



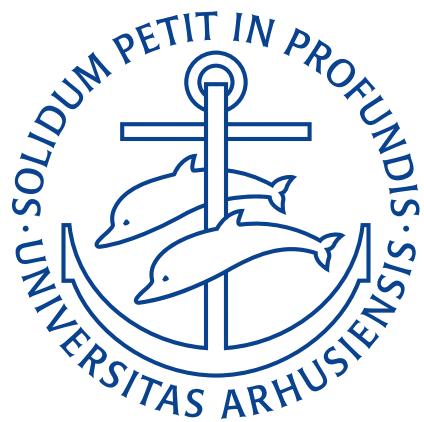
AARHUS  
UNIVERSITY

# Fluid Responsiveness Prediction During Surgery

- Physiological and Methodological  
Limitations and Considerations

PhD dissertation

Johannes Aagaard Enevoldsen



Health  
Aarhus University  
2022

# **Fluid Responsiveness Prediction During Surgery**

**– Physiological and Methodological  
Limitations and Considerations**

PhD dissertation

Johannes Aagaard Enevoldsen

Health  
Aarhus University  
Department of Clinical Medicine

For my girls

## Acknowledgements

First, thanks to my supervisor, Simon Tilma Vistisen. I have interrupted Simon’s work endlessly for close to four years, and he has kept his door open regardless. Many half-baked ideas, questions and frustrations have been resolved over a quick (rarely), unsolicited redecoration of Simon’s whiteboard. I suspect that few PhD students are lucky to have a supervisor who tolerates, and even engages in, biweekly musings on the wonders of *mixed models*.

Also, a big thanks to my co-supervisors, Thomas Scheeren, Stephen Edward Rees and Peter Juhl-Olsen. They helped design a clinical study, that was both safe, clinically relevant and practical to conduct, and were always available to answer questions.

Thanks to Thomas Scheeren for inviting me to Groningen—a stay that was canceled due to COVID19. And thanks to Stephen Rees and Dan Karbing for supervising my overly ambitious project on cardiovascular models at Aalborg University. Despite not even getting close to finishing the project, I have never learned so much in one month.

Thanks to the members of the assessment committee, Professor Hanne Berg Ravn and Professor Anders Åneman, for agreeing to assess this dissertation and the public defence, and to Professor Ebbe Bødtkjær, chairman at the public defence.

I am grateful to the patients who agreed to participate in the clinical study comprising a large part of this PhD.

Thanks to Aarhus University and *Holger & Ruth Hesse's mindefond* for supporting this work.

I wish to thank all my colleagues and collaborators. Thank you for interesting scientific discussions and entertaining lunch conversation. Thanks to the clinical staff who welcomed me and my data cables in the operating room, and helped me conduct the clinical study.

Also, thanks to all the kind strangers on the internet who selflessly helped me with various statistical and technical questions 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 Papers 2 and 3. Also, special thanks

to Gavin L. Simpson, who after helping out with several of the questions I posted on Stackoverflow.com and Twitter, agreed to co-write Paper 2. The template used for this dissertation is [oxforddown](#) by Ulrik Lyngs, John McManigle, Sam Evans and Keith Gillow. Thank you for sharing.

Thanks to friends and family for laying ear to both frustrations and excitement over nerdy details, and for your support in general.

Anna-Elisabeth, thank you for being a loving wife and mother, and for your admirable patience during periods with frequent absent-minded gazes from your husband. And thanks Senja and Gry for making every day meaningful, interesting and joyful.

Johannes Aagaard Enevoldsen  
Aarhus University  
Fall, 2022

# Abstract

## Background

Administration of intravenous fluids is a frequent medical intervention during surgery. However, fluids have side effects, and increasing concern is being raised about the widespread use of intravenous fluids. To reduce side effects, fluids should only be administered to patients who can mobilise the added fluid via an increase in the heart's stroke volume (SV)—this is called a fluid response. If a patient's SV does not respond to fluid, they do not benefit from the fluid and can only experience side effects.

Several methods for predicting a patient's response to fluid have been proposed. One of the best indicators that a patient's SV will increase with fluid administration, is high ventilator-induced pulse pressure variation (PPV): the cyclic change in pulse pressure caused by mechanical ventilation. However, after this method was developed patients are generally being ventilated with lower tidal volumes ( $V_T$ ) and higher respiratory rates (RR), and it has been shown that this *lung-protective* ventilation lowers the predictive accuracy of PPV.

Another approach to fluid responsiveness prediction is the mini-fluid challenge (MFC), where a patient's fluid-responsiveness status is tested with a small amount of fluid. Only if the patient responds to the MFC, should a larger bolus of fluid be given. The hypothesis that the MFC-response predicts the response to a larger fluid bolus has been investigated in several studies.

## Aims

The aim with this PhD project was to tackle challenges in fluid responsiveness prediction during surgery. Specifically, I wanted to develop methods for overcoming current clinical limitations to the use of PPV, and to understand and describe a methodological problem present in most MFC studies.

## Papers

This dissertation is composed of three papers:

Paper 1 describes how the design used in most MFC studies creates a mathematical coupling between the predictor (the MFC-response) and the outcome to predict (the full

fluid response). This causes an overestimation of the MFC's predictive ability. An improved design is suggested.

Paper 2 introduces generalized additive models (GAMs) as a tool for analysing medical time series and waveforms. We demonstrate that GAMs can be used to calculate PPV in situations where the classical method fails, and that GAMs can decompose a central venous pressure waveform into physiologically meaningful components.

Paper 3 presents a clinical study of the effects of  $V_T$  and RR on PPV. The results indicate that using a GAM-derived PPV may help overcome some current limitations to the use of PPV, and that PPV's predictive ability might be improved by adjusting PPV for  $V_T$ .

# Dansk Resumé

## Baggrund

Intravenøs indgift af væske er en hyppig medicinsk behandling under operationer. Væskebehandling har dog bivirkninger, og der er stigende bekymring omkring den udbredte brug af intravenøs væske. For at begrænse bivirkninger, bør væskebehandling kun bruges hos patienter, der er i stand til at mobilisere den ekstra væske ved en øgning af hjertets slagvolumen (SV)—dette kaldes et væskerespons. Hvis en patients SV ikke responderer på væske, har patienten ingen gavn af væsken, men kan stadig få bivirkninger.

Der er foreslået flere forskellige metoder til at forudsige en patients respons på en væskeingift. En af de bedste indikatorer for at en patients SV vil øges af en væskeindgift, er en høj respiratorinduceret pulstryksvariation – den cykliske variation i pulstryk forårsaget af respiratorbehandling. Siden denne metode blev udviklet, er man begyndt at bruge lavere tidalvolumen ( $V_T$ ) og højere respirationsrate (RR) under respiratorbehandling, og det er blevet vist at pulstryksvariationen har lavere prædiktiv værdi ved disse indstillinger.

