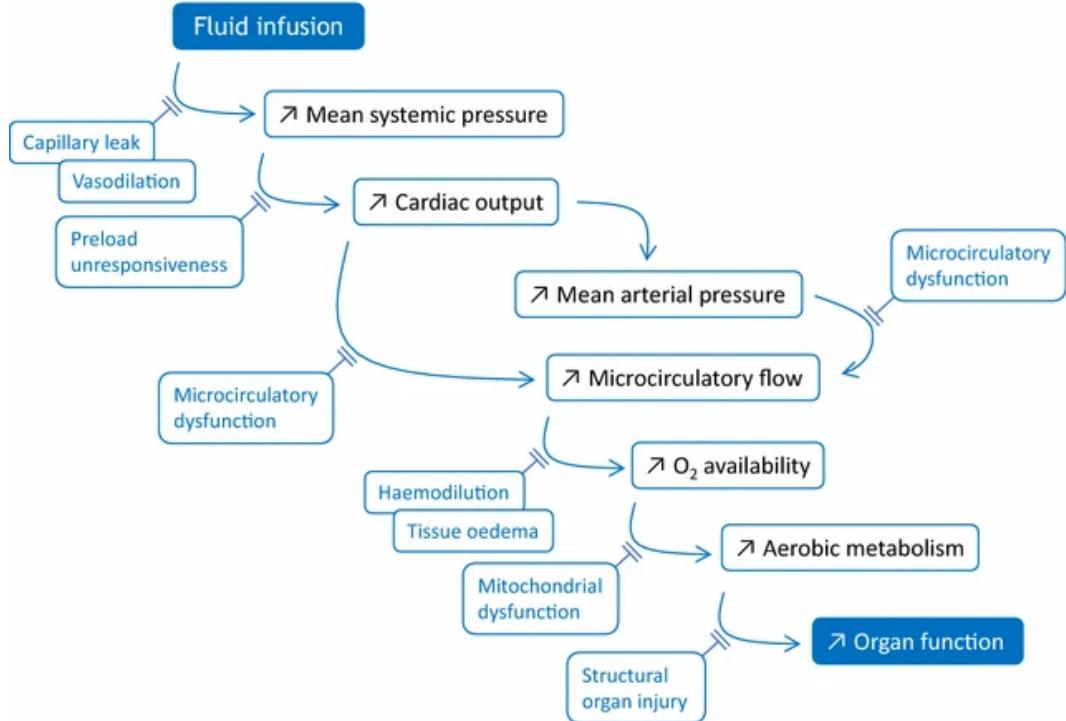


Figure 2.1: Illustration of the physiological steps from fluid administration to benefit. Reprinted from Monnet et al, 2018 [10] (CC BY).

and hygiene, the safety of IV fluid administration has increased, and the indication for treatment has widened accordingly—today, IV fluid is even available as a drop-in or home-delivery hangover remedy [6]. Naturally, debates about the appropriate use of IV fluids continue.

2.2 Why give IV fluids?

Fluid should only be administered when the patient is likely to benefit from it [9]. As a treatment for hypovolemia, the goal is that the infused fluid will increase the circulating volume and thereby the mean systemic pressure. This gives an increase in cardiac preload and, through the Frank-Starling mechanism (see section ref:frank-starling), an increase in CO. This should increase microcirculation and, in turn, the oxygen available to organs, increasing (or retaining) organ function (see Figure 2.1) [10]. As discussed in section ref:fluid-bolus-physiology, this goal is not always achieved.

2. Background

Another common aim with IV fluid therapy is to resuscitate dehydration, characterised by hypertonicity (excess salt) due to loss of water—e.g., due to gastroenteritis or diabetic polyuria. While hypovolemia can be corrected rapidly with IV infusion of an isotonic fluid, hypertonicity should be corrected slowly through oral intake of water or with IV hypotonic fluid [1]. This dissertation focuses on fluid as a treatment of hypovolemia, and dehydration/hypertonicity will not be discussed further.

During surgery, there is a continuous loss of circulation fluid through urination, bleeding, perspiration and *redistribution*. If this fluid is not replaced, the patient will gradually become hypovolemic and, eventually, organ failure will occur. Since blood loss and urination is accurately recorded throughout the surgery, the unknown volume of fluid to replace is from perspiration and redistribution [5]. An additional unknown is the patient's fluid status on arrival to the operating room. Preoperative hypovolemia due to extended fasting before operations used to be a relevant concern, and was treated with significant volumes of fluid during or before induction of anaesthesia [2,5]. With today's more liberal guidelines for pre-surgery fluid intake, allowing clear fluids until two hours before surgery, the preoperative deficit is probably lower [5,9,12].

With the risk of organ failure from hypovolemia, why not simply give the patient some extra fluid to ensure that there is enough? ## Concerns about liberal use of IV fluids

The statement that IV fluid administration should be treated as any other prescription medication has become a trope in fluid resuscitation literature [cite miller 2019; Myburgh 2013; Hoste 2014; Malbrain 2020]. For good reasons. In addition to the intended effect summarised in the section above, fluid administration can cause side effects. The principal concerns with excessive fluid administrations can be divided into hypervolemia and non-volume-related effects. The non-volume-related effects depend on the type of fluid. Notable examples include hyperchloremic acidosis from normal saline and kidney injury associated with excessive? HES infusion [cite Myburg 2013; perner 2012]. Hypervolemia is a more general issue

2. Background

with excessive fluid administration, causing oedemas of tissue and lungs. The pathophysiology of hypervolemia involves both cardiovascular mechanics and microcirculation, so this will be a good place for an introduction to the physiology related to fluid administrations.

2.3 The physiology and pathophysiology of a fluid bolus

A fluid bolus should increase stroke volume (SV) and, hence, CO. This requires that the fluid remains in circulation to increase cardiac preload, and that this increase in cardiac preload causes an increase in SV. We will start with the latter condition: the Frank-Starling mechanism. #### The Frank-Starling mechanism {#frank-starling}

The relation between cardiac preload and SV is often termed the “Frank-Starling mechanism” or the “Frank-Starling law of the heart”, after Otto Frank and Ernest Starling, who are commonly attributed the discovery of the relationship [cite frank1895; patterson1914]. It has been noted, though, that the phenomenon had been observed several decades earlier [cite zimmer2002]. The Frank-Starling mechanism describes that increasing the length of a cardiomyocyte will increase the force generated when the muscle is activated; or, as a consequence, that increased filling of a ventricle will increase stroke volume. The relation occurs only until a certain length/volume, where the curve flattens and the effect eventually reverses. This was clearly demonstrated in experiments by Otto Frank, where he measured the pressure generated through isovolumetric contractions of frog hearts (see figure ref:background-frank1895).

(Figure: background-frank1895): Otto Frank’s tracings of intraventricular pressure during isovolumetric contractions. In each panel, ventricular volume increases with the number on the tracing. Right panel, tracing 4 demonstrates the decrease in maximum pressure with overdistension. Reproduction of figures from Frank 1895, [cite frank1895] public domain.

2. Background

A clinical consequence of the Frank-Starling mechanism is that a patient's SV can only increase from a fluid bolus, if the heart is currently functioning on the rising section of the Frank-Starling curve (see figure ref:background-starling-curve). A heart's Frank-Starling curve is, however, not constant, but can be raised with higher sympathetic tone or sympathomimetic drugs (positive inotropes).

(Figure: background-starling-curve): Illustration of the Frank-Starling mechanism—the relation between the end-diastolic volume and the stroke volume. If end-diastolic volume is increased in a heart operating on the steep part of the curve (e.g. by a fluid bolus) stroke volume will increase. If the heart is already operating on the flat part of the curve, no increase in stroke volume can be expected.

A few concepts are used interchangeably for both the cause and the effect in the Frank-Starling relationship. Starling et al. note that the physiological relationship must be between the length of a piece of cardiac muscle and the tension it exerts when it contracts, but, since they are not able to measure these in an intact heart, assume that ventricular volume is linearly related to muscle length, and that ventricular pressure is linearly related to muscle tension. They appropriately note that this assumption will become increasingly incorrect with distension and a more globular shape of the ventricle [cite patterson1914]. Other terms used as proxies for end-diastolic muscle length are preload and end-diastolic pressure. Preload should be synonymous with end-diastolic volume or muscle length, but the term is often not well defined. End-diastolic pressure is of course related to muscle length, but it also depends on static mechanical factors (e.g. fibrosis), external pressure and the shape and size of the heart. Proxies for the systolic muscle tension include afterload, systolic ventricular pressure, stroke volume and mechanical work. ### Venous return and mean systemic filling pressure In steady state, blood circulates continuously from the heart to the arteries, through tissues, to the veins and back to the heart. The circulating volume is constant and CO is equal to the venous return to the heart. We can consider a simple model of this system, where blood is pumped from a small elastic compartment into a larger elastic compartment and returned again to the smaller elastic compartment through

2. Background

a tube (see figure {ref:background-venous-return-simple}). The pump is the heart, the large compartment represents the capacitance of venules and veins—completely neglecting the capacitance of the arteries—and the smaller compartment represents the right atrium and large veins immediately upstream from the heart. Arteries are neglected in this model because of their low compliance relative to the venous system (for an illustration of how the arterial system would fit in this model, see figure fig:background-venous-return-art) [cite magder 2016 volume]. The pressure in the large compartment is the venous pressure (P_V), the pressure in the smaller compartment is the right atrial pressure (P_{RA}) and the resistance in the tube is the resistance to venous return (R_V). If the pump is stopped, both compartments will reach an equilibrium pressure: the mean systemic filling pressure (P_{MSF}).

(Figure: background-venous-return-simple): A simple model illustrating the concepts of venous return (Q_V) and mean systemic filling pressure (P_{MSF}). The pressure is defined by the height of the fluid surface and the compliance is proportional to the width of the compartment. CO, cardiac output. P_V , pressure of the compartment representing venules and veins. R_V , resistance to venous return. P_{RA} , pressure of the compartment representing the right atrium.

(Figure: background-venous-return-art): An illustration of how the arterial system could be represented in the simple model illustrated in figure fig:background-venous-return-simple. MAP, mean arterial pressure.

From this simple model, we can appreciate some factors that determine CO. First, the heart's ability to pump is the absolute limitation to cardiac output. However, the heart also cannot pump more than what is returned from the veins. This venous return is determined by the resistance to venous return and the pressure difference between the venous compartment and the right atrial compartment:

$$CO = Q_V = \frac{P_V - P_{RA}}{R_V}.$$

Thus, increasing venous pressure or lowering resistance to venous return *allows* a higher CO. If venous compliance is constant, a fluid bolus can increase venous

2. Background

pressure. Alternatively, we can use an α -adrenergic agonist such as noradrenaline to decrease venous compliance and thereby increase pressure without adding fluid (venoconstriction will also increase resistance to venous return, but the effect on compliance seems to dominate) [cite persichini2022]. Models of venous return often divide compartments into stressed volume and unstressed volume. The unstressed volume is the volume of fluid that will not create any pressure in the compartment—essentially the volume that will remain if the circulation is lacerated. Unstressed volume is effectively inert, and only the stressed volume has any influence on venous return. Unstressed volume can, however, become stressed through vasoconstriction. While the concept of unstressed volume makes sense anatomically (a vessel can be filled to a certain volume without exerting an elastic recoil), it will often be difficult to differentiate between the effects of *unstressed volume becoming stressed* and “a decrease in compliance of already stressed volume* (both increase P_{MSF}).

As described in the section above, the Frank-Starling curve describes the relationship between ventricular filling and CO (via SV). Ventricular filling is positively related to the right atrial pressure, while the right atrial pressure is inversely related to CO, since a high CO will tend to empty the right atrial compartment. The right atrial pressure and CO where venous return and cardiac output are in equilibrium can be found as the intersection between the venous return curve (the relationship between venous return and right atrial pressure) and a variant of the Frank-Starling curve with CO rather than SV on the y-axis (see figure fig:guyton). This graphical solution was first proposed by Arthur Guyton [cite Guyton1957].

(Figure: background-guyton): A) The relationship between venous return (Q_V), right atrial pressure (P_{RA}) and cardiac output (CO). If CO (and hence Q_V) drops to zero, P_{RA} will equal the mean systemic filling pressure (P_{MSF}). The circulation is at steady state at the intersection of the venous return function and the cardiac function (when $Q_V = CO$). This illustration corresponds to spontaneous breathing, where the intrathoracic pressure is negative. Therefore, the cardiac function starts at a negative pressure where the ventricular transmural pressure ($P_{tm,RV}$) is zero. B) Change in inotropy or heart rate. C) Change in resistance to venous return

2. Background

(R_V). D) Change in P_{MSF} ; either via change in stressed volume or compliance of capacitance vessels.

The simple model depicted in figure fig:venous-return-simple and Guyton's graphical solution the steady state CO, provides a basis for understanding clinical interventions that impact CO. One category of interventions target the heart directly by increasing inotropy or chronotropy (increase in cardiac function). A common drug with this effect is dobutamine, which has both positive inotrope and chronotrope effects. In isolation, positive inotropy or chronotropy will increase CO and lower P_{RA} , as depicted in figure fig:guyton B. From this figure, we can also identify the theoretical maximum CO obtainable from inotropy or chronotropy: when the heart essentially pumps the right atrium "dry" faster than the venous return can refill it. Since veins are flaccid, they cannot have a transmural pressure below zero. Lowering P_{RA} below zero (only possible because the intrathoracic pressure is below zero) will not further increase venous return as extrathoracic veins will collapse in proportion to the lower P_{RA} (depicted as the left steady state point in figure fig:guyton B). This is known as the "waterfall effect", since it is analogous to how changing the lower water level in a waterfall will not affect the flow over the waterfall (cite:Permutt1963).

A second target for optimising CO, is resistance to venous return (R_V). This resistance can be greatly increased with liver cirrhosis, and alleviation by transjugular intrahepatic portosystemic shunt (TIPS) increases CO (cite: wong1995). Late stages of pregnancy can also increase R_V via compression of the inferior vena cava when the mother is in supine position. Increase in R_V does not impact P_{MSF} but reduces P_{RA} and thereby CO (see figure fig:guyton C).

The last point of intervention is P_{MSF} . A fluid bolus will increase P_{MSF} by increasing stressed volume while maintaining compliance of capacitance veins. Venoconstriction (e.g. with noradrenaline) decreases compliance of capacitance veins, and may additionally mobilise previously unstressed volume; both effects increase P_{MSF} . An increase of P_{MSF} increases P_{RA} and thereby CO on the condition that

2. Background

the heart is operating on the ascending part of the Frank-Starling curve (see figure fig:guyton C). #### Fluid distribution and oedema formation

A principal adverse effect of fluid administration is oedema: a pathological build-up of fluids in the intercellular tissue or within alveoli. Additional fluid in the interstitium increases the diffusion distance between the capillary blood and the cells, decreasing the rate of oxygen delivery to the mitochondria [cite dunn2016]. Pulmonary oedema has a similar detrimental effect on gas exchange in the alveoli.

The classic understanding of oedema formation

The mechanism for oedema formation is classically described with Ernest Starling's understanding of capillary physiology: The interstitial fluid is in an equilibrium between the colloid-osmotic (oncotic) pressure from the macromolecules in blood (pull) and the hydrostatic pressure across the capillary membrane (push). An increase in stressed volume increases transcapillary pressure, driving fluid into the interstitium until a new equilibrium is reached [cite boronMicrocirculation2016]. Adding to this, crystalloid fluids (e.g. normal saline or acetated Ringer's solution) dilute plasma, which lowers the oncotic pressure and further promotes the formation of oedema.

A revised mechanism of oedema formation: the endothelial glycocalyx

In recent years, increasing focus has been on the endothelial glycocalyx layer's role in fluid resuscitation and oedema formation. The glycocalyx is a gel of macromolecules (mainly glycoproteins, hyaluronans and proteoglycans) lining the vascular endothelium (cite:weinbaum2007). One function of this layer is to form a semipermeable membrane that, in addition to retaining plasma proteins, also retains water to a variable degree. In this model, the flow of water from vaculatere to tissue is determined less by oncotic pressure difference and more by the current state of the glycocalyx layer (cite:milfordRecusitationfluid2019). The permeability of the glycocalyx layer may be impacted by volume loading. A proposed mechanism for this regulation is that volume loading increases right atrial pressure, causing

2. Background

release of atrial natriuretic peptide (ANP). ANP increases water filtration and may, directly or indirectly, damage the glycocalyx (chappallHypervolemia2014). This mechanism has, however, later been questioned (cite:damenAtrial2021). Another important cause of glycocalyx degradation is inflammation—especially related to sepsis (cite:ibaDerangement2019).

How long does fluid remain in circulation?

Fluids must remain in circulation to benefit the patient's hemodynamic status. Both patient and fluid specific factors impact how long we can expect a fluid bolus to exert its intended effect. The intravascular half-life of a crystalloid infusion is around 20 to 40 minutes in conscious volunteers, while the half-life is more than doubled in surgery with general anaesthesia. Colloids are reported to expand plasma volume with a half-life of two to three hours for both healthy subjects and during surgery (the half-life of the macromolecules themselves in synthetic starches (HES) are much longer than the effect on volume expansion). Generally, a hypovolemic state is associated with a more persistent effect of a volume expansion (cite: hahnHalflife2016).

2.4 How much fluid should we give and when?

There are two overall strategies for fluid management: to replace fluids according to an estimated loss or deficit, or to give fluids until a specific hemodynamic target is reached.

The fluid replacement strategy is commonly investigated by comparing a *restrictive* strategy against a *liberal* strategy. The terms *restrictive* and *liberal* are, of course, relative, and through years with superior results from *restrictive* fluid regimens, both terms have referred to successively lower volumes. This trend seems to have been concluded with the RELIEF trial (cite: millerPerioperative2019).

In the RELIEF trial, 3000 patients undergoing abdominal surgery were randomised to either a *liberal* fluid regimen, expected to give a positive fluid balance,

2. Background

or to a *restrictive* fluid regimen, expected to give a neutral fluid balance. The *liberal* group received a median of 6.1 litres of fluid, while the *restrictive* group received a median of 3.7 litres. There was no difference in disability-free survival between the groups, but the *restrictive* group had a higher rate of acute kidney injury. The *liberal* group had a calculated fluid balance of +3.1 litres and gained 1.6 kg weight; the *restrictive* group had a +1.4 litre fluid balance and gained 0.3 kg (weight gain was only measured in one third of the patients). (mylesRestrictive2018) Overall, this suggests that a positive fluid balance of 1-2 litres is preferable in major abdominal surgery (cite brandstrupFinding2018).

The alternative—or complementary—strategy is goal directed hemodynamic therapy (GDT). Here, patients are treated with fluid, and often vasopressors, to reach a specific hemodynamic target. The aim with GDT is to individualise treatment to ensure that hypovolemic patients get enough fluid, while avoiding fluid overload. A common GDT target is CO optimisation: fluid is given in boluses until CO stops increasing. This is interpreted as the patient's heart having reached the plateau top of the Frank-Starling curve, and that further fluid administration will be futile. An example of GDT was investigated in the OPTIMISE trial (cite: pearseEffect2014).

The OPTIMISE trial randomised 724 high-risk, abdominal surgery patients to either CO-guided GDT or *usual care*. The GDT intervention consisted of a fluid administration algorithm where first, a patient's target SV was determined by administering colloid fluid in 250 ml boluses until a new bolus no longer caused a sustained increase in SV above 10%. Afterwards, fluids were administered to maintain this target SV. Additionally, dopexamine (inotrope) was infused at a low rate ($0.5 \text{ }\mu\text{g kg}^{-1} \text{ min}^{-1}$). The study results were inconclusive: they were suggestive of a protective effect of the GDT protocol on adverse events and mortality, though the results were also compatible with there being no difference between groups.

Generally, the effect of GDT is difficult to assess. Both because protocols are numerous and heterogenous and because *usual care* continues to assimilate the GDT protocols under investigation (cite: millerPerioperative2019). A recent systematic

2. Background

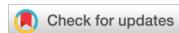
review of 76 randomised GDT trials had a conclusion similar to the OPTIMISE trial: GDT might work (cite: jessenGoaldirected2022).

There are some disadvantages to the SV-optimising approach to fluid therapy. First, it requires continuous monitoring of SV or CO. Recent technological advances have broadened the availability of continuous CO monitoring, though the accuracy of these technologies are debated (see section (ref:methods-co)). Second, it can be (??has been) argued that if fluid is given until it no longer increases SV, then the last fluid bolus was unnecessary and should not have been given. This could be avoided, if the response to a fluid bolus could be predicted.

Appendices

A

Paper 1



Received: 26 April 2021 | Revised: 2 July 2021 | Accepted: 19 July 2021
DOI: 10.1111/aas.13965

REVIEW



Existing fluid responsiveness studies using the mini-fluid challenge may be misleading: Methodological considerations and simulations

Johannes Enevoldsen^{1,2} | Thomas W. L. Scheeren³ | Jonas M. Berg^{1,2} |
Simon T. Vistisen^{1,2}

¹Department of Clinical Medicine, Aarhus University, Aarhus, Denmark

²Department of Anaesthesiology and Intensive Care, Aarhus University Hospital, Aarhus, Denmark

³Department of Anesthesiology, University of Groningen, University Medical Centre Groningen, Groningen, The Netherlands

Correspondence

Johannes Enevoldsen, Department of Clinical Medicine, Aarhus University, Palle Juul-Jensens Boulevard 82, 8200 Aarhus N, Denmark.

Email: enevoldsen@clin.au.dk

Abstract

Background: The mini-fluid challenge (MFC) is a clinical concept of predicting fluid responsiveness by rapidly infusing a small amount of intravenous fluids, typically 100 ml, and systematically assessing its haemodynamic effect. The MFC method is meant to predict if a patient will respond to a subsequent, larger fluid challenge, typically another 400 ml, with a significant increase in stroke volume.

Methods: We critically evaluated the general methodology of MFC studies, with statistical considerations, secondary analysis of an existing study and simulations.

Results: Secondary analysis of an existing study showed that the MFC could predict the total fluid response (MFC + 400 ml) with an area under the receiver operator characteristic curve (AUROC) of 0.92, but that the prediction was worse than random for the response to the remaining 400 ml (AUROC = 0.33). In a null simulation with no response to both the MFC and the subsequent fluid challenge, the commonly used analysis could predict fluid responsiveness with an AUROC of 0.73.

Conclusion: Many existing MFC studies are likely overestimating the classification accuracy of the MFC. This should be considered before adopting the MFC into clinical practice. A better study design includes a second, independent measurement of stroke volume after the MFC. This measurement serves as reference for the response to the subsequent fluid challenge.

1 | INTRODUCTION

The term *mini-fluid challenge* (MFC) was coined by Muller et al about a decade ago¹ as a new way to predict fluid responsiveness. At the time, common fluid infusion practice consisted of 'let's give some fluid and see what happens' as highlighted by the accompanying editorial.² That 'some fluid' was a fluid challenge of around 500 ml as identified by the FENICE study.³ Motivated by the finding that fluid is not harmless and may induce fluid overload, Muller et al suggested the MFC: the haemodynamic effect of a rapid infusion of a small

amount of fluid could guide whether or not a larger amount of fluid should be given. The authors tested whether the change in aortic velocity time integral (VTI; an echocardiographic measure correlated with stroke volume [SV]) induced by the MFC (100 ml within 1 min) could predict the effect of a 'normal' fluid challenge of 500 ml, specifically, the combined effect of the MFC and another 400 ml. The method was highly predictive (area under the receiver operating characteristic [ROC] curve, AUROC, of 0.92).¹ Others have since investigated and validated the MFC, and a recent systematic review including seven MFC studies (368 fluid challenges in 324 patients)^{1,4–9}

identified a pooled AUROC of 0.91 for the MFC method.¹⁰ Since the systematic review, more MFC studies have been published, all pointing to the same compelling conclusion: that the method is accurate in predicting fluid responsiveness.¹¹⁻¹⁵

In 2018, we published a correspondence debating the way MFC studies were designed.¹⁶ The correspondence raised clinical and statistical issues with the most adopted methodology. Yet, the notion that optimal MFC methodology may not be completely settled has hardly influenced methodology in subsequent publications. In this paper, we will:

- explain in simple terms the problems with the most frequently used MFC method
- demonstrate, by secondary analysis of an existing study and by simulations, the potential magnitude of the problem
- discuss strengths and limitations of less frequently used designs
- give recommendations on the way forward for researching this otherwise compelling method.

1.1 | A representative MFC study design

To simplify the key message, we will consider and discuss a representative MFC study design as depicted in Figure 1: 100 ml fluid is infused within 1 min (the MFC), the haemodynamic response (relative SV change) of that MFC is evaluated, and subsequently another 400 ml fluid (totalling 500 ml) is infused over 15 min. The final response (outcome) is evaluated as a relative SV change from baseline (i.e. before any fluid administration) to after the full amount of 500 ml. While we use SV in the examples, the arguments can be generalised to any method for estimating SV or cardiac output.

2 | METHOD

Figure 1 identifies that calculations of the haemodynamic response to the MFC (ΔSV_{100}) and the response to the full fluid challenge (ΔSV_{500}) both include the haemodynamic variable measured at baseline, that is before the MFC. Specifically, ΔSV_{100} and ΔSV_{500} are calculated as

$$\Delta SV_{100} = \frac{SV_{100} - SV_{\text{baseline}}}{SV_{\text{baseline}}}$$

and

$$\Delta SV_{500} = \frac{SV_{500} - SV_{\text{baseline}}}{SV_{\text{baseline}}}.$$

This shared baseline causes the problem.¹ It introduces two effects that, in addition to a true classification accuracy, can explain the high classification accuracy found in several MFC studies:

1. The predictor and the outcome share measurement error, creating a spurious correlation.

Editorial Comment

This review presents a detailed assessment of methodological aspects of studies assessing clinical effects of a form of intravascular fluid administration challenge. Findings are presented which demonstrate how many clinical reports in this area of inquiry can contain bias related to the choice of assessment variables, which must be considered when interpreting results. The authors suggest possible means to improve reliability for results related to methodological choices.

2. The predictor (ΔSV_{100}) is also a part of the outcome we try to predict (ΔSV_{500}).

2.1 | Shared error

Any measurement is associated with uncertainty (error). This can be subdivided into a systematic error (often referred to as bias) and a random error (often referred to as variance and defining precision).^{17,18} It is useful to think of a 'true' SV and a random error around this value. The 'true' SV is what the clinician wants to measure, and what they hope to increase with a fluid infusion. The random error comprises both the imprecision of the monitoring equipment and minor temporal (minute-wise) physiologic changes in haemodynamics that are effectively noise in the context of evaluating a fluid response. It is the random error on the baseline measurement that causes the problem. In the following equations, each measured SV is divided into a 'true' SV and a random measurement error.

$$\begin{aligned}\Delta SV_{100} &= \frac{(SV_{100} + \epsilon_{SV,100}) - (SV_{\text{baseline}} + \epsilon_{SV,\text{baseline}})}{SV_{\text{baseline}} + \epsilon_{SV,\text{baseline}}} \\ &= \frac{SV_{100} + \epsilon_{SV,100}}{SV_{\text{baseline}} + \epsilon_{SV,\text{baseline}}} - 1,\end{aligned}$$

and

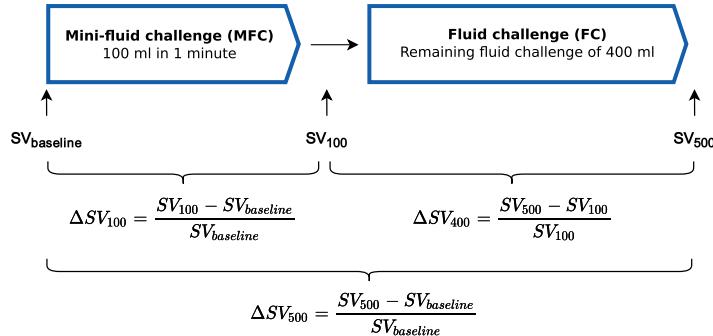
$$\Delta SV_{500} = \frac{SV_{500} + \epsilon_{SV,500}}{SV_{\text{baseline}} + \epsilon_{SV,\text{baseline}}} - 1.$$

These two equations essentially depict the problem: the random error ($\epsilon_{SV,\text{baseline}}$) is part of the denominator in the calculation of both the predictor (ΔSV_{100}) and the outcome (ΔSV_{500}), making them spuriously correlated, and therefore more likely to agree.¹⁹

2.2 | The predictor (ΔSV_{100}) is also a part of the outcome we try to predict (ΔSV_{500})

The MFC should be used as a predictive method, that is to decide whether to administer the remaining 400 ml fluid or not. Thus,

FIGURE 1 Representation of the design used in most mini-fluid challenge (MFC) studies. In this design, stroke volume (SV) is measured three times: (1) At baseline, (2) after the MFC and (3) after the full fluid challenge



when evaluating the accuracy of the MFC as a predictor of fluid responsiveness, only the effect of the last 400 ml should define the outcome. The naive solution would be to use the MFC (ΔSV_{100}) to predict ΔSV_{400} (see Figure 1). Unfortunately, this does not solve the shared error problem. The random variation of the SV_{100} measurement will introduce a similar problem, since that variable is now a constituent of both the predictor (ΔSV_{100}) and outcome (ΔSV_{400}) variables (see Figure 1). In this case, the random variation in SV_{100} will make the predictor and outcome variables less likely to agree, by creating a spurious, negative, correlation, leading to an underestimation of the true classification accuracy.