En anden metode til at forudsige et væskerespons er med en *mini-fluid challenge* (MFC). Her testes en patients evne til at respondere på væske først med lille mængde væske. Kun hvis patienten responderer på dette, giver man efterfølgende en større mængde væske. Hypotesen om at et MFC-respons kan prædiktere responset på en efterfølgende større mængde væske er afprøvet i flere studier.

## Formål

Formålet med dette ph.d-projekt var at arbejde med de udfordringer, der aktuelt er i at forudsige patienters væskerespons under operationer. Specifikt ville jeg udvikle metoder til at afhjælpe nogle af de kliniske begrænsninger der eksisterer for brugen af pulstryksvariation, og jeg ville forstå og beskrive et metodemæssigt problem, som går igen i de fleste MFC-studier.

## Artikler

Denne afhandling består af tre artikler:

Artikel 1 beskriver at der opstår en matematisk kobling mellem prediktoren (responset til en MFC) og det endepunkt der skal forudsiges (det totale væskerespons). Dette giver en overestimering af den prædiktive værdi af en MFC. Derfor foreslås et bedre design.

Artikel 2 introducerer generaliserede additive modeller (GAM) som et værktøj til at analysere medicinske tidsserier og kurvedata. Vi demonstrerer at en GAM kan bruges til at beregne pulstryksvariation, i situationer hvor den klassiske metode ikke virker, og at man kan bruge en GAM til at opdele en centralvenetrykskurve i fysiologisk forståelige komponenter.

Artikel 3 beskriver et klinisk forsøg, hvor vi undersøgte effekten af  $V_T$  og RR på pulstryksvariation. Resultaterne tyder på at GAM-afledt pulstryksvariation kan afhjælpe nogle af de kliniske begrænsninger for brugen af pulstryksvariation, og at den prædiktive værdi af pulstryksvariation kan øges ved at lave en justering for  $V_T$ .

# Contents

<b>Acknowledgements</b>	<b>ii</b>
<b>Abstract</b>	<b>iv</b>
<b>Dansk Resumé</b>	<b>vi</b>
<b>Contents</b>	<b>viii</b>
<b>List of Figures</b>	<b>ix</b>
<b>List of Abbreviations</b>	<b>x</b>
<b>Papers</b>	<b>1</b>
<b>1 Introduction</b>	<b>2</b>
<b>2 Background</b>	<b>4</b>
2.1 History . . . . .	4
2.2 Why give IV fluids? . . . . .	5
2.3 Concerns about liberal use of IV fluids . . . . .	6
2.4 The physiology of a fluid bolus . . . . .	6
2.5 When should fluid be given—and how much? . . . . .	13
2.6 Fluid responsiveness prediction . . . . .	15
2.7 Passive leg raise . . . . .	15
2.8 The mini-fluid challenge . . . . .	16
2.9 Pulse pressure variation and heart-lung interaction . . . . .	16

## *Contents*

2.10 Other dynamic preload indicators . . . . .	20
2.11 Aims . . . . .	21
<b>3 Methods</b>	<b>22</b>
3.1 An introduction to <i>models</i> . . . . .	22
3.2 Paper 1 – The mini-fluid challenge . . . . .	24
3.3 Paper 2 – Generalized additive models . . . . .	26
3.4 Paper 3 – Effects of Ventilator Settings on PPV . . . . .	31
<b>4 Results and discussion</b>	<b>41</b>
4.1 Paper 1 – Most mini-fluid challenge studies overestimate predictive performance . . . . .	41
4.2 Paper 2 – GAMs can extract PPV at ventilation with low HR/RR and decompose the heart-lung interaction in a CVP waveform . . . . .	43
4.3 Paper 3 – There is a near-linear relationship between PPV and $V_T$ . . . . .	46
<b>5 Conclusion and future research</b>	<b>51</b>
5.1 Future research . . . . .	52
5.2 General challenges in intraoperative fluid responsiveness prediction . . . . .	53
<b>References</b>	<b>55</b>
<b>Appendices</b>	
<b>A Paper 1</b>	<b>66</b>
<b>B Paper 2</b>	<b>75</b>
<b>C Paper 3</b>	<b>89</b>
C.1 Paper 3 - Supplementary figures and tables . . . . .	115
<b>D Declarations of co-authorship</b>	<b>121</b>

## List of Figures

2.1	Steps from fluid administration to patient benefit . . . . .	5
2.2	Otto Frank's tracings . . . . .	7
2.3	Illustration of the Frank-Starling mechanism . . . . .	8
2.4	Simple model of venous return . . . . .	9
2.5	The arterial system in the simple model of venous return . . . . .	10
2.6	Guyton's venous return graph . . . . .	11
2.7	Passive leg raise manoeuvre . . . . .	15
2.8	Example of ventilator induced PPV . . . . .	17
2.9	Heart-lung interaction physiology . . . . .	18
2.10	PPV reflects the current slope on the Frank-Starling curve . . . . .	18
3.1	The common mini-fluid challenge study design . . . . .	25
3.2	Regression towards the mean . . . . .	26
3.3	MFC simulation without response . . . . .	27
3.4	A wooden spline . . . . .	28
3.5	Linear model vs GAM . . . . .	28
3.6	A cyclic spline . . . . .	29
3.7	Penalised regression spline . . . . .	29
3.8	Fitting a GAM pulse pressure . . . . .	31
3.9	Fitting a GAM to CVP . . . . .	32
3.10	Ventilator settings in Paper 3 . . . . .	33
3.11	Bayesian updating . . . . .	37

*List of Figures*

3.12 Bayesian updating with two-dimentional visualisation of distributions . . . . .	38
4.1 Alternative analysis of the results from Muller et al., 2011 . . . . .	42
4.2 short caption . . . . .	42
4.3 A GAM fitted to PP at low HR/RR . . . . .	44
4.4 CVP waveform . . . . .	45
4.5 A GAM fit of CVP before and after a fluid bolus . . . . .	46
4.6 Bland-Altman plot of $PPV_{Classic}$ vs $PPV_{GAM}$ . . . . .	47
4.7 Difference between $PPV_{Classic}$ and $PPV_{GAM}$ by HR/RR . . . . .	48
4.8 PPV (GAM and Classic) for all ventilator settings . . . . .	48
4.9 Effects of ventilator settings on PPV . . . . .	49

## List of Abbreviations

<b>ABP</b>	Arterial blood pressure.
<b>ANP</b>	Atrial natriuretic peptide.
<b>ARDS</b>	Acute respiratory distress syndrome.
<b>AUROC</b>	Area under the receiver operating characteristic curve.
<b>CI</b>	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.
<b>CO</b>	Cardiac output.
<b>CVP</b>	Central venous pressure.
<b>ECG</b>	Electrocardiogram.
<b>EEOT</b>	End-expiratory occlusion test.
<b>GAM</b>	Generalized additive model.
<b>GDT</b>	Goal-directed hemodynamic therapy.
<b>HES</b>	Hydroxyethyl-starch.
<b>HR</b>	Heart rate.
<b>ICU</b>	Intensive care unit.
<b>IV</b>	Intravenous.
<b>LVOT</b>	Left ventricular outflow tract.
<b>MAP</b>	Mean arterial pressure.
<b>MFC</b>	Mini-fluid challenge.
<b>P<sub>MSF</sub></b>	Mean systemic filling pressure.
<b>PBW</b>	Predicted body weight.
<b>PEEP</b>	Positive end-expiratory pressure.
<b>PLR</b>	Passive leg raise.
<b>PP</b>	Pulse pressure.