Both the problems described above arise from *mathematical coupling* of the predictor and outcome.^{20,21}

2.3 | Secondary analysis of an existing study

To illustrate what happens with classification, if we try to predict ΔSV_{400} instead of ΔSV_{500} , we extracted ΔSV_{100} and ΔSV_{500} from plot 3A in the pioneering study by Muller et al and calculated the corresponding ΔSV_{400} .¹ Data were captured using DataThief III (version 1.7, datathief.org). Although the study reported relative VTI changes (ΔVTI), we will continue to use the SV term for consistency.

ΔSV_{400} is defined as

$$\Delta SV_{400} = \frac{SV_{500} - SV_{100}}{SV_{100}}.$$

If ΔSV_{100} and ΔSV_{500} are known, we can calculate ΔSV_{400} :

$$\Delta SV_{500} + 1 = (\Delta SV_{100} + 1) \cdot (\Delta SV_{400} + 1).$$

Therefore,

$$\Delta SV_{400} = \frac{\Delta SV_{500} + 1}{\Delta SV_{100} + 1} - 1.$$

We then analysed ΔSV_{100} 's (MFC) ability to predict $\Delta SV_{400} > 15\%$.

2.4 | Simulations

Simulations can reveal how shared error can introduce a significant bias to the result of MFC studies. The magnitude of the problem in existing studies is impossible to calculate exactly, since some relevant variables have to be estimated, but a simulation can provide a ballpark estimate.

Using R (4.0.4) and R packages, pROC and Tidyverse,^{22–24} we simulated SV measurements at all three measurement points in Figure 1 (baseline, after 100 ml and after 500 ml fluid) for 2000 subjects. Annotated code generating the simulations is available from the digital Supplementary Material S1, and an interactive tool that allows changing simulation parameters is available from <https://johanenesne.shinyapps.io/mini-fluid-challenge-simulation/>.

2.4.1 | Simulation 1

First, we simulated how the MFC methodology performs in virtual patients whose SV are entirely unresponsive to fluid, but with random variation in SV measurements. Since there is nothing to predict, any apparent predictive ability is a statistical artefact. Each patient was assigned a constant 'true' SV for all three windows (mean = 75 ml, SD = 10 ml), with an additional random variation (mean = 0, SD = 3 ml) that was independent between time windows (see Figure 2). A random error with a SD of 3 ml gives an 8% precision at 75 ml SV. This was chosen to match the between examination variability in VTI measurements performed by the same observer (although the magnitude of this variation will only effect the results of simulation 2).²⁵ From these three simulated measurements of a 'constant' SV (but with random measurement error added), we calculated ΔSV_{100} , ΔSV_{400} and ΔSV_{500} . We also simulated a second independent SV_{100} measurement (SV_{100b}) to serve as the reference for an independent outcome measure ($\Delta SV_{400b} = (SV_{500} - SV_{100b})/SV_{100b}$). In this initial simulation, we regarded any increase in SV as a positive fluid response. Using ROC analysis, we showed how well ΔSV_{100} predicted an increase in SV with either ΔSV_{500} , ΔSV_{400} or $\Delta SV_{400b} > 0\%$ as the outcome measure. Since SV varies randomly, half of patients should be responders by this definition, and because the variation is independent between the time windows, it should

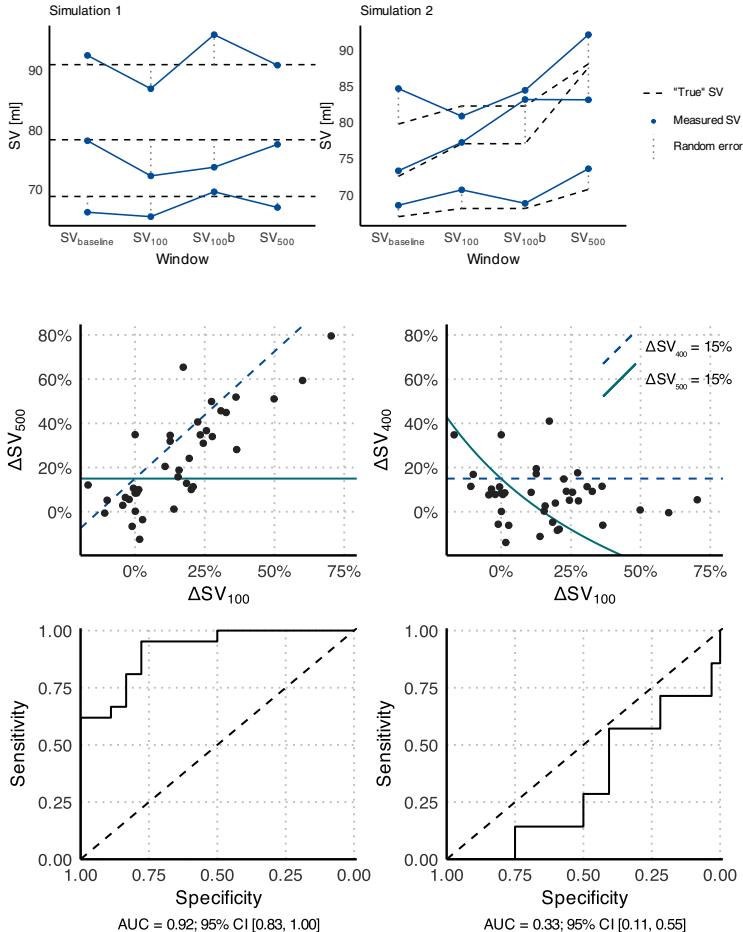


FIGURE 2 Illustration of how stroke volume (SV) measurements were simulated. Each panel shows three of 2000 simulated subjects. The dotted lines indicate the added random error at each time point

FIGURE 3 Reconstruction of data from figure 3A from Muller et al. (2011).¹ Upper panels: Scatter plots of the relation between ΔSV_{500} and ΔSV_{100} (left) and the relation between ΔSV_{400} (derived) and ΔSV_{100} (right). The full line represents the level at which ΔSV_{500} is 15% and the dashed line represents the level at which ΔSV_{400} is 15%. Lower panels: Corresponding ROC classification curves of ΔSV_{100} predicting $\Delta SV_{500} > 15\%$ and $\Delta SV_{400} > 15\%$ respectively

be impossible to predict which patients will have an increase in SV after 500 ml.

2.4.2 | Simulation 2

In a second, more realistic, simulation we simulated a 'true' response, still with additional random variation. Each subject was assigned an individual fluid response, which is the 'true' relative change from SV_{baseline} to SV_{500} (the 'true' ΔSV_{500}). The simulated fluid response was drawn from a normal distribution (mean change = 15%, SD = 10%). To keep the simulation simple, the 'true' ΔSV_{100} was defined as 30% of this 'true' ΔSV_{500} :

'True' SV_{baseline} was drawn from a normal distribution (mean = 75 ml, SD = 10 ml).

$$\text{'True' } SV_{500} = \text{'true' } SV_{\text{baseline}} \cdot (1 + \text{individual fluid response}).$$

$$\text{'True' } SV_{100} = \text{'true' } SV_{\text{baseline}} \cdot (1 + 0.3 \text{ individual fluid response}).$$

Independent random variation was subsequently added to each of these three 'true' measurements (mean = 0, SD = 3 ml) (see Figure 2). Again, we also simulated a second independent SV_{100} measurement (SV_{100b}) to serve as the reference measurement for an independent outcome measure (ΔSV_{400b}). An increase in SV of >15% was considered a significant positive fluid response in this clinical simulation.

3 | RESULTS

3.1 | Secondary analysis of an existing study

In Figure 3, plots are shown for ΔSV_{100} 's ability to predict $\Delta SV_{500} > 15\%$ (left panels) and ΔSV_{100} 's ability to predict $\Delta SV_{400} > 15\%$ (right panels). It is evident from Figure 3 that the classification goes from excellent (AUROC: 0.92) to worse than random (AUROC: 0.33) if SV_{100} is used as the reference value for the subsequent fluid response (ΔSV_{400}).

3.2 | Simulations

3.2.1 | Simulation 1

In a simulated population with no 'true' response to fluid, the commonly used MFC methodology (prediction of $\Delta SV_{500} > 0\%$ using ΔSV_{100}) predicted a fluid response with an AUROC of 0.73 (see Figure 4). Conversely, the prediction of $\Delta SV_{400} > 0\%$ (AUROC = 0.26) showed an equally large underestimation of the expected AUROC of 0.5. The independent outcome $\Delta SV_{400b} > 0\%$ was predicted by ΔSV_{100} with an AUROC of ~0.5, appropriately matching that variation in SV was random in this simulation.

3.2.2 | Simulation 2

In this simulation of a 'true' fluid response, ΔSV_{100} predicted $\Delta SV_{500} > 15\%$ with an AUROC of 0.78, and $\Delta SV_{400} > 15\%$ with an AUROC of 0.47 (see Figure 5). With a new, independent measurement after 100 ml (SV_{100b}), ΔSV_{100} predicted $\Delta SV_{400b} > 15\%$ with an AUROC of 0.65.

4 | DISCUSSION

This study demonstrates that the MFC study design most widely used in the literature (Figure 1) is problematic. Results from studies

with such problematic designs may overestimate the true classification accuracy of an MFC. This should be considered before adopting the MFC into clinical practice. Still, there are aspects of the above simulations that are worth discussing, and other study designs that should be considered in the search for the optimal MFC methodology.

4.1 | Simulations vs secondary analysis of an existing study

The simulations above were designed to illustrate only the shared error problem that arises, when the same random error is included in both predictor and outcome variables. Simulation 2 assumes a proportional relationship between the 'true' MFC response and the 'true' full response ('true' ΔSV_{100} is 30% of 'true' ΔSV_{500}). Translated into physiology, the model implies a straight Frank-Starling curve, that never plateaus. A real patient, on the other hand, can have a 'true' response to the MFC, but no 'true' response to the subsequent fluid administration, because the plateau of the Frank-Starling curve was already reached with the MFC. Indeed, in the study by Muller et al., most of the fluid response took place with the MFC, indicating that many patients were no longer fluid responsive after the MFC. But since the MFC response is also a part of the outcome (ΔSV_{500}), classification accuracy is high. This physiological circumstance (unmodelled in our simulation)

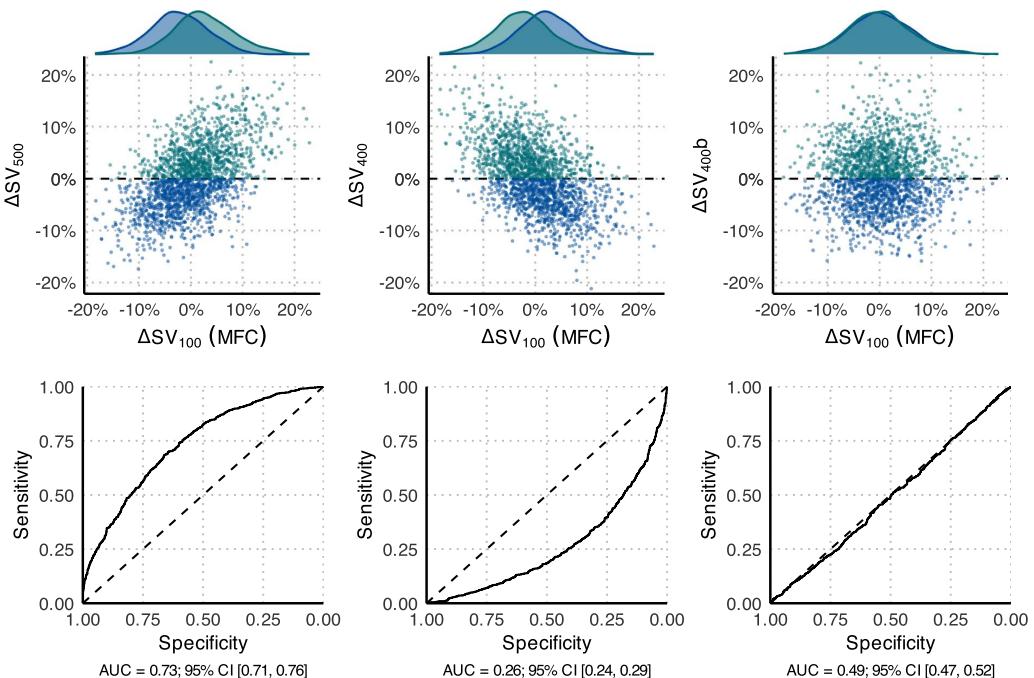


FIGURE 4 Results of simulation 1. Upper panels are scatter plots of the simulated data ($n=2000$) along with distributions of the responder and non-responder subpopulations. Lower panels are the corresponding ROC classification curves of ΔSV_{100} predicting fluid responsiveness (ΔSV_{500} , ΔSV_{400} and $\Delta SV_{400b} > 0\%$). The changes in stroke volume (ΔSV) are only random variation, so any correlation is a statistical artefact

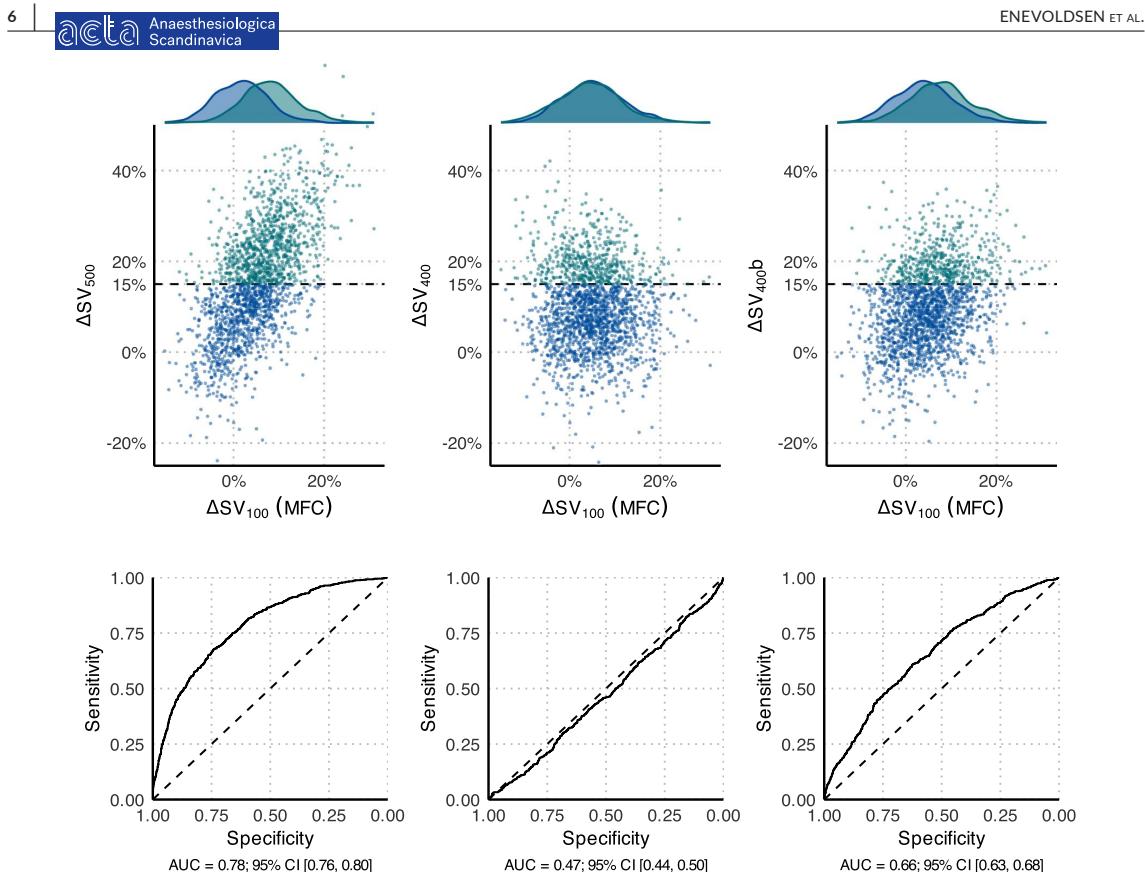


FIGURE 5 Results of simulation 2. Upper panels are scatter plots of the simulated data ($n=2000$) along with distributions of the responder and non-responder subpopulations. Lower panels are the corresponding ROC classification curves of ΔSV_{100} predicting fluid responsiveness (ΔSV_{500} , ΔSV_{400} and $\Delta SV_{400b} > 15\%$). The simulation identifies the same problem highlighted in Figure 3, although at a lower magnitude, indicating that the assumptions for the statistical modelling may be too conservative in comparison with the behaviour of real-world data

theoretically gives rise to a further overestimation of the classification accuracy in studies using the problematic MFC design, compared to the demonstrated overestimation in our simulation. The difference between predicting $\Delta SV_{500} > 15\%$ and $\Delta SV_{400} > 15\%$ is larger in the study by Muller et al than that in our simulations (see Figures 3–5). This can be explained by the combination of the shared error problem and the relatively large MFC response in the study by Muller et al. It is important to note that while ΔSV_{400} is a more clinically meaningful outcome to predict, we discourage using ΔSV_{400} as the outcome given the mathematical coupling still present due to the shared constituent value (SV_{100}). Neither of the two ROC curves in Figure 3 reveal the ‘truth’.

4.2 | Designs with different monitoring modalities for predictor and outcome variables

In one study, authors used different monitoring modalities for predictor and outcome variables: changes in pulse pressure variation

(ΔPPV) predicting fluid responsiveness (defined as change in cardiac output).⁸ This approach has the advantage that baseline measurements of PPV and thermodilution-derived cardiac output (CO_{TD}) have separate measurement errors:

$$\text{Predictor: } \Delta PPV = PPV_{100} - PPV_{\text{baseline}},$$

$$\text{Outcome: } \Delta CO_{TD} = CO_{TD,500} - CO_{TD,\text{baseline}}.$$

This reduces the concern about spurious correlation/mathematical coupling. However, while measurement errors are no longer shared, fluctuating physiology over time may still couple different haemodynamic modalities measured simultaneously. Also, this design still includes the response to the MFC in the outcome. Unlike other fluid responsiveness approaches such as the passive leg raising (PLR) manoeuvre, the MFC induces an irreversible physiologic change (because 100 ml fluid is not subsequently removed from the bloodstream).

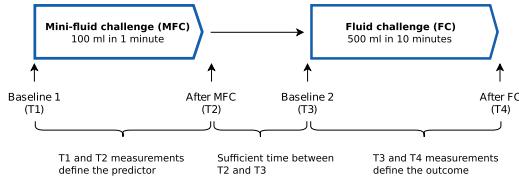


FIGURE 6 An illustration of the MFC study design used by Guinot et al.⁵ In this design, the predictor and outcome are NOT mathematically coupled

4.3 | A new reference measurement after the MFC

To date, the study with the most appropriate design is that by Guinot et al.⁵ Importantly, these authors incorporated an additional SV measurement 5 min after the MFC, to serve as reference for defining the outcome (see conceptual design in Figure 6). In that study, all four SV measurements were obtained by thoracic impedance cardiography (NICCOMO, Imedex, France). A spurious (negative) correlation could, in theory, remain, provided that the error (measurement and physiological) at T2 is correlated with the error at T3. However, it seems plausible that a 5-min window is sufficient to consider errors independent between T2 and T3. This is supported by the data, since any spurious correlation should theoretically reduce classification accuracy, which was probably not encountered in the study by Guinot et al.,⁵ reporting an AUROC of 0.93. Other monitoring modalities than NICCOMO may have data-stabilising moving-average algorithms implemented making a 5-min window insufficient. An extreme case of this is the continuous cardiac output (CCO) measurement from thermodilution pulmonary artery catheters that is only (truly) updated every 4–12 min due to a moving-average algorithm.^{26,27}

The time window between T2 and T3 is not without concern though. On average, the effect of the MFC is likely to subside during this period, making the patients more fluid responsive at T3 than at T2. Essentially, this design is using an MFC to predict the response to fluid given 5 min later. In clinical practice, the remaining fluid will likely be given immediately if the MFC response is above a certain threshold. While it may be reasonable to give the fluid right away, if the patient will respond in 5 min, this discrepancy between the study design and clinical practice should be kept in mind. This may be a necessary trade-off to avoid the statistical problems described in this paper.

4.4 | Additional considerations

Infusion rates and timing of the SV measurements can impact the results. Most MFC studies infuse the MFC in 1–2 min and the remaining fluid in 10–30 min, making the infusion rate considerably higher during the MFC.¹⁰ Prather et al show, from fluid expansions of dogs, that cardiac output returns to baseline faster than circulatory volume, and note that rapid infusion results in markedly higher peak cardiac output compared to slower infusion.²⁸ In a human study, 250 ml crystalloid was infused over 5 min and cardiac output had largely returned to baseline 10 min after end infusion.²⁹ The effect

on circulating volume is longer: it takes about 30 min before infused crystalloid is distributed between plasma and interstitial fluid, and the elimination half-life is around 20–40 min in conscious humans and several times longer during general anaesthesia.^{30,31} Because of the different infusion rates and durations, the MFC is not simply a ‘mini’ version of the full fluid infusion. It is possible that most healthy hearts will respond to a rapid fluid infusion, while some degree of hypovolaemia may be necessary for a lasting response to a slow infusion. Thus, infusion rates and timing of the SV measurements should be carefully considered in the design of an MFC study.

In most fluid responsiveness studies (incl. MFC studies), the outcome (e.g. ΔSV_{500}) is dichotomised into ‘responder’ (e.g. $\Delta SV_{500} \geq 15\%$) or ‘non-responder’. While this approach simplifies analysis and interpretation, the threshold is more-or-less arbitrary. Dichotomisation of continuous variables is generally not recommended.^{32,33} For normally distributed data, it results in a loss of power equivalent to at least a 36% reduction in sample size, and considerably more if the split is not balanced.³⁴ MFC studies, and fluid responsiveness studies in general, would benefit from keeping variables on a continuous scale.

Lastly, it may be possible to do a statistically valid analysis on data from a study with only three SV measurements (as in Figure 1). Unfortunately, we have not yet seen an example of this, nor found a satisfactory solution ourselves.

5 | CONCLUSION AND RECOMMENDATIONS

The vast majority of published MFC studies used designs that are problematic. These probably overestimate the accuracy of using MFC to guide fluid therapy.

We strongly recommend that a study design separating the predictor from the outcome is applied in the future studies. This is exemplified by the study by Guinot et al as depicted in Figure 6. Here, two separate measurements were obtained after the MFC—one to evaluate the MFC response and one to serve as a new reference for the remaining fluid infusion.

We recommend that specific attention is paid to ensure that outcome and predictor variables are indeed separated by a sufficient time window between the T2 and T3 measurements (see Figure 6). An appropriate time window will depend on the used monitoring modality and its underlying algorithms and time resolution.

Researchers should strongly consider keeping both the predictor and outcome on a continuous scale, and be cautious of spurious correlations when analysing changes.

CONFLICT OF INTEREST

TWLS received research grants and honoraria from Edwards Lifesciences (Irvine, CA, USA) and Masimo Inc. (Irvine, CA, USA) for consulting and lecturing, and from Pulsion Medical Systems SE (Feldkirchen, Germany) for lecturing. JE, JMB and STV have no conflict of interests to declare.

AUTHOR CONTRIBUTIONS

JE: Conception, manuscript preparation, simulation, artwork and revision. TWL: Conception, manuscript preparation and revision. JMB: Manuscript preparation and revision. STV: Conception, manuscript preparation, simulation, artwork and revision.

ORCID

Johannes Enevoldsen  <https://orcid.org/0000-0002-9190-6566>
 Thomas W. L. Scheeren  <https://orcid.org/0000-0002-9184-4190>
 Jonas M. Berg  <https://orcid.org/0000-0001-9056-7470>
 Simon T. Vistisen  <https://orcid.org/0000-0002-1297-1459>

REFERENCES

- Muller L, Toumi M, Bousquet P-J, et al. An increase in aortic blood flow after an infusion of 100 ml colloid over 1 minute can predict fluid responsiveness: the mini-fluid challenge study. *Anesthesiology*. 2011;115:541-547.
- Vincent JL. "Let's Give Some Fluid and See What Happens" versus the "Mini-fluid Challenge". *Anesthesiology*. 2011;115:455-456.
- Cecconi M, Hofer C, Teboul JL, et al. Fluid challenges in intensive care: the FENICE study: a global inception cohort study. *Intensive Care Med*. 2015;41:1529-1537.
- Smorenberg A, Cherpanath TGV, Geerts BF, et al. A mini-fluid challenge of 150mL predicts fluid responsiveness using Modelflow (R) pulse contour cardiac output directly after cardiac surgery. *J Clin Anesth*. 2018;46:17-22.
- Guinot P-G, Bernard E, Defrancq F, et al. Mini-fluid challenge predicts fluid responsiveness during spontaneous breathing under spinal anaesthesia: an observational study. *Eur J Anaesthesiol*. 2015;32:645-649.
- Biais M, de Courson H, Lanchon R, et al. Mini-fluid challenge of 100 ml of crystalloid predicts fluid responsiveness in the operating room. *Anesthesiology*. 2017;127:450-456.
- Xiao-ting W, Hua Z, Da-wei L, et al. Changes in end-tidal CO₂ could predict fluid responsiveness in the passive leg raising test but not in the mini-fluid challenge test: a prospective and observational study. *J Crit Care*. 2015;30:1061-1066.
- Mallat J, Meddour M, Durville E, et al. Decrease in pulse pressure and stroke volume variations after mini-fluid challenge accurately predicts fluid responsiveness. *Br J Anaesth*. 2015;115:449-456.
- Wu Y, Zhou S, Zhou Z, Liu B. A 10-second fluid challenge guided by transthoracic echocardiography can predict fluid responsiveness. *Crit Care*. 2014;18:R108.
- Messina A, Dell'Anna A, Baggiani M, et al. Functional hemodynamic tests: a systematic review and a metanalysis on the reliability of the end-expiratory occlusion test and of the mini-fluid challenge in predicting fluid responsiveness. *Crit Care*. 2019;23:264.
- Ali A, Dorman Y, Abdullah T, et al. Ability of mini-fluid challenge to predict fluid responsiveness in obese patients undergoing surgery in the prone position. *Minerva Anestesiol*. 2019;85:981-988.
- Mukhtar A, Awad M, Elayashy M, et al. Validity of mini-fluid challenge for predicting fluid responsiveness following liver transplantation. *BMC Anesthesiol*. 2019;19:56.
- Lee C-T, Lee T-S, Chiu C-T, Teng H-C, Cheng H-L, Wu C-Y. Mini-fluid challenge test predicts stroke volume and arterial pressure fluid responsiveness during spine surgery in prone position: A STARD-compliant diagnostic accuracy study. *Medicine (Baltimore)*. 2020;99:e19031.
- Fot EV, Izotova NN, Smetkin AA, Kuzkov VV, Kirov MY. Dynamic tests to predict fluid responsiveness after off-pump coronary artery bypass grafting. *J Cardiothorac Vasc Anesth*. 2020;34:926-931.
- Messina A, Lionetti G, Foti L, et al. Mini fluid chAllenge aNd End-expiratory occlusion test to assess fLUid responsiVEness in the opeRating room (MANEUVER study): a multicentre cohort study. *Eur J Anaesthesiol EJA*. 2021;38:422-431.
- Vistisen ST, Scheeren TWL. Challenge of the mini-fluid challenge: filling twice without creating a self-fulfilling prophecy design. *Anesthesiology*. 2018;128:1043-1044.
- Squa P, Scheeren TWL, Aya HD, et al. Metrology part 1: definition of quality criteria. *J Clin Monit Comput*. 2021;35:17-25.
- Squa P, Scheeren TWL, Aya HD, et al. Metrology part 2: procedures for the validation of major measurement quality criteria and measuring instrument properties. *J Clin Monit Comput*. 2021;35:27-37.
- Fugitt GV, Lieberson S. Correlation of ratios or difference scores having common terms. *Social Methodol*. 1973;5:128.
- Archie J. Mathematic coupling of data: a common source of error. *Ann Surg*. 1981;193:296-303.
- Stratton HH, Feustel PJ, Newell JC. Regression of calculated variables in the presence of shared measurement error. *J Appl Physiol*. 1987;62:2083-2093.
- R Core Team. R: A Language and Environment for Statistical Computing [Internet]. 2021 Available from: <https://www.r-project.org/>
- Robin X, Turck N, Hainard A, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*. 2011;12:77.
- Wickham H, Averick M, Bryan J, et al. Welcome to the Tidyverse. *J Open Source Softw*. 2019;4:1686.
- Jozwiak M, Mercado P, Teboul J-L, et al. What is the lowest change in cardiac output that transthoracic echocardiography can detect? *Crit Care*. 2019;23:116.
- Haller M, Zollner C, Briegel J, Forst H. Evaluation of a new continuous thermodilution cardiac output monitor in critically ill patients: a prospective criterion standard study. *Crit Care Med*. 1995;23:860-866.
- Bootsma IT, Boerma EC, Scheeren TWL, de Lange F. The contemporary pulmonary artery catheter. Part 2: measurements, limitations, and clinical applications. *J Clin Monit Comput*. 2021;1:15.
- Prather J, Taylor A, Guyton A. Effect of blood volume, mean circulatory pressure, and stress relaxation on cardiac output. *Am J Physiol-Leg Content*. 1969;216:467-472.
- Aya HD, Ster IC, Fletcher N, Grounds RM, Rhodes A, Cecconi M. Pharmacodynamic analysis of a fluid challenge. *Crit Care Med*. 2016;44:880-891.
- Hahn RG, Lyons G. The half-life of infusion fluids: an educational review. *Eur J Anaesthesiol EJA*. 2016;33:475-482.
- Hahn RG. Understanding volume kinetics. *Acta Anaesthesiol Scand*. 2020;64:570-578.
- Dawson NV, Weiss R. Dichotomizing continuous variables in statistical analysis: a practice to avoid. *Med Decis Making*. 2012;32:225-226.
- Altman DG, Royston P. The cost of dichotomising continuous variables. *BMJ*. 2006;332:1080.
- Fedorov V, Mannino F, Zhang R. Consequences of dichotomization. *Pharm Stat*. 2009;8:50-61.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Enevoldsen J, Scheeren TWL, Berg JM, Vistisen ST. Existing fluid responsiveness studies using the mini-fluid challenge may be misleading: Methodological considerations and simulations. *Acta Anaesthesiol Scand*. 2021;00:1-8. <https://doi.org/10.1111/aas.13965>

B

Paper 2

B. Paper 2

Journal of Clinical Monitoring and Computing
<https://doi.org/10.1007/s10877-022-00873-7>

ORIGINAL RESEARCH



Using generalized additive models to decompose time series and waveforms, and dissect heart–lung interaction physiology

Johannes Enevoldsen^{1,2} · Gavin L. Simpson³ · Simon T. Vistisen^{1,2}

Received: 23 March 2022 / Accepted: 2 May 2022
© The Author(s) 2022

Abstract

Common physiological time series and waveforms are composed of repeating cardiac and respiratory cycles. Often, the cardiac effect is the primary interest, but for, e.g., fluid responsiveness prediction, the respiratory effect on arterial blood pressure also convey important information. In either case, it is relevant to disentangle the two effects. Generalized additive models (GAMs) allow estimating the effect of predictors as nonlinear, smooth functions. These smooth functions can represent the cardiac and respiratory cycles' effects on a physiological signal. We demonstrate how GAMs allow a decomposition of physiological signals from mechanically ventilated subjects into separate effects of the cardiac and respiratory cycles. Two examples are presented. The first is a model of the respiratory variation in pulse pressure. The second demonstrates how a central venous pressure waveform can be decomposed into a cardiac effect, a respiratory effect and the interaction between the two cycles. Generalized additive models provide an intuitive and flexible approach to modelling the repeating, smooth, patterns common in medical monitoring data.