*List of Abbreviations*

- PPV** . . . . . Pulse pressure variation.
- RA** . . . . . Right atrium.
- ROC** . . . . . Receiver operating characteristics.
- RR** . . . . . Respiratory rate.
- RV** . . . . . Right ventricle.
- SD** . . . . . Standard deviation.
- SV** . . . . . Stroke volume.
- SVI** . . . . . Body weight indexed stroke volume.
- SVV** . . . . . Stroke volume variation.
- TAPSE** . . . . Tricuspid annular plane systolic excursion.
- TD** . . . . . Thermodilution.
- TIPS** . . . . . Transjugular intrahepatic portosystemic shunt.
- V<sub>T</sub>** . . . . . Tidal volume (generally in ml kg(PBW)<sup>-1</sup>).
- VTI** . . . . . (Left ventricular outflow tract) velocity time integral.

# 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** The Effects of Respiratory Rate and Tidal Volume on Pulse Pressure Variation in Healthy Lungs—A Generalized Additive Model Approach May Help Overcome Limitations. *Submitted for publication.*

“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 throughout the operation.

Like every treatment, intravenous (IV) fluid should only be given to patients who will benefit from the fluid. This is the setup for a prediction problem: can we, before we give the fluid, 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: 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.

## *1. Introduction*

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

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

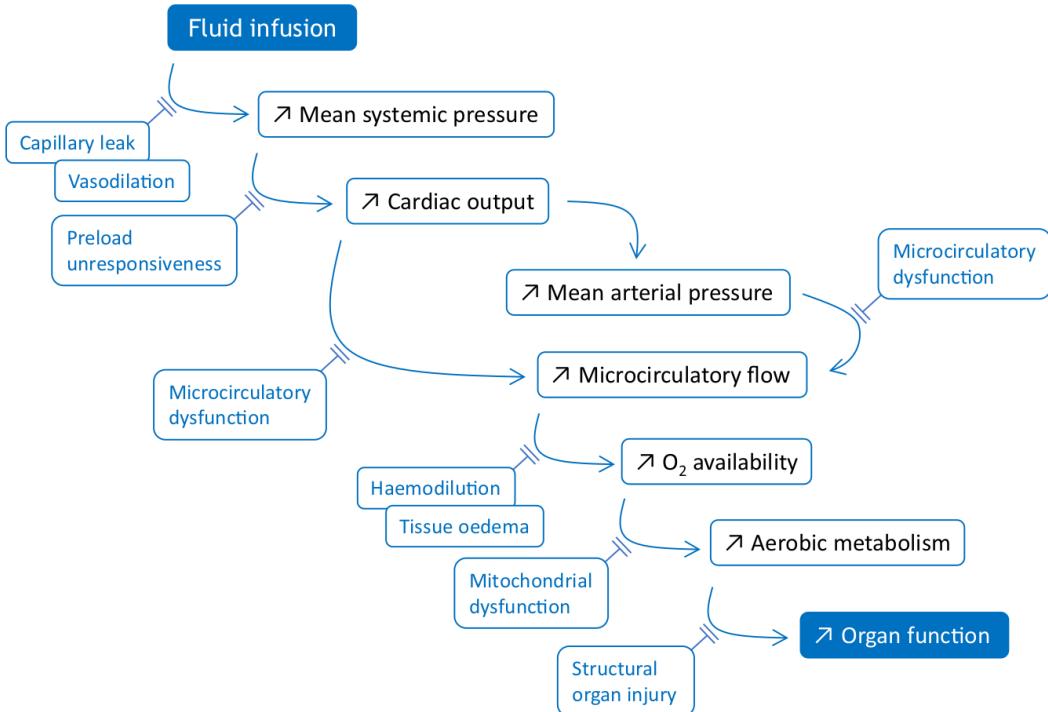
### 2.1 History

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” [1,2]. However, 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 contemporary sceptics were probably highly relevant—the fluid was both unsterile and unphysiological (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 [4,5]. This was followed by a number of reports of successful IV fluid resuscitations in humans [5].

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 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. Background



**Figure 2.1:** Illustration of the physiological steps from fluid administration to benefit. Reprinted from Monnet et al., 2018 [8] (CC BY).

## 2.2 Why give IV fluids?

Fluid should only be administered, when the patient is likely to benefit from it [7]. As a treatment for hypovolemia, the goal is that the infused fluid will increase the circulating volume and thereby the mean systemic filling pressure (see section 2.4.2). This gives an increase in cardiac preload and may, through the Frank-Starling mechanism (see Section 2.4.1), increase CO. This should increase microcirculation and, in turn, the oxygen available to organs, thereby increasing (or retaining) organ function (see Figure 2.1) [8]. As discussed in Section 2.4, this goal is not always achieved.

Another common aim with IV fluid therapy is to resuscitate dehydration, characterised by plasma hypertonicity 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 [9]. 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 circulating 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

## *2. Background*

urination is accurately recorded throughout the surgery, the unknown volume of fluid to replace is from perspiration and redistribution [10]. 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 [10,11]. 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 [7,10,12].

With the risk of organ failure from hypovolemia, why not simply give the patient some extra fluid to ensure that there is enough?

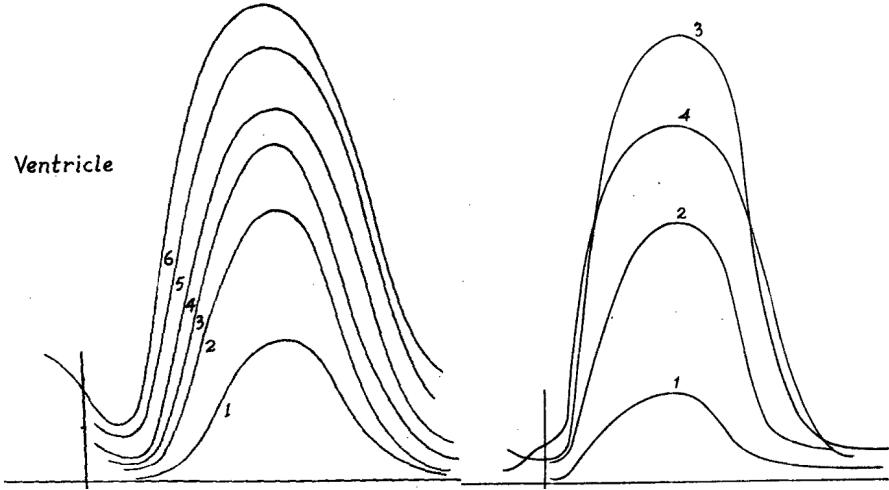
### **2.3 Concerns about liberal use of IV fluids**

The statement that IV fluid administration should be treated as any other prescription medication, has become a trope in fluid resuscitation literature [7,13–15]. For good reasons. In addition to the intended effect on CO, 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 hydroxyethyl-starch (HES) infusion [13,16]. Hypervolemia is a more general issue with excessive fluid administration, causing oedemas of tissue and lungs. The pathophysiology of hypervolemia involves both cardiovascular mechanics and endothelial function, so this will be a good place for an introduction to the physiology of fluid administrations.

### **2.4 The physiology 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.

## 2. Background



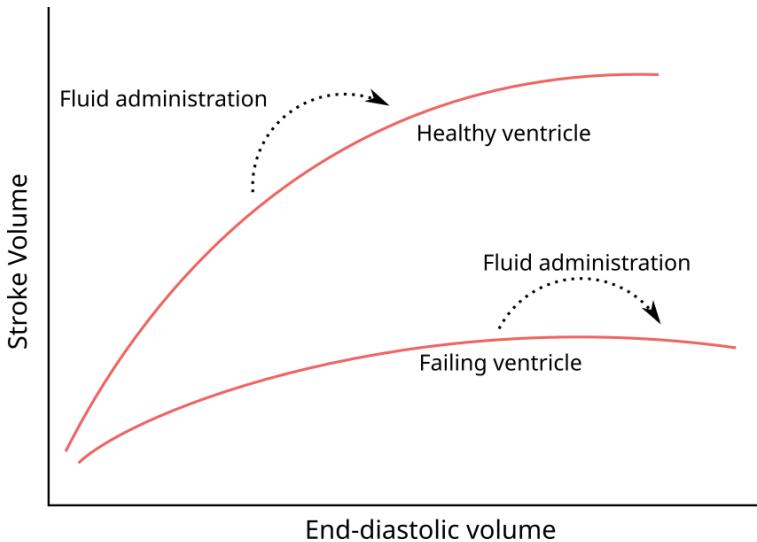
**Figure 2.2:** Otto Frank’s tracings of intraventricular pressure during isovolumetric contractions. In each panel, ventricular volume increases with the number on the tracing. In the right panel, tracing 4 demonstrates the decrease in maximum pressure with overdistension. Reproduction of figures from Frank, 1895 [17], public domain.

### 2.4.1 The Frank-Starling mechanism

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 its discovery [17,18]. It has been noted, though, that the phenomenon had been observed several decades earlier [19]. 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 or 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 2.2). A number of cellular mechanisms for this length-force relationship has been proposed, but the most important contributors seems to be the following: When cardiac sarcomeres are stretched, the lateral distance between actin and myosin decreases, which increases cross-linking between the filaments and increases  $\text{Ca}^{++}$  sensitivity. Also, stretching cardiac myocytes increase their  $\text{Ca}^{++}$  permeability. The drop in force with further stretching has been attributed to a decreasing overlap between filaments, though this mechanism is controversial [20].

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 2.3). A heart’s Frank-Starling curve is, however,

## 2. Background



**Figure 2.3:** 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.

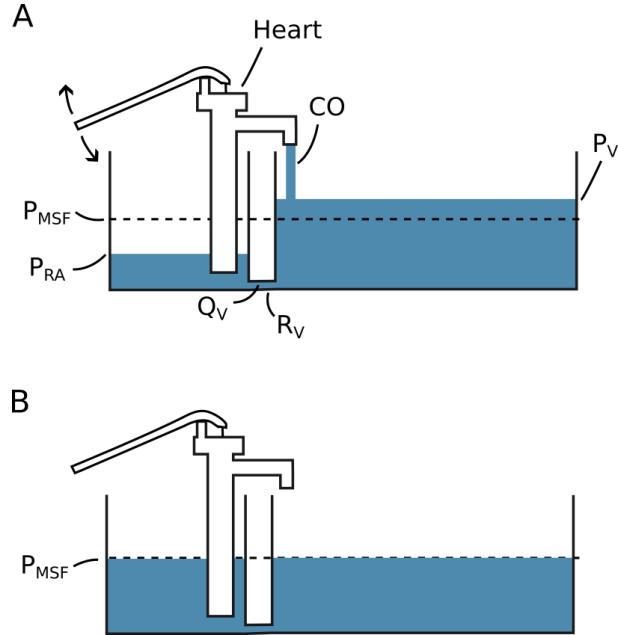
not constant, but can be raised with higher sympathetic tone or sympathomimetic drugs (positive inotropes).

A few concepts are used interchangeably for both the cause (end-diastolic muscle length) and the effect (contraction) in the Frank-Starling relationship. Starling et al., 1914, 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 [18]. 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 (contraction) include afterload, systolic ventricular pressure, stroke volume and mechanical work ( $\text{pressure} \times \text{volume}$ ).

### 2.4.2 Venous return and mean systemic filling pressure

Blood circulates from the heart to the arteries, through tissues, to the veins and back to the heart. In steady state, the circulating volume is constant and CO is equal to

## 2. Background

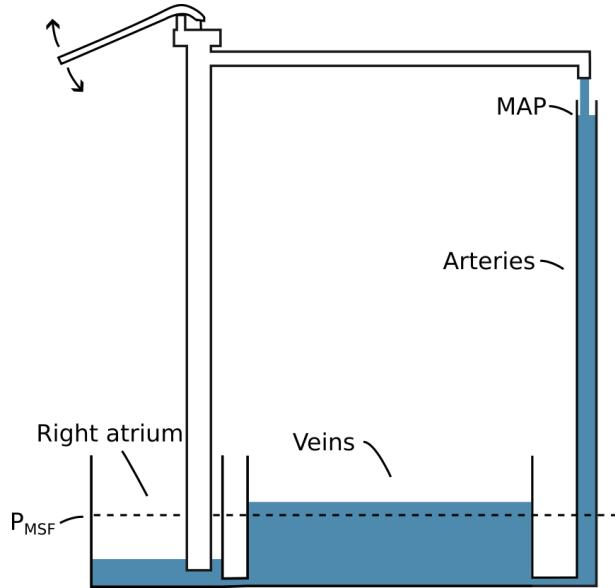


**Figure 2.4:** A simple model illustrating the concepts of venous return ( $Q_V$ ) and mean systemic filling pressure ( $P_{MSF}$ ). The pressure is defined by the height of the fluid surface and the compliance is proportional to the width of the compartment. **A)** The system with a constant cardiac output (CO). **B)** The system with cardiac arrest ( $CO = 0$ ).  $P_V$ , pressure of the compartment representing venules and veins.  $R_V$ , resistance to venous return.  $P_{RA}$ , pressure of the compartment representing the right atrium.