Keywords Hemodynamic monitoring · Central venous pressure · Mechanical ventilation · Signal processing · Statistical modelling

1 Introduction

Medical waveforms of physiological measurements, like electrocardiogram (ECG), invasive arterial blood pressure (ABP), photoplethysmogram (pleth) and central venous pressure (CVP), are ubiquitous in settings with closely monitored patients, notably in intensive care units and operating rooms. While waveforms of these signals are often displayed on a bedside monitor, they are rarely interpreted directly by the clinician (the ECG being a notable exception). Instead, simple summary characteristics, e.g. heart rate, respiratory rate and standard blood pressure features, are automatically

calculated by the bedside monitor and presented beside the waveforms.

The main signal in these waveforms comes from the heart. In addition, respiration impacts the waveform, and the cyclic respiratory effect can convey important information about patient physiology. This is especially recognised in fluid responsiveness research where “dynamic” fluid responsiveness indicators such as the pulse pressure variation (PPV) have repeatedly outperformed “static” indicators [1, 2]. However, the details of the cyclic respiratory effects can be difficult to disentangle, illustrated by the ventilation-related limitations to PPV such as tidal volume, respiratory rate and respiratory system compliance [3].

Researchers have developed several methods for analysing medical waveforms and derived time series: e.g. pulse pressure variation (PPV), cardiac output estimation, hypotension prediction index, etc. While many of these measures are useful and often implemented in commercial monitors, they do not always reflect what the clinician expects them to (e.g. a high PPV from a patient with a subtle arrhythmia). Generally, these complicated algorithms are difficult to understand and typically proprietary. This makes it difficult

✉ Johannes Enevoldsen
enevoldsen@clin.au.dk

¹ Department of Clinical Medicine, Aarhus University, Palle Juul-Jensens Boulevard 82, 8200 Aarhus N, Denmark

² Department of Anaesthesiology and Intensive Care, Aarhus University Hospital, Palle Juul-Jensens Boulevard 99, 8200 Aarhus N, Denmark

³ Department of Animal Science, Aarhus University, Tjele, Denmark

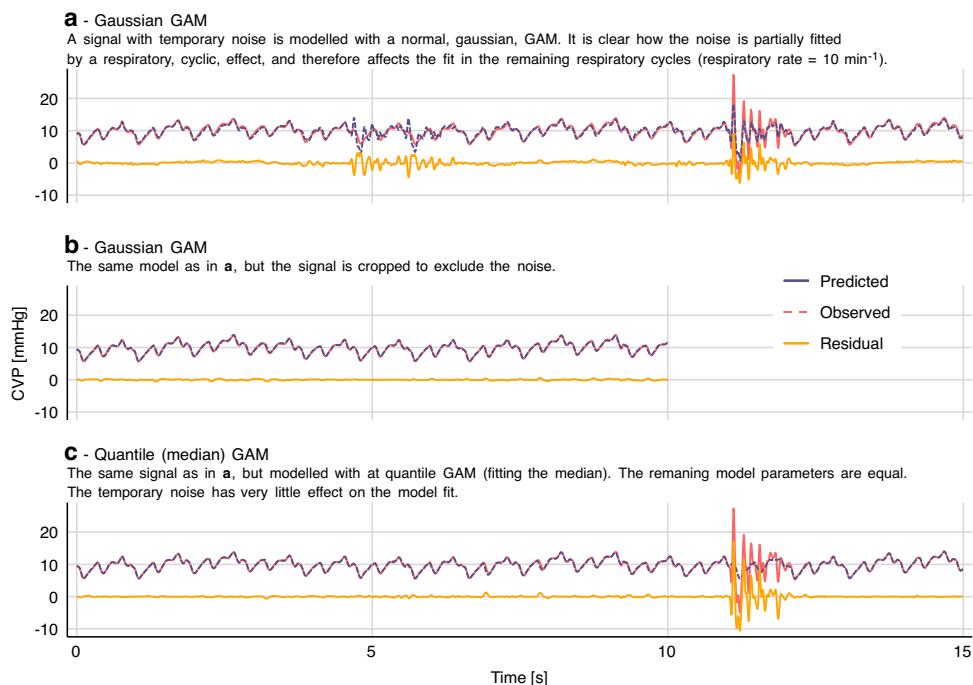


Fig. 7 Quantile generalized additive models (QGAM) robustly fit medical signals with non-normal errors. The models correspond to the model shown in Fig. 5

2.3.3 Dealing with signal noise

In the examples above, we have fitted data on the assumption that errors (residuals) are normally distributed. In practice, extreme outliers are much more common than expected from a normal distribution. Most of the time, the measured signal (e.g. CVP) will reflect the true state with very little noise. However, temporary large deflections of the waveform are common (e.g. due to manipulation of transducer or tubing). Together, these two sources of noise give rise to errors that are both non-normal and heteroscedastic (with non-constant spread). If we try to fit this data with a model that assumes homoscedastic, normally distributed errors, we will likely encounter overfitting. This is illustrated in Fig. 7a, where the noise at 12 s is also predicted one respiratory cycle earlier—a least squares regression will prioritise being a little wrong twice over being doubly wrong once (since the errors are squared).

To remedy the problem with noise having a high impact on the fit, an effective approach is to fit the median of the signal with a *quantile* GAM. When fitting the median, there

is no assumption about the conditional distribution of the dependent variable, and outliers (e.g. from noise) have a much lower impact on the model fit (see Fig. 7c). The *qgam* package by Fasiolo et al. extends *mgcv* to allow fitting quantile models [22].

3 Discussion

In this methodological paper, we demonstrate how GAMs can be used as a flexible tool for modelling cyclic medical time series and waveforms. We give two heart-lung interaction examples: The first is a specific use-case: a robust calculation of pulse pressure variation from a time series of pulse pressure measurements. The second is a demonstration of how we can use a relatively simple model to fit the CVP waveform, with very little preconception of the shape of the waveform.

3.1 Possible applications of GAMs

Currently, GAMs are research tools that may aid investigation of complex, yet deterministic patterns in medical time series and waveforms. Respiratory variation in hemodynamic variables is often just regarded as a potential source of error, sometimes dealt with by reporting only end-expiatory measurements. There may be clinically relevant information in the respiratory variation of measurements and GAMs give researchers a powerful tool for visualising and describing the effects of ventilation on their measurement of interest. It would be interesting to see GAMs like those demonstrated here for the CVP waveform and its changes during a respiratory cycle correlated to echocardiographic measurements like tricuspid annular plane systolic excursion (TAPSE) or other measures of right ventricular function. In particular, one hypothesis is that the x' descent and its dynamics during a respiratory cycle reflect right ventricular contraction against varying afterload [23]. Another CVP feature of interest is the y descent, whose magnitude is related to the rate of right ventricular filling during diastole. A large y descent has been proposed to indicate a *non-fluid-responsive heart* [24]. This hypothesis, and the respiratory variation in the y descent, could be further investigated using GAMs of CVP waveforms. CVP morphology has not had a prominent place in the scientific literature for decades, although venous return and mean systemic filling pressure are gaining more interest [25, 26]. The detailed dynamics of the CVP waveform during mechanical ventilation may reflect “upstream aspects” of venous return, mean systemic filling pressure and conditions for outflow of organs such as the kidneys. These might be elucidated by the diastolic parts of the CVP waveform.

A GAM of the arterial blood pressure waveform (and not just PPs) could give a more nuanced picture of the variation in left ventricular contraction.

As a clinical tool, estimation of PPV using a GAM could be implemented in a bedside monitor. The PPV could be presented along with a visualisation of the model fit (similar to Fig. 2c and d) for a clinician to decide if, e.g., a high PPV should be interpreted as noise or a true respiratory variation. Such interpretation, however, may require more than basic understanding of the physiologic determinants of PPV.

Another intriguing use case is that by Wyffels et al. They use a GAM to separate the seemingly random PPV from patients with atrial fibrillation into variation caused by ventilation and variation caused by the atrial fibrillation [4]. In this regard, both the respiratory component as well as the atrial fibrillation component may offer insights concerning

fluid responsiveness, because blood pressure changes induced by filling time changes (induced by extrasystoles) have also predicted fluid responsiveness with acceptable accuracy in the intensive care unit [27, 28].

3.2 Limitations

In the examples, we use synchronised data from both the ventilator and the bedside monitor. This is rarely available in data that is not recorded specifically to study heart–lung interactions. It is possible to fit these models if only the respiratory rate is known (by using the modulo operation of time over respiration length), though the phase of the respiratory effect will be arbitrary [4]. In many cases, the respiratory rate can be assessed by frequency analysis; Fourier analysis for recordings with a constant sample rate (e.g. CVP) or Lomb-Scargle analysis for irregular time series (e.g. pulse pressure).

The models presented here assume that all respiratory cycles are equivalent. This requires deeply sedated, mechanically ventilated subjects. Therefore, the models presented here are most suitable in the setting of general anaesthesia. It is possible that the models could be extended to account for spontaneous ventilation efforts, e.g., by including esophageal- or airway pressure as independent variables in the model.

The CVP model uses a non-cyclic spline to model the effect of a cardiac cycle. We expect that the CVP at the end of one cardiac cycle continues smoothly into the following cycle, but this expectation is not enforced in our model. We cannot simply use a cyclic spline, as they require a fixed cycle length, while the cardiac cycle length varies with respiration. We could use the relative position in the cardiac cycle (from 0 to 1) as the independent variable in a cyclic spline, but this assumes that the cardiac cycle effect scales linearly with cardiac cycle length (i.e. if the cardiac cycle length is 10% longer, the time from, e.g., the ‘ a peak’ to the ‘ v peak’ should be 10% longer), which is not the case. Using non-cyclic splines to model the cardiac cycle gives the model some “unnecessary” degrees of freedom, and a better solution may exist.

It can be computationally expensive to fit GAMs, especially with large, high-resolution data sets and when interaction terms are introduced. The CVP model used in Fig. 5 takes ~60 s to fit on a modern laptop, currently making it infeasible for real time implementation. The quantile model used in Fig. 7 takes ~300 s for just 15 s of signal (1875 samples). The PP model in Fig. 2 takes only ~30 ms.

4 Conclusion

Generalized additive models provide an intuitive and flexible approach to modelling the repeating signals common to medical monitoring data. We hope researchers will use this introduction as a starting point for including GAMs in their data analyses. Both to answer specific research questions, and as a tool to explore and visualise the cardiac effects and respiratory effects on hemodynamic measurements and the effect of heart–lung interactions.

5 Recommended reading

Generalized Additive Models, An Introduction with R by Simon Wood [29].

GAMs in R by Noam Ross, A Free, Interactive Course using mgcv (<https://noamross.github.io/gams-in-r-course/>).

Modelling Palaeoecological Time Series Using Generalised Additive Models [20]. An introduction to GAMs with a more detailed description of the statistical considerations related to modelling time series and the inferences that can be drawn from the models.

Hierarchical generalized additive models in ecology: an introduction with mgcv [30]. The present paper only describes models fitted to data from one individual. A relevant next step is to fit one model across multiple individuals.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10877-022-00873-7>.

Acknowledgements The authors thank Dr. John George Karippacheril for an effective collaborative effort on extending his open source software, VitalSignCapture, to support exporting real time data from Drager ventilators.

Author contributions JE: Conception of study, design of data analyses, data collection, data analysis, interpretation of data and results, writing up of the first draft of the paper, revising the manuscript for important intellectual content, final approval of the version to be published.GLS: Design of data analysis, interpretation of data and results, revising the manuscript for important intellectual content, final approval of the version to be published.STV: Conception of study, interpretation of data and results, revising the manuscript for important intellectual content, final approval of the version to be published.

Funding JE is supported by Aarhus University and *Holger & Ruth Hesse's Mindefond*. GLS is supported by an Aarhus University Research Foundation Starting Grant.

Declarations

Conflict of interest STV is associate editor of Journal of Clinical Monitoring and Computing. JE and GLS report no competing interests.

Ethical approval Data was recorded as part of a project registered on ClinicalTrials.gov, NCT04298931 with regional ethical committee approval, case: 1-10-72-245-19.