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 a tube (see Figure 2.4). The pump is the heart, the large compartment represents the capacitance of venules and veins, 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 2.5) [21]. 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 ( $P_{MSF}$ ). Often,  $P_V$  is considered equal to  $P_{MSF}$  in this model, since the right atrium and arteries have relatively small volumes, and therefore have little impact on  $P_{MSF}$ .

From this simple model, we can appreciate some factors that determine CO. First, the heart's ability to pump is an absolute limitation to CO. However, the heart also cannot pump more than what is returned from the veins. This venous return ( $Q_V$ ) is

## 2. Background



**Figure 2.5:** An illustration of how the arterial system could be represented in the simple model illustrated in Figure 2.4. MAP, mean arterial pressure.

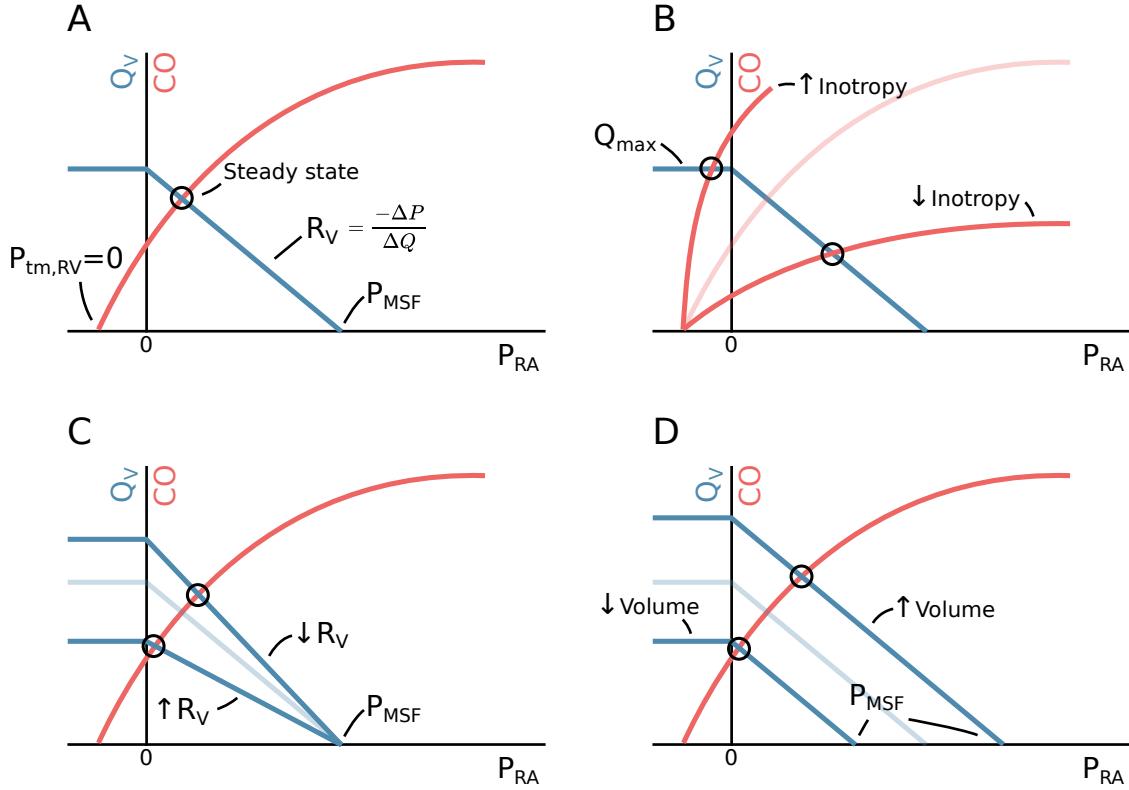
determined by the resistance to venous return and the pressure difference between the venous compartment and the right atrial compartment:

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

Thus, increasing  $P_V$  or lowering  $R_V$  allows a higher CO. If venous compliance is constant, a fluid bolus increases  $P_V$ . Alternatively, we can use an  $\alpha$ -adrenergic agonist such as noradrenaline to decrease venous compliance and thereby increase  $P_V$  without adding fluid (venoconstriction will also increase  $R_V$ , but the effect on compliance seems to dominate) [22]. 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  $Q_V$ . 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  $P_{RA}$ , while the  $P_{RA}$  is inversely related to CO, since a high CO will tend to empty the right atrial compartment. The  $P_{RA}$  and CO where  $Q_V$  and CO are in equilibrium can

## 2. Background



**Figure 2.6:** **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 ( $R_V$ ). **D** Change in  $P_{MSF}$ , either via change in stressed volume (fluid bolus) or compliance of capacitance vessels (venoconstriction).

be found as the intersection between the venous return curve (the relationship between venous return and  $P_{RA}$ ) and the Frank-Starling curve (a variant with  $CO$  rather than  $SV$  on the y-axis). This graphical solution, illustrated in Figure 2.6, was first proposed by Arthur Guyton [23].

The simple model depicted in Figure 2.4 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 inotropic and chronotropic effects. In isolation, positive inotropy or chronotropy will increase  $CO$  and lower  $P_{RA}$ , as depicted in Figure 2.6B. 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

## 2. Background

below zero. Lowering  $P_{RA}$  below zero (only possible because the intrathoracic pressure is below zero during spontaneous breathing) 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 2.6B). 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 [24].

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 [25]. 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 2.6C).

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  $P_{RA}$  (by a smaller amount) and thereby CO on the condition that the heart is operating on the ascending part of the Frank-Starling curve (see Figure 2.6C).

### 2.4.3 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 [26]. Pulmonary oedema has a similar detrimental effect on gas exchange in the alveoli.

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 pressure (oncotic pressure) from the macromolecules in blood and the hydrostatic pressure across the capillary membrane. An increase in stressed volume increases transcapillary pressure, driving fluid into the interstitium until a new equilibrium is reached [27]. 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.

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 lining

## 2. Background

the vascular endothelium [28]. 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. The flow of water from vasculature to tissue may be determined less by oncotic pressure difference and more by the current state of the glycocalyx layer [29]. The permeability of the glycocalyx layer seems to be impacted by volume loading. A proposed mechanism for this regulation is that volume loading increases right atrial pressure, causing release of atrial natriuretic peptide (ANP). ANP increases water filtration and may degrade the glycocalyx [30]. This proposed mechanism has, however, been challenged [31]. Another important cause of glycocalyx degradation is inflammation—especially related to sepsis [32].

### 2.4.4 How long does a fluid bolus remain in circulation?

Fluids must remain in circulation to improve 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 [33].

## 2.5 When should fluid be given—and how much?

There are two overall strategies for fluid management: to replace fluids according to a measured or 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 [7].

In the RELIEF trial, 3000 patients undergoing abdominal surgery were randomised to either a *liberal* fluid regimen, expected to give a positive fluid balance, or to a *restrictive* fluid regimen, expected to give a neutral fluid balance. The *liberal* group received a

## 2. Background

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) [34]. An overall interpretation has been that a positive fluid balance of 1-2 litres is preferable in major abdominal surgery [35].

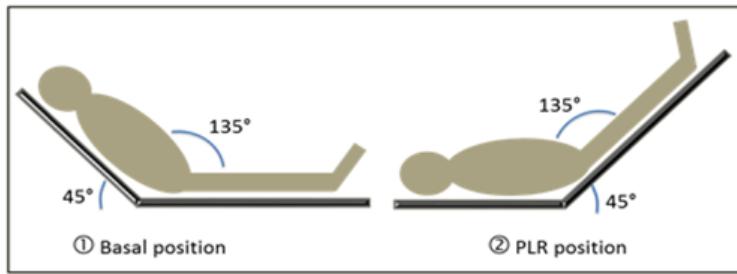
The alternative—or complementary—strategy is goal-directed hemodynamic therapy (GDT). Here, patients are treated with fluid, and often vasopressors, to reach a specific hemodynamic target. The aim with GDT is to individualise treatment to ensure that hypovolemic patients get enough fluid, while avoiding fluid overload. A common GDT target is SV optimisation: fluid is given in boluses until SV stops increasing. This is interpreted as the patient's heart having reached the plateau of the Frank-Starling curve, and that further fluid administration will be futile. An example of GDT was investigated in the OPTIMISE trial [36].

The OPTIMISE trial randomised 724 high-risk, abdominal surgery patients to either SV-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 ( $0.5 \mu\text{g kg}^{-1} \text{ min}^{-1}$ ). 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 [7]. A recent systematic review of 76 randomised GDT trials had a conclusion similar to the OPTIMISE trial: GDT might work [37]. The OPTIMISE II trial recently completed inclusion ( $n = 2502$ ) and will hopefully shed more light on the effects of using a relatively simple SV-optimising protocol [38].

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 3.4.3). Second, it can be 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.

## 2. Background



**Figure 2.7:** Illustration of the passive leg raise (PLR) manoeuvre. From Wikimedia by Patricia.Pineda.Vidal ([CC BY](#))

## 2.6 Fluid responsiveness prediction

Fluid responsiveness prediction attempts to answer the question: will this patient benefit from a fluid bolus? More precisely, it attempts to answer whether fluid will increase the patient's CO, which is a necessary—but not sufficient—condition for benefitting from a fluid bolus (see Figure 2.1).

Clinical observations and measurements have long been used as indications that a patient might benefit from fluid administration. Examples are: tachycardia, hypotension, low central venous pressure (CVP) and high plasma lactate. These clinical findings are all common in hypovolemia, but they also have several other causes. Generally, these findings are poorly related to fluid responsiveness [39–41].

Instead, several *dynamic* methods for predicting fluid responsiveness have been proposed. Here, *dynamic* refers to the concept of manipulating the heart's working position on the Frank-Starling curve, and evaluating the effect on SV (either directly or *via* a surrogate measure).

## 2.7 Passive leg raise

The passive leg raise manoeuvre attempts to evaluate the effect of a fluid bolus, without giving any fluid. By raising the patient's legs, unstressed volume is mobilised, increasing  $P_{MSF}$  analogous to a fluid challenge (see Figure 2.7). If the patient responds to this *autotransfusion*, they are also expected to respond to a subsequent fluid challenge [42]. Unfortunately, the passive leg raise is rather impractical during surgery.

## *2. Background*

### **2.8 The mini-fluid challenge**

In the OPTIMISE trial, all fluid challenges were 250 ml, and each bolus was used to determine whether the next bolus should be given. This can be seen as a form of fluid responsiveness prediction: if the patient does not respond to the bolus, there is a low probability that a subsequent bolus will elicit a response; if the patient does respond, the probability is higher, and a subsequent bolus is given (and so on ...). If these fluid challenges are used frequently to test fluid responsiveness in a non-responsive patient, it can accumulate to a significant amount of ineffective fluid. A potential solution was proposed by Muller et al. in 2011: the mini-fluid challenge (MFC) [43].

With a MFC, a low volume of fluid, 100 ml, is used to test fluid responsiveness. If the patient responds, 400 ml of fluid is subsequently given. Thus, if a patient is non-responsive to fluid, only 100 ml of ineffective fluid have been given.

The accuracy of the MFC in predicting fluid responsiveness was summarised in a systematic review from 2019 including seven studies [44]. The pooled area under the receiver operating characteristic curve (AUROC) was 0.91, 95% confidence interval (CI) [0.85; 0.97], indicating a highly accurate prediction.

#### **2.8.1 Limitations to the mini-fluid challenge**

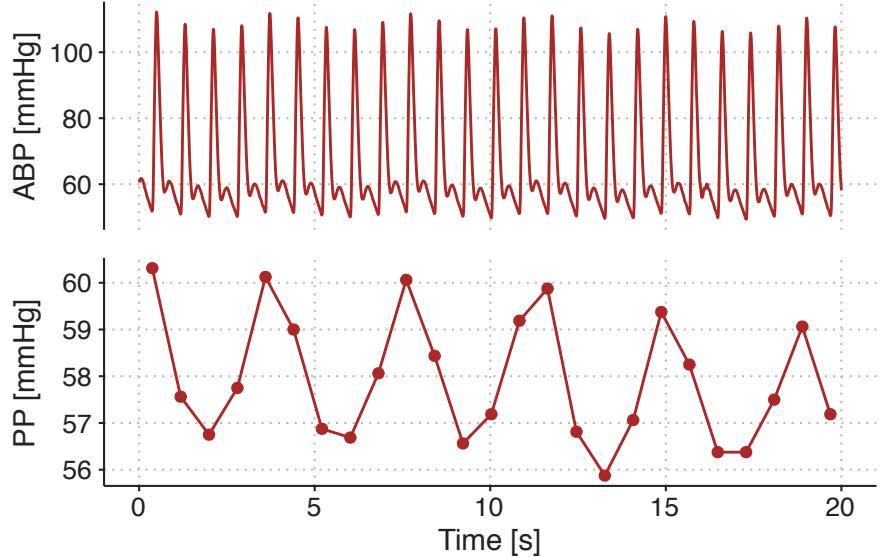
In most studies investigating the predictive ability of the MFC, SV is measured thrice: before the MFC, after the MFC and after the 400 ml bolus. The SV response to the MFC is used to predict the response to the entire 500 ml of fluid given, and the accuracy of this prediction is evaluated. Essentially one change (from 0 ml to 100 ml) is used to predict another change (from 0 ml to 500 ml). Change scores are notoriously difficult to analyse correctly, and most MFC studies do indeed encounter statistical problems. These problems are the topic of Paper 1.

### **2.9 Pulse pressure variation and heart-lung interaction**

Ventilator induced pulse pressure variation (PPV) is the cyclic variation in pulse pressure (PP) caused by mechanical ventilation (see Figure 2.8).