Informed consent All participants provided written informed consent.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Marik PE, Cavallazzi R, Vasu T, Hirani A. Dynamic changes in arterial waveform derived variables and fluid responsiveness in mechanically ventilated patients: a systematic review of the literature. Crit Care Med. 2009;37:2642–7. <https://doi.org/10.1097/CCM.0b013e3181a590da>.
2. Guerin L, Monnet X, Teboul J-L. Monitoring volume and fluid responsiveness: from static to dynamic indicators. Best Pract Res Clin Anaesthesiol. 2013;27:177–85. <https://doi.org/10.1016/j.bpa.2013.06.002>.
3. Michard F, Chemla D, Teboul J-L. Applicability of pulse pressure variation: how many shades of grey? Crit Care. 2015;19:15–7. <https://doi.org/10.1186/s13054-015-0869-x>.
4. Wyffels PAH, De Hert S, Wouters PF. New algorithm to quantify cardiopulmonary interaction in patients with atrial fibrillation: a proof-of-concept study. Br J Anaesth. 2021;126:111–9. <https://doi.org/10.1016/j.bja.2020.09.039>.
5. Hastie T, Tibshirani R. Generalized additive models. Stat Sci Inst Math Stat. 1986;1:297–318.
6. Hastie T, Tibshirani R, Friedman JH. The elements of statistical learning: data mining, inference, and prediction. New York: Springer; 2009.
7. Wood SN. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. J R Stat Soc Ser B. 2011;73:3–36.
8. R Core Team. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2021.
9. Simpson GL. gratia: graceful ggplot-based graphics and other functions for GAMs fitted using mgcv; 2022. <https://gavinsimpsongithub.io/gratia/>.
10. Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, et al. Welcome to the tidyverse. J Open Source Softw. 2019;4:1686. <https://doi.org/10.21105/joss.01686>.
11. Lee H-C, Jung C-W. Vital Recorder—a free research tool for automatic recording of high-resolution time-synchronised physiological data from multiple anaesthesia devices. Sci Rep. 2018;8:1527. <https://doi.org/10.1038/s41598-018-20062-4>.
12. Karippacheril JG, Ho TY. Data acquisition from S/5 GE datex anesthesia monitor using VSCapture: an open source.NET/

B. Paper 2

Journal of Clinical Monitoring and Computing

- mono tool. *J Anaesthesiol Clin Pharmacol.* 2013;29:423–4. <https://doi.org/10.4103/0970-9185.117096>.
- 13. Michard F. Changes in arterial pressure during mechanical ventilation. *Anesthesiology.* 2005;103:419–28. <https://doi.org/10.1097/000000542-200508000-00026>.
 - 14. Aboy M, McNames J, Thong T, Phillips CR, Ellenby MS, Goldstein B. A novel algorithm to estimate the pulse pressure variation index/spl Delta/PP. *IEEE Trans Biomed Eng.* 2004;51:2198–203. <https://doi.org/10.1109/TBME.2004.834295>.
 - 15. Cannesson M, Slieker J, Desebbe O, Bauer C, Chiari P, Hénaine R, et al. The ability of a novel algorithm for automatic estimation of the respiratory variations in arterial pulse pressure to monitor fluid responsiveness in the operating room. *Anesth Analg.* 2008;106:1195–200. <https://doi.org/10.1213/01.ane.0000297291.01615.5c>.
 - 16. Vistisen ST, Struijk JJ, Larsson A. Automated pre-ejection period variation indexed to tidal volume predicts fluid responsiveness after cardiac surgery. *Acta Anaesthesiol Scand.* 2009;53:534–42. <https://doi.org/10.1111/j.1399-6576.2008.01893.x>.
 - 17. De Backer D, Ph D, Taccone FS, Holsten R, Ibrahim F, Vincent J, et al. Influence of respiratory rate on stroke volume variation in mechanically ventilated patients. *Anesthesiology.* 2009;110:1092–7. <https://doi.org/10.1097/ALN.0b013e31819db2a1>.
 - 18. Mackenzie J. The interpretation of the pulsations in the jugular veins. *Am J Med Sci.* 1907;134:12–34. <https://doi.org/10.1097/0000441-190707000-00002>.
 - 19. Constant J. The X prime descent in jugular contour nomenclature and recognition. *Am Heart J.* 1974;88:372–9. [https://doi.org/10.1016/0002-8703\(74\)90474-8](https://doi.org/10.1016/0002-8703(74)90474-8).
 - 20. Simpson GL. Modelling palaeoecological time series using generalised additive models. *Front Ecol Evol.* 2018. <https://doi.org/10.3389/fevo.2018.00149>.
 - 21. van Rij J, Hendriks P, van Rijn H, Baayen RH, Wood SN. Analyzing the time course of pupillometric data. *Trends Hear.* 2019. <https://doi.org/10.1177/2331216519832483>.
 - 22. Fasiolo M, Wood SN, Zaffran M, Nedellec R, Goude Y. Fast calibrated additive quantile regression. *J Am Stat Assoc.* 2020. <https://doi.org/10.1080/01621459.2020.1725521>.
 - 23. Raut MS, Maheshwari A. “x” descent of CVP: an indirect measure of RV dysfunction ? *J Anaesthesiol Clin Pharmacol.* 2014;30:430–1. <https://doi.org/10.4103/0970-9185.137289>.
 - 24. Magder S, Erice F, Lagonidis D. Determinants of the Y descent and its usefulness as a predictor of ventricular filling. *J Intensive Care Med.* 2000;15:262–9. <https://doi.org/10.1177/08850666000150005>.
 - 25. de Keijzer IN, Scheeren TWL. Perioperative hemodynamic monitoring: an overview of current methods. *Anesthesiol Clin.* 2021;39:441–56. <https://doi.org/10.1016/j.anclin.2021.03.007>.
 - 26. Meijis LPB, van Houtte J, Conjaerts BCM, Bindels AJGH, Bouwman A, Houterman S, et al. Clinical validation of a computerized algorithm to determine mean systemic filling pressure. *J Clin Monit Comput.* 2021. <https://doi.org/10.1007/s10877-020-00636-2>.
 - 27. Vistisen ST, Krog MB, Elkmann T, Vallentin MF, Scheeren TWL, Sølling C. Extrasystoles for fluid responsiveness prediction in critically ill patients. *J Intensive Care.* 2018;6:52. <https://doi.org/10.1186/s40560-018-0324-6>.
 - 28. Vistisen ST. Using extra systoles to predict fluid responsiveness in cardiothoracic critical care patients. *J Clin Monit Comput.* 2017;31:693–9. <https://doi.org/10.1007/s10877-016-9907-8>.
 - 29. Wood SN. Generalized additive models: an introduction with R. London: Chapman and Hall; 2017. <https://doi.org/10.1201/9781315370279>.
 - 30. Pedersen EJ, Miller DL, Simpson GL, Ross N. Hierarchical generalized additive models in ecology: an introduction with mgcv. *PeerJ.* 2019;7:e6876. <https://doi.org/10.7717/peerj.6876>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

for the clinician to critically consider the algorithm's analysis of the waveform.

The task of analysing physiological data both comprehensively and transparently seems a perfect fit for generalized additive models (GAMs). A recent paper by Wyffels et al. demonstrates how GAMs can be used to isolate the respiratory component of PPV in subjects with atrial fibrillation [4]. An elegant solution that may be used to guide fluid therapy in this patient group.

The aim of this paper is to demonstrate how GAMs can be used to decompose waveforms or time series recorded in mechanically ventilated patients into separate, physiologically relevant, components. This allows analysts to focus on each component individually. We give a short introduction to splines and GAMs, and then demonstrate the method using two examples. First, we use a time series of pulse pressure measurements to give a robust estimate of PPV in mechanically ventilated patients with sinus rhythm (a simplified version of the model presented by Wyffels et al. [4]). Second, we decompose the CVP waveform into separate, physiologically relevant, effects. Finally, we summarise and discuss how GAMs might be used in future research and in clinical monitoring.

1.1 What is a GAM?

Generalized additive models are both flexible and interpretable. In the space of statistical models, they reside somewhere between simple but rigid methods like linear regression and flexible but complex methods like neural networks. With GAMs, we can build transparent models, with components that represent known physiology.

Hastie and Tibshirani introduced GAMs in 1986, as extensions of generalized linear models [5]. Instead of fitting straight lines, GAMs can fit any smooth function. In the basic form of a GAM, a smooth function is fitted for each independent variable in the model. These functions are added together to give the model's prediction of the dependent variable:

$$Y_{predicted} = \alpha + f(X_1) + f(X_2),$$

where α is a constant value and f can be any smooth function (continuous and with no kinks). In this paper, we do not introduce link functions, and we mainly use models with a Gaussian conditional distribution.

1.1.1 Cubic splines

Several types of smooth functions can be used to fit data. In this paper, we use one type: the *cubic spline*. A cubic spline is built by combining a number of third-order polynomials. Each polynomial fits its individual section of the data (e.g., a period of time if time is the independent variable) and is constrained to join smoothly to the adjacent polynomial(s). The intersections between adjacent polynomials are called *knots*. Smoothness at the knots is ensured by constraining adjacent polynomials to align at the knots. Specifically, the values of adjacent cubic polynomials' 0th, 1st and 2nd derivatives must be equal at the knots. The knots can be placed at will, but a common choice is to position knots at the quantiles (including at minimum and maximum) of the independent variable, giving the same number of observations in each segment (see Fig. 1a). Cubic splines are often additionally constrained by fixing the second and third derivative at the

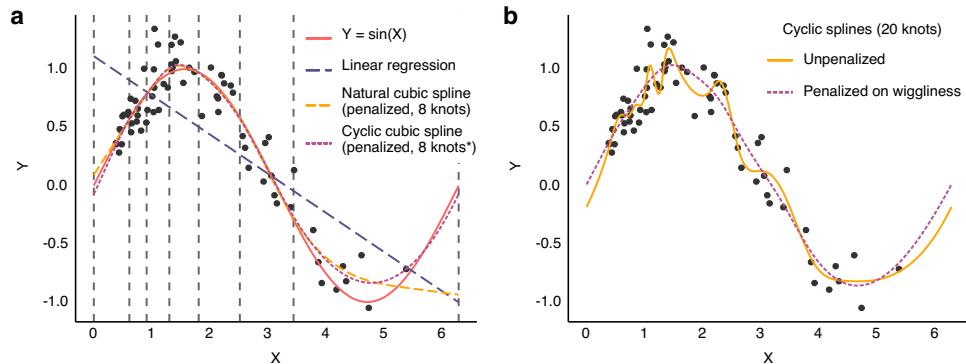


Fig. 1 Splines fitted to simulated data ($n=70$). The data-generating function is $Y = \sin(X)$ with added normally distributed noise. **a** Vertical dashed lines show the position of the 8 knots. *In the cyclic spline there are effectively 7 knots, since the first and last line represent a single knot, joining the ends of the spline. **b** Comparison of a penal-

ised and an unpenalised spline fitted to the same data. The unpenalised spline with 20 knots is clearly too wiggly and overfits the data. Penalising the spline on wigginess reduces the risk of overfitting, but keeps the model flexible in case the data demand it

outer knots to zero (making them linear outside the outer knots). This is termed a *natural cubic spline* [6].

To reduce the risk of overfitting, splines can be *penalised* according to their wigginess (by default defined as the integral of the squared 2nd derivative). A penalised spline is fitted to optimise the tradeoff between goodness of fit (e.g. high likelihood) and complexity (measured by the wigginess of the function) (see Fig. 1b). The relative weight of fit and wigginess in this tradeoff is controlled with a *smoothing parameter*. This smoothing parameter can be automatically optimised to prevent overfitting (e.g. using a *restricted maximum likelihood* approach [7]) or be chosen manually. A manual smoothing parameter can be useful if there is prior knowledge about the smoothness of one or more splines in the model (e.g. the effect of ventilation is expected to be very smooth).

1.1.2 Modelling interaction between variables

Interaction terms can be included in two principal ways. In the simplest case, one term is continuous (X_1) and one is categorical (X_2). Individual smooth functions are then fit for each category [$f(X_1)$ for each X_2]. If both terms are continuous, the interaction can be represented as $f(X_1, X_2)$: a function that takes two values and returns one value. This can be visualised as a smooth plane where each combination of X_1 and X_2 corresponds to an output (the elevation of the plane) (see Fig. 5e.1).

1.1.3 Modelling cyclic data

Some variables repeat cyclically without a marked distinction between the end of one cycle and the beginning of the next. An example is compass direction, where $0^\circ \equiv 360^\circ$. Likewise, we expect CVP at the end of one respiratory cycle to continue smoothly into the next cycle. We can model the effect of a cyclic variable with a *cyclic cubic spline*. A cyclic cubic spline is a special case of the cubic spline where the first and last knot are treated as one. The beginning and end are effectively adjacent, and the respective splines match up to the 2nd derivative (see Fig. 1a).

2 Examples

Examples are analysed using R 4.1.0 [8] with packages: *mgcv* 1.8–36 [7], *gratia* [9] and *tidyverse* [10]. While the paper aims to be language agnostic, sample data and annotated R code are supplied in Online Resource 1 (<https://doi.org/10.5281/zenodo.6375221>).

2.1 Example data

The data for these demonstrations are recorded during abdominal surgery from three consenting patients on pressure control ventilation (recorded as part of a project registered on ClinicalTrials.gov, NCT04298931 with regional ethical committee approval, case: 1-10-72-245-19). Haemodynamic waveforms (125 Hz) were recorded from a Philips MX550 using Vital Recorder [11] and ventilator data (timestamps for each inspiration start) were recorded from a Dräger Perseus A100 using VSCaptureDrgVent [12].

2.2 Example 1: Pulse pressure

In recent years, more complex waveform analysis is being implemented in the monitors. One example is ventilator-induced pulse pressure variation (PPV): a measure commonly used to predict fluid responsiveness [13]. While it is possible to manually calculate PPV from an arterial pressure waveform, it is neither trivial nor reproducible. Also, manually calculated PPV may differ substantially from the PPV automatically calculated by the monitor. This is due to a sophisticated analysis of the arterial waveform that takes multiple respiratory cycles into account [14, 15]. The PPV calculated automatically by, e.g., Philips monitors is robust to noise and outliers [14], but the steps between the ABP waveform and the automatically calculated PPV are probably unclear to most clinicians.

In the individual, pulse pressure (PP = systolic pressure – diastolic pressure) is highly correlated with stroke volume; and like stroke volume, PP varies between heart beats. The main cause of the short-term variation in PP is respiration, and the effect is especially pronounced during controlled mechanical ventilation. A beat's position in the respiratory cycle is associated with a specific effect on PP (see Fig. 2c). Around the end of the inspiration, PP is above average; and during expiration, it drops below average (the phase depends on respiratory cycle length).

Variation in pulse pressure (PP) can be understood as the sum of three separate effects. First, the effect of ventilation: with each breath, PP rises and then decreases. This is caused by the breath's combined effect of both preload and afterload on both ventricles [13]. It is the size of this effect that is related to the response to fluid therapy. Second, PP varies over longer periods, e.g. with changes in vascular tone. Third, there is also a fast, effectively random, variation in PP: e.g. measurement noise and subtle ‘random’ fluctuations in cardiac contractility). This decomposition of PP into three separate effects can be described with the equation:

$$PP = \alpha + f(pos_{ventilationcycle}) + f(time) + \epsilon.$$

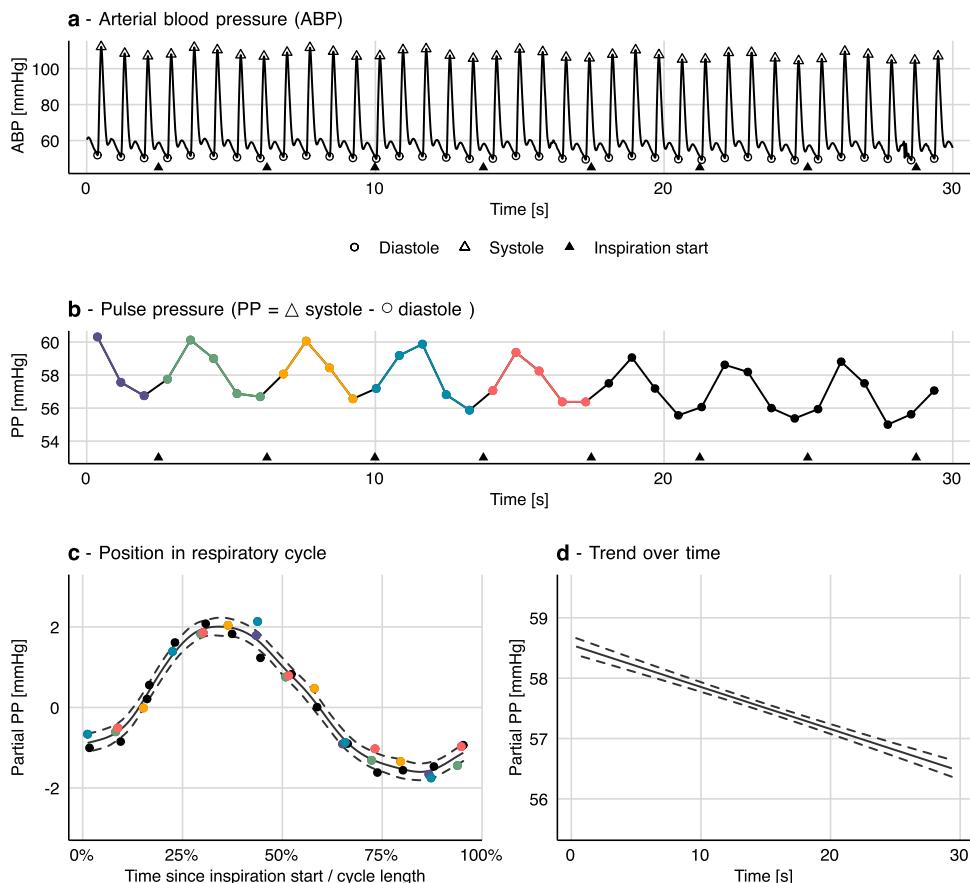


Fig. 2 How a generalized additive model (GAM) can be fitted to a series of pulse pressure measurements (derived from the arterial waveform). **a** and **b** For each beat, systolic and diastolic pressure are detected, and pulse pressure (PP) is calculated. A GAM with two smooths **c** and **d** is fitted to the PP time series (**b**). **c** This first smooth represents the variation in PP explained by the beats' position in the respiratory cycle. Coloured points (beats) correspond between panels

b and **c**. **d** The second smooth represents the trend in PP over time with the model constant (α) added. The sum of these two smooths (**b** and **c**) gives the model prediction. Residuals of the model (e) are the vertical distance from the smooth to the points in panel **c** (i.e. the scatters are partial residuals). Dashed curves represent 95% confidence intervals

$f(pos_{ventilationcycle})$ describes the relationship between a heart beat's position in the respiratory cycle and the produced PP at that heart beat. $f(time)$ represents the trend in PP over time, and α is the mean PP over the entire sample. e represents the remainder: noise, 'random' fluctuation, etc.

The individual observations in this analysis are heart beats. For each heart beat, we need to know the time it occurred, its position in the respiratory cycle (time since the start of the latest inspiration/respiratory cycle length) and the pulse pressure of the beat. The timing of each beat was assigned the time of the diastole,¹ and pulse pressure was calculated as systolic minus diastolic pressure (see Fig. 2a

and b). With this data, the model can be fitted as a GAM where $f(pos_{ventilationcycle})$ is a cyclic cubic spline and $f(time)$ is a natural cubic spline.

¹ Alternatively, QRS-complexes from the ECG could be used to mark the time of each heart beat. Pulse transit time is around 200 ms and it varies approximately 10–20 ms with ventilation [16]. Therefore, using QRS-complexes to time each heart beat would create a slight leftwards phase shift of the respiratory cycle smooth (Fig. 2c) and a probably unnoticeable effect of the variation in pulse transit time. For patients with cardiac arrhythmia, using QRS-complexes could aid the analysis. Both because it may be difficult to identify individual heart beats from the ABP waveform alone, and because pulse transit time might vary significantly between beats.

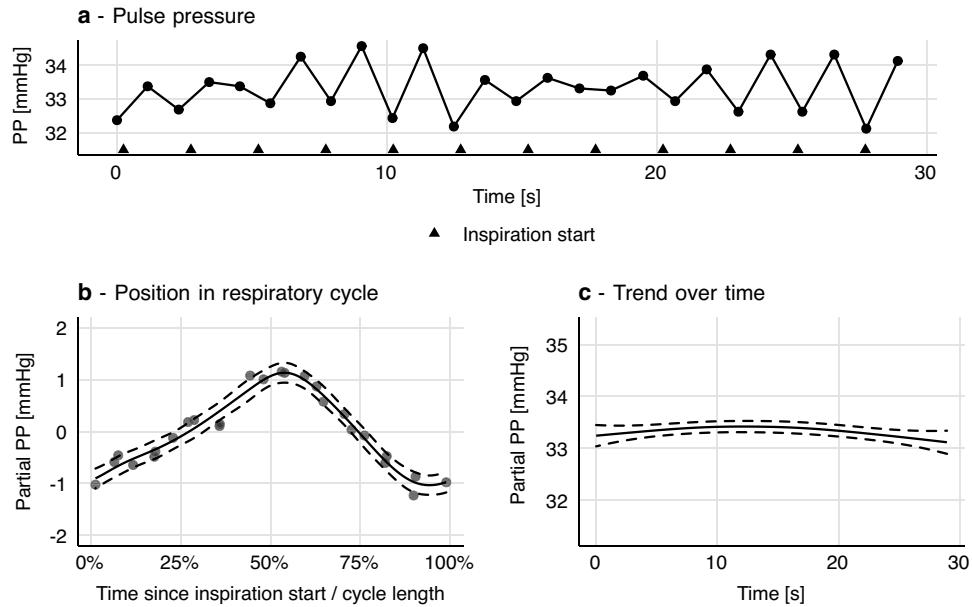


Fig. 3 This patient has a heart-rate-to-respiratory rate ratio just beyond 2:1 (52:24). **a** From the pulse pressure (PP) plot, it is difficult to assess pulse pressure variability (PPV), and it seems to be chang-

ing. **b** When PP is modelled as a smooth function of each beat's position in the respiratory cycle, a tight relationship between respiration and PP is revealed. Dashed curves represent 95% confidence intervals

After fitting the model, we can inspect the model by plotting the smooth functions over a relevant interval (usually the interval containing the original observations) (see Fig. 2c and d). In our model of pulse pressure, $f(pos_{ventilationcycle})$ represents the variation in pulse pressure with each respiratory cycle. Thus, we can use $f(pos_{ventilationcycle})$ to calculate PPV.

$$PPV = \frac{\max(f(pos_{ventilationcycle})) - \min(f(pos_{ventilationcycle}))}{\alpha},$$

where $pos_{ventilationcycle}$ is between 0 and 100%.

Since α is the mean PP, this is equivalent to the classic formula for PPV:

$$PPV = \frac{PP_{max} - PP_{min}}{(PP_{max} + PP_{min})/2}.$$

Calculation of a confidence interval for PPV is described in Online Resource 1.

Essentially, a GAM facilitates the “step” from panel b to panel c in Fig. 2, where the highly deterministic effect of heart–lung interactions on pulse pressure is uncovered. Calculating PPV from a GAM model takes every beat in our sample into account. This makes the PPV estimate less sensitive to outliers (min and max being inherently very sensitive to outliers). Also, PPV estimated from individual

respiratory cycles will tend to be lower than PPV calculated from a GAM, by a somewhat random amount. Heart beats occur at varying positions in the respiratory cycle; often not at the positions giving both the maximum and minimum pulse pressure. This is especially important in conditions with few beats per ventilation [17] (see Fig. 3). Details about the shape and phase of $f(pos_{ventilationcycle})$ may also contain important information about the heart–lung interaction, though this has not yet been investigated.

2.3 Example 2: Central venous pressure

Hemodynamic waveforms are affected by both the heart and the lungs. The CVP waveform has a fast period with the length of one cardiac cycle and a slower period with the length of one respiratory cycle. For each cardiac cycle, well-defined features represent atrial contraction (*a*), tricuspid valve closing (*c*), ventricular contraction (*x'*), atrial filling during ventricular systole (*v*) and tricuspid valve opening (*y*) [18, 19] (CVP landmarks are shown in Fig. 4b). If the patient is on a ventilator, the entire CVP waveform will rise with the inspiration and fall with the expiration (see Fig. 4a). A third effect is the interaction between the cardiac cycle and the respiratory cycle. A cardiac cycle during inspiration produces a CVP waveform that is different from what is produced during expiration. Lastly, a number of factors

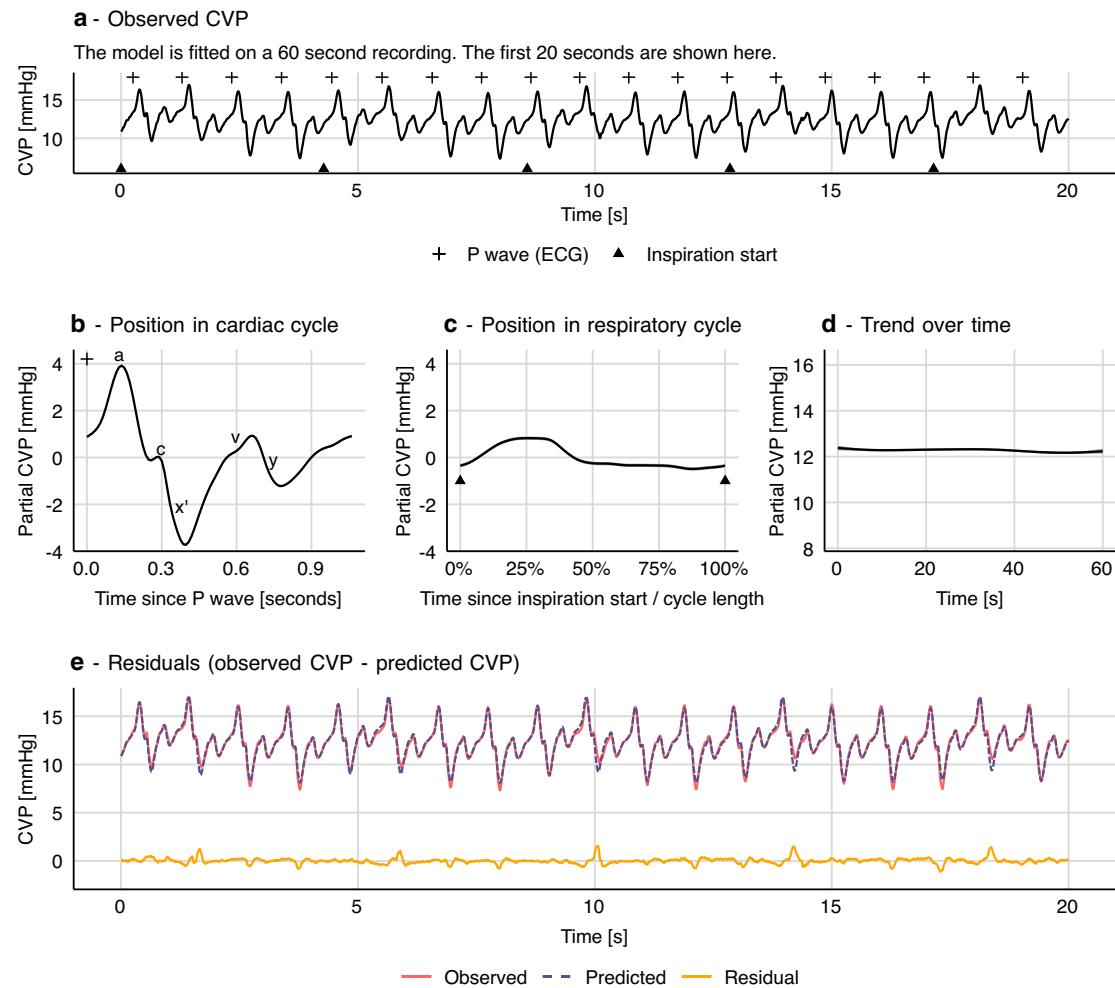


Fig. 4 Generalized additive model of central venous pressure (CVP). Variation in CVP is explained by the effects of the cardiac cycle and the respiratory cycle. In this model there is no interaction between the

two effects. Grey shades in **b**, **c** and **d** represent 95% confidence intervals (often too narrow to be visible)

influence CVP and can change over longer periods. These include, but are not limited to: surgical activity, autonomic regulation and medication.

In this example, we model the entire waveform; not just a time series of derived measurements as in the above example with pulse pressure. The unit observations are individual samples of a 125 Hz CVP recording. Each sample has a value (CVP) and a time. Using this sample time, the timing of P waves from the ECG and the timing of each inspiration start, we can compute two additional features: the sample's position in the cardiac cycle (time since the latest P wave) and its position in the respiratory cycle (similar to example

1). Timing of P waves was calculated by subtracting a constant, manually measured, PR interval from algorithmically determined QRS complex timings. The exact length of the subtracted interval is not very important. It simply ensures that the atrial contraction is placed in the beginning of a cardiac cycle rather than in the end of the previous cycle. We model the effect of the cardiac cycle with a non-cyclic spline, since cardiac cycles vary slightly in length (due to respiratory sinus arrhythmia).

A first approach to modelling CVP from these three features could be a simple extension of the PP model proposed in example 1:

$$CVP = \alpha + f(pos_{cardiac}) + f(pos_{ventilation}) + f(time) + \epsilon$$

This is a strictly additive model and therefore assumes no interaction between the effect of ventilation and heart beat on CVP; i.e. this model assumes that every heartbeat produces the same CVP pattern. This pattern is simply raised and lowered with ventilation (see Fig. 4).

This model describes most of the variation in CVP, but the depth of the x' descent (corresponding to the ventricular contraction) is systematically off at specific places in the respiratory cycle. Clearly, the pattern in CVP produced by a heart beat depends on its position in the respiratory cycle. To address this, we introduce a smooth interaction term to the model.

$$\begin{aligned} CVP = & \alpha + f(pos_{cardiac}) + f(pos_{ventilation}) \\ & + f(pos_{cardiac}, pos_{ventilation}) + f(time) + \epsilon. \end{aligned}$$

$f(pos_{cardiac}, pos_{ventilation})$ is a smooth function that represents the interaction between the cardiac and respiratory cycles. It is based on a non-cyclic spline in the x -direction (cardiac cycle) and a cyclic spline in the y -direction (respiratory cycle). It can be visualised as a surface (or more specifically, a cylinder, since it is cyclic in the Y direction), where the x -axis represents the cardiac cycle, the y -axis represents the respiratory cycle, and the z -axis represents the effect of the interaction on CVP (see Fig. 5).