Mechanical ventilation affects circulation in multiple ways [45]. The four principal factors contributing to PPV are illustrated in Figure 2.9. 1) During inspiration, alveolar pressure squeezes out blood from lung capillaries and veins, increasing left-ventricular preload.

## 2. Background



**Figure 2.8:** Example of ventilator-induced PPV in a patient ventilated with a  $V_T$  of  $8 \text{ ml kg}^{-1}$  and a RR of  $17 \text{ min}^{-1}$ . ABP, arterial blood pressure. PP, pulse pressure.

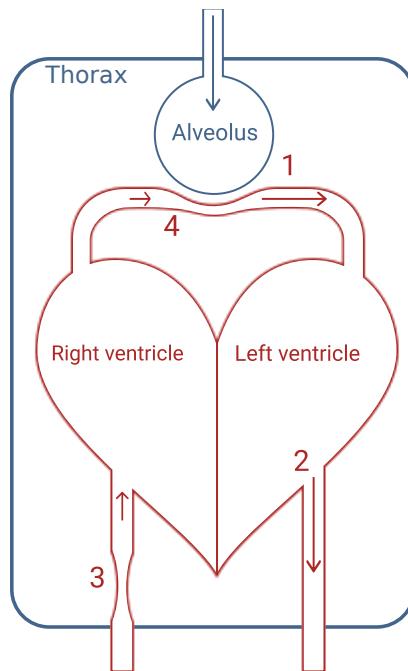
2) Lung inflation raises thoracic pressure and aids the flow of blood from the thorax, reducing left-ventricular afterload. Both effects (1 and 2) tend to increase left-ventricular SV during inspiration. Conversely, (3) right-ventricular preload is decreased, since the increased thoracic pressure reduces venous return, and (4) right-ventricular afterload is increased by the alveolar pressure during inspiration. These effects (3 and 4) decrease right-ventricular SV during inspiration. A decrease in right-ventricular SV will decrease left-ventricular preload, but this effect is delayed through the pulmonary circulation. The decreased SV from the right ventricle affects the left ventricle a few seconds later—at the beginning of the expiration—making all four effects contribute to respiratory variation in left-ventricular SV. In most patients, the effect on the right-ventricular preload (3) is the main mediator of ventilator-induced stroke volume variation (SVV) [46].

Pulse pressure is directly proportional to SV for a given arterial compliance [47]. The cardiac response to the small preload changes caused by the ventilator is therefore reflected in the arterial blood pressure waveform as PPV. Essentially, for a given variation in left-ventricular preload, PPV is proportional to the slope of the Frank-Starling curve at the current preload (see Figure 2.10). Therefore, patients with a high PPV should get an increase in SV if they receive a fluid bolus.

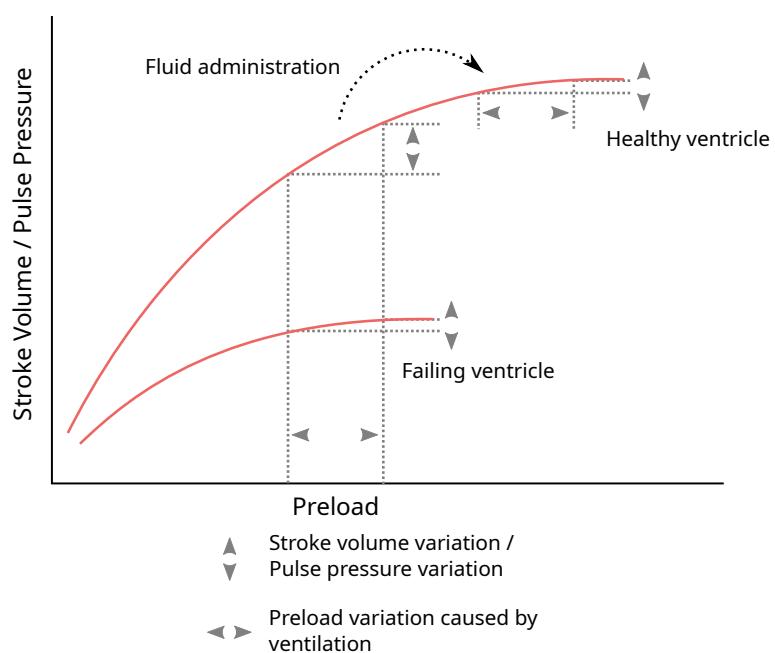
### 2.9.1 Pulse pressure variation as a predictor of fluid responsiveness

In the late 1900s, a number of studies found that systolic pressure variation increased with lower circulating volume in both dogs [48,49] and humans [50,51]. In 1998, Tav-

## 2. Background



**Figure 2.9:** Illustration of the three principal effects. 1) Squeeze of peri-alveolar vessels. 2) Higher thoracic pressure facilitates blood flow out of the thorax. 3) Decreased venous return to the right-ventricle due to higher thoracic pressure. 4) Alveolar pressure increases right-ventricular afterload.



**Figure 2.10:** Illustration of how PPV reflects the current slope on the Frank-Starling curve, and, therefore, is an indicator of fluid responsiveness.

## 2. Background

ernier et al. investigated the more clinically relevant question: did high systolic pressure variation predict fluid responsiveness?—it did [52]. The following couple years, results by Michard et al. indicated that PPV was an even better predictor [53,54]. This finding was supported in a systematic review from 2009 that reported a pooled AUROC of 0.94, 95%CI [0.92; 0.96] [55].

### 2.9.2 Limitations to pulse pressure variation

The use of blood pressure swings induced by heart-lung interactions by mechanical ventilation gained significant interest and scientific credibility in the 1990's and early 2000's [56]. In 2005, De Backer et al. published what would become a landmark study regarding limitations to the use of PPV: predicting fluid responsiveness with PPV was less accurate when the  $V_T$  was  $< 8 \text{ ml kg}^{-1}$  predicted body weight [41]. This study of 60 patients was soon adopted as firm evidence of the existence of a strict  $V_T < 8 \text{ ml kg}^{-1}$  limitation. The study, however, had some limitations itself. Patients that were ventilated with a low  $V_T$  ( $< 8 \text{ ml kg}^{-1}$ ) also had other unique and possibly confounding characteristics: all but one had acute respiratory distress syndrome (ARDS), while only one had ARDS in the higher  $V_T$  group. Therefore, lung-mechanics, including compliance, were very different between the groups. Respiratory rates were likely also higher in the low  $V_T$  group to maintain adequate ventilation. In later studies the effects of RR and respiratory system compliance was investigated. De Backer et al., 2019 reported that a heart-rate-to-respiratory-rate ratio (HR/RR)  $< 3.6$  markedly reduced PPV [57] (though this effect could also be caused by RR directly), and Monnet et. al. showed in 2012 that low respiratory system compliance ( $< 30 \text{ ml cmH}_2\text{O}^{-1}$ ) reduced PPV's ability to predict fluid responsiveness [58]. Today, all these aspects are considered individual limitations to PPV [56,59]. However, low  $V_T$ , high RR and low respiratory compliance often appear together (e.g. in ARDS), and one or more of these circumstances may confounded the others when identifying them as limitations.