To aid comprehension of the model—CVP as the interaction of two repeating cycles—we attach an animation of the model's prediction, simultaneously on a time scale and projected onto a plane with cardiac cycle position and respiratory cycle position as independent variables (see Online Resource 2). The plane is equivalent to the contour plot in Fig. 6b, before 250 ml fluid.

2.3.1 Autocorrelation

Like other regression models, a GAM assumes that observations are independent, conditional on the model (i.e. that the residuals are independent). First, if there is some pattern remaining in the residuals, it is important to consider that the model may have underfitted the data (as in the example without an interaction term; shown in Fig. 4). But, even with an “optimal” fit, models of high resolution waveforms will likely have a high degree of autocorrelation in the residuals, as noise itself is often autocorrelated in these waveforms. To correct for this, we have included in the CVP models a first-order autoregressive model [AR(1)] for the residuals (see Online Resource 1 for details). Failure to deal with

autocorrelation will give too narrow confidence intervals and can cause overfitting [20, 21].

2.3.2 How the CVP waveform changes after a fluid bolus

To illustrate the type of responses that can be estimated, we fitted a GAM to two one-minute sections of a CVP recording: the first section before administration of a 250 ml fluid bolus and the other after. Separate splines were fitted to each section:

$$\begin{aligned} CVP = & \alpha + \beta_s + f_s(pos_{cardiac}) + f_s(pos_{ventilation}) \\ & + f_s(pos_{cardiac}, pos_{ventilation}) + f_s(time) + \epsilon, \end{aligned}$$

where β_s is an additional constant, that is zero for the pre-fluid section, and f_s is a spline for each section of data (before or after 250 ml fluid). This model also extends the previous model (Fig. 5) by using an *adaptive* smoothing spline to estimate $f_s(pos_{ventilation})$. An adaptive smoothing spline allows the spline's smoothing parameter to vary across the range of the independent variable. This allows the spline to adapt to the sharp transition between inspiration and expiration, and to fit a subtle disturbance at the beginning of the expiration² while remaining smooth in areas where there is no change in the effect of the independent variable on the response. The model is visualised in Fig. 6. We see that after fluid, this subject's CVP varies more over a cardiac cycle, but less over a respiratory cycle, compared to before fluid. This is clearest in Fig. 6d. In Fig. 6c, we show the predicted CVP at end expiration and at end inspiration before and after fluid. This lets us compare how the interaction between the cardiac cycle and the respiratory cycle changes with fluid administration. The pressure during atrial contraction (a wave in Fig. 6c) increases with fluid, but the effect of ventilation on this pressure is lower after fluid. Another interesting difference is the shape of the v wave, representing the pressure in the right atrium before the tricuspid valve opening. Before fluid, the v wave has a flat peak, but after fluid, it increases gradually and reaches a higher pressure. This difference disappears at end-inspiration.

² The small disturbance at the beginning of the inspiration corresponds to the closing of the ventilator solenoid valve at end-inspiration. The sudden drop in pressure makes the ventilator tubing move and disturb the adjacent CVP line. It is most visible in Fig. 6d, but can also be recognized in the residuals in Fig. 5f.

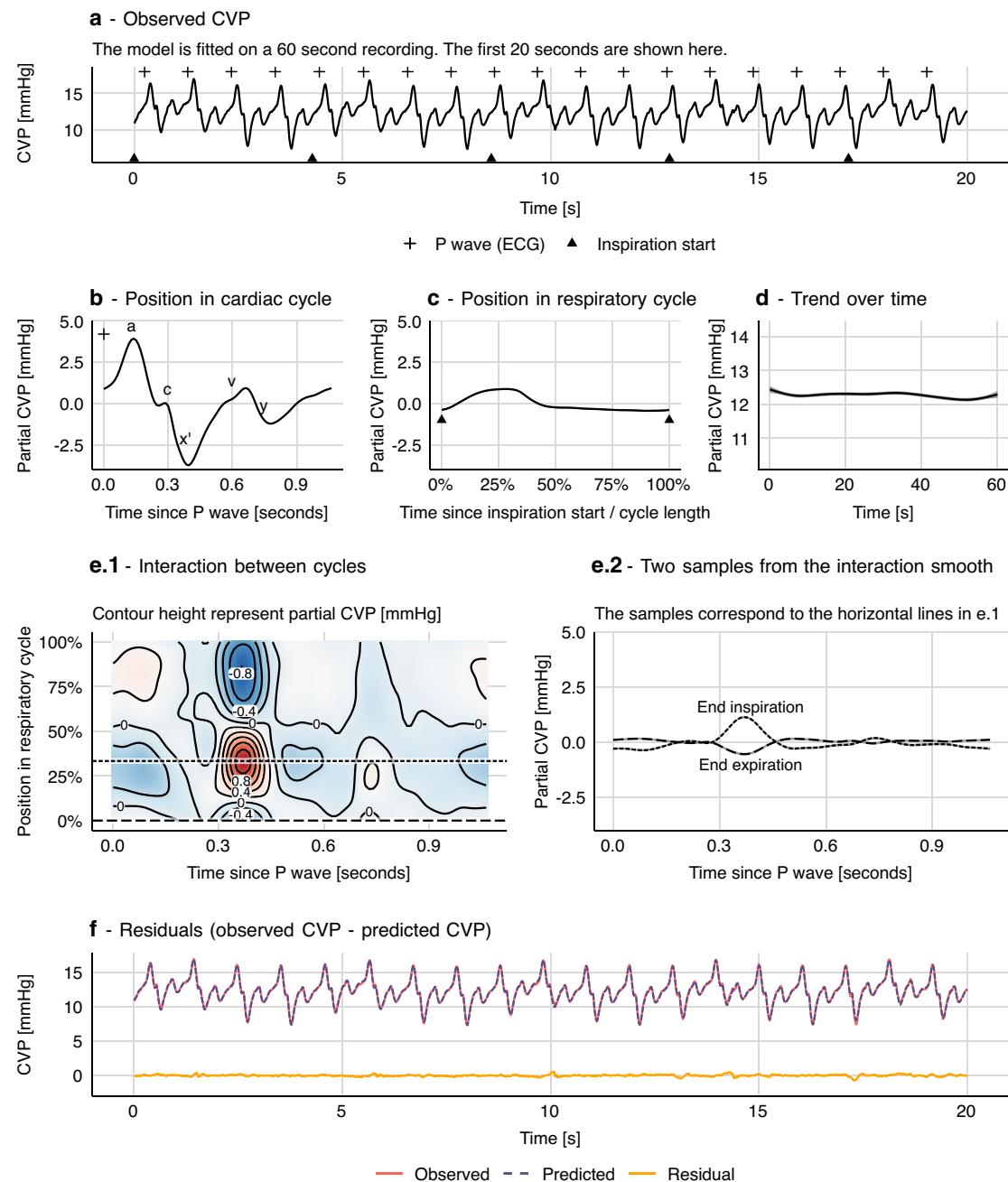


Fig. 5 How a generalized additive model (GAM) can be fitted to a CVP waveform. **a** Each sample from a 125 Hz CVP waveform is represented with three predictor variables: position in cardiac cycle, position in respiratory cycle and time (seconds since sample start). A GAM is fitted giving the smooth functions **b** to **e** (the model con-

stant (α) is added to the smooth function in **d**. **f** Model fit including residuals that are markedly reduced compared to the model without an interaction term, visualised in Fig. 4. Grey shades in panel b, c and e represent 95% confidence intervals (often too narrow to be visible)

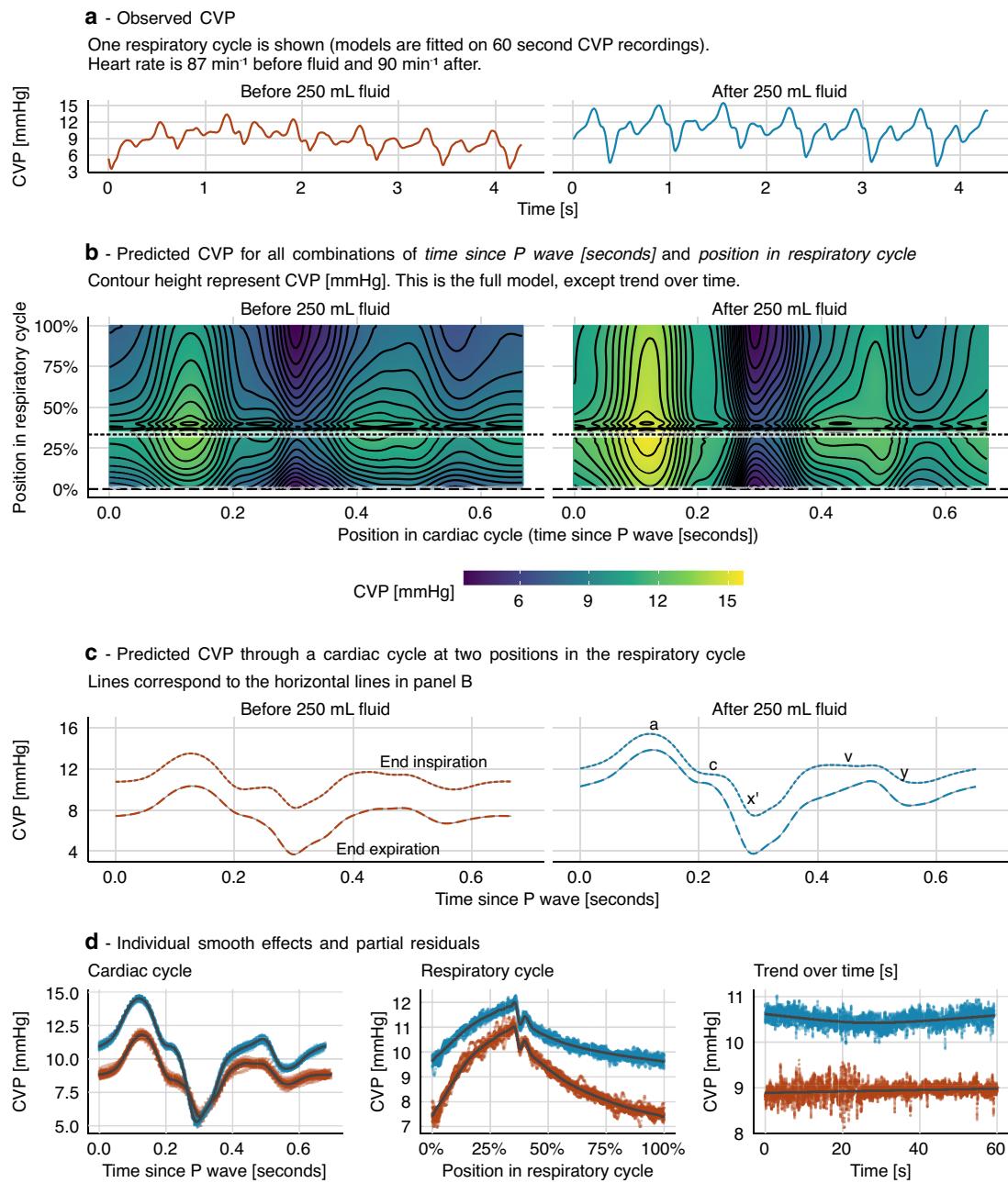


Fig. 6 Comparison of sections of a generalized additive model (GAM). The model is fitted to two one-minute sections of central venous pressure (CVP) recordings. One section before a 250 ml fluid

infusion and one after. A grey shade in panel c represents 95% confidence intervals (often too narrow to be visible). The constant terms (α and β_s) are included in all predictions (b, c and d)

C

Paper 3

References

- [1] Gautam Bhave and Eric G. Neilson. 2011. Volume Depletion Versus Dehydration: How Understanding the Difference Can Guide Therapy. *American Journal of Kidney Diseases* 58, 2 (August 2011), 302–309. DOI:<https://doi.org/10.1053/j.ajkd.2011.02.395>
- [2] A.j. Coe and B. Revanäs. 1990. Is crystalloid preloading useful in spinal anaesthesia in the elderly? *Anaesthesia* 45, 3 (1990), 241–243. DOI:<https://doi.org/10.1111/j.1365-2044.1990.tb14696.x>
- [3] J. E. Cosnett. 1989. The origins of intravenous fluid therapy. *Lancet (London, England)* 1, 8641 (April 1989), 768–771. DOI:[https://doi.org/10.1016/s0140-6736\(89\)92583-x](https://doi.org/10.1016/s0140-6736(89)92583-x)
- [4] B. A. Foëx. 2003. How the cholera epidemic of 1831 resulted in a new technique for fluid resuscitation. *Emergency Medicine Journal* 20, 4 (July 2003), 316–318. DOI:<https://doi.org/10.1136/emj.20.4.316>
- [5] Matthias Jacob, Daniel Chappell, and Markus Rehm. 2009. The 'third space'—fact or fiction? *Best Practice & Research. Clinical Anaesthesiology* 23, 2 (June 2009), 145–157. DOI:<https://doi.org/10.1016/j.bpa.2009.05.001>
- [6] Anne Kadet. 2015. House Calls for Hangovers. *Wall Street Journal* (August 2015).
- [7] Thomas Latta. 1832. MALIGNANT CHOLERA.: DOCUMENTS COMMUNICATED BY THE CENTRAL BOARD OF HEALTH, LONDON, RELATIVE TO THE TREATMENT OF CHOLERA BY THE COPIOUS INJECTION OF AQUEOUS AND SALINE FLUIDS INTO THE VEINS. *The Lancet* 18, 457 (June 1832), 274–280. DOI:[https://doi.org/10.1016/S0140-6736\(02\)80289-6](https://doi.org/10.1016/S0140-6736(02)80289-6)
- [8] R Lewins. 1832. Injection of Saline Solutions into the Veins. *BOSTON MEDICAL AND SURGICAL JOURNAL* VI, 24 (1832), 373–374.
- [9] Timothy E. Miller and Paul S. Myles. 2019. Perioperative Fluid Therapy for Major Surgery. *Anesthesiology* 130, 5 (May 2019), 825–832. DOI:<https://doi.org/10.1097/ALN.0000000000002603>

C. Paper 3

- [10] Xavier Monnet and Jean-Louis Teboul. 2018. My patient has received fluid. How to assess its efficacy and side effects? *Annals of Intensive Care* 8, 1 (December 2018), 54. DOI:<https://doi.org/10.1186/s13613-018-0400-z>
- [11] Georg A Petroianu. 2021. On saline infusion, clonus, molecules and forgotten scientists: Who was Dr Julius Sander (1840–1909)? *Journal of Medical Biography* (December 2021), 09677720211065357. DOI:<https://doi.org/10.1177/09677720211065357>
- [12] Ian Smith, Peter Kranke, Isabelle Murat, Andrew Smith, Geraldine O’Sullivan, Eldar Søreide, Claudia Spies, and Bas in’t Veld. 2011. Perioperative fasting in adults and children: Guidelines from the European Society of Anaesthesiology. *European Journal of Anaesthesiology / EJA* 28, 8 (August 2011), 556–569. DOI:<https://doi.org/10.1097/EJA.0b013e3283495ba1>