Dissecting the effects of  $V_T$  and RR on PPV is the topic of Paper 3, and Paper 2 presents an alternative method for calculating PPV, which may be less sensitive to low HR/RR.

The ventilation-related limitations to PPV presented above are not the only situations where PPV may be unreliable. Some additional limitations are presented here, but these will not be addressed further in this dissertation. The most common limitations to the use of PPV has been summarised with the **LIMITS** acronym [59]: **L** is for low HR/RR, which is one of the limitations addressed in this dissertation. **I** is for irregular heart beats, representing that arrhythmia and frequent extrasystoles causes PPV that is not

## 2. Background

just related to ventilatory preload variation, but also related to variable intervals between beats. Wyffels et al., 2021 present a promising method for extracting ventilatory PPV in patients with atrial fibrillation, but this method has not yet been clinically validated [60]. **M** is for mechanical ventilation with  $V_T < 8 \text{ ml kg}^{-1}$ : the other limitation addressed in this dissertation. **I** is for increased abdominal pressure, which primarily refers to laparoscopic surgery. High abdominal pressure decreases thoracic compliance, potentially increasing PPV for a given  $V_T$  [61]. **T** is for open thorax, where variation in intrathoracic pressure is eliminated. And **S** is for spontaneous breathing, which also refers to minimally sedated patients whose ventilation is supported by a ventilator. Spontaneous breathing reduces or inverts the ventilatory variation in thoracic pressure, making any resulting PPV difficult to interpret.

### 2.10 Other dynamic preload indicators

During the past two decades, other *dynamic* methods for predicting fluid responsiveness have emerged. These include the end-expiratory occlusion test (EEOT) and the tidal volume ( $V_T$ ) challenge [62].

Of these, the EEOT is the best validated. The method utilises that an transient end-expiratory occlusion manoeuvre (15 seconds) increases the pressure gradient for venous return (by lowering  $P_{RA}$ ), causing a right-shift on the Frank-Starling curve [63]. If the flow response to the EEOT is high, patients are likely to be fluid responders. A systematic review of this method reported an AUROC of 0.96, 95%CI [0.92; 1.00], for predicting fluid responsiveness [44]. The drawback of the EEOT is that it requires active intervention, and patients must be mechanically ventilated with no or little breathing effort. It is best validated with continuous flow monitoring as the predictor, though an increase in PP with end-expiratory occlusion seems to predict fluid responsiveness as well [62]. Also, it may require that the baseline  $V_T$  is  $>8 \text{ ml kg}^{-1}$  [64]. Most EEOT studies have been conducted in the ICU, and end-expiratory occlusion may not be possible on all anaesthesia machines, potentially limiting the applicability for intraoperative use [65].

The  $V_T$  challenge consists of changing the  $V_T$  from  $6 \text{ ml kg}^{-1}$  to  $8 \text{ ml kg}^{-1}$ , causing an increase in PPV [64]. The manoeuvre increases PPV and the increase should be higher in fluid responders. In a recent systematic review that only included a few  $V_T$  challenge studies, the  $V_T$  challenge predicted fluid responsiveness with an AUROC of 0.92 (no CI reported) [66]. The  $V_T$  challenge requires an active intervention, and has the same limitations as PPV itself. When the  $V_T$  challenge is done, and  $V_T$  is  $8 \text{ ml kg}^{-1}$ , the  $V_T$  limitation is of course removed. The PPV at a  $V_T$  of  $8 \text{ ml kg}^{-1}$  is in itself predictive

## *2. Background*

of fluid responsiveness, though the  $V_T$  challenge might offer some additional predictive accuracy [64]. However, the performance of the  $V_T$  challenges has rarely been compared to the performance of PPV at  $V_T=8 \text{ ml kg}^{-1}$ , and Xu et al., 2022 showed no difference in performance between the two predictors [67].

### **2.11 Aims**

The aim of this PhD was to tackle limitations to two popular methods for fluid responsiveness prediction in the intra-operative setting: PPV and MFC. For PPV, I address the limitations that apply to deeply sedated patients in sinus rhythm undergoing open, non-thoracic surgery.

The aim of Paper 1 was to explain a statistical problem with the design used in most studies of the MFC's ability to predict fluid responsiveness, and to propose a better design.

The aim of Paper 2 was to introduce generalized additive models (GAMs) as a tool for analysing medical time series and waveforms, to propose how PPV derived from a GAM could help overcome the low HR/RR limitation, and to propose more advanced and experimental use cases of GAMs.

The aim of Paper 3 was to analyse the effects of  $V_T$  and RR on PPV, and to investigate how these ventilator settings impacted PPV's ability to predict fluid responsiveness.

*“Few terms are used in popular and scientific discourse more promiscuously than ‘model’. A model is something to be admired or emulated, a pattern, a case in point, a type, a prototype, a specimen, a mock-up, a mathematical description—almost anything from a naked blonde to a quadratic equation.”*

— Nelson Goodman, 1968 (Languages of Art, p. 171)

# 3

## Methods

This chapter introduces and discusses the methods used in the three Papers. In all three Papers, *models* have a central role; both as communicative tools and as specific statistical models (Papers 2 and 3). Therefore, the first section is a general introduction to the scientific use of the term *model*.

### 3.1 An introduction to *models*

In science, *model* is often used as a synonym for both a statistical analysis, a simple representation of a complex system or a sophisticated computer simulation. However, the word’s exact meaning is often not specified, and has to be assumed from context. I will use this section to present the different meanings of the term *model*, focusing on those that are relevant for the present dissertation.

#### 3.1.1 A model can be a mental and communicative tool

Consider the four effects of ventilation causing pulse pressure variation, illustrated in Figure 2.9. To recap, in the right ventricle, inspiration lowers SV by limiting venous return and increasing afterload; in the left ventricle, inspiration increases pulmonary venous return by squeezing blood from the lungs and reduces afterload by increasing thoracic pressure relative to the remaining arterial pressure. Together these effects cause PPV. Of course, real physiology is more complex, but this simple *model* may aid a scientist in designing a study investigating the effect of e.g. PEEP on PPV, and it can

### 3. Methods

give colleagues a common reference when discussing the interpretation of a high PPV in a patient with right ventricular failure. Without a common model, miscommunication will likely happen as each colleague has their own private mental model of the situation, and may assume that their peer's mental model is identical.

This conceptual model of PPV is, however, not detailed enough, to enable quantitative predictions of what will happen to PPV, if we change ventilator settings; it is only a qualitative description of the relationship between ventilation and SV. To make quantitative predictions from a model, we need to mathematically define the relationship between the variables in the model.

#### 3.1.2 Mathematical models

A mathematical model is a set of equations that, unambiguously, describes how a system behaves, often over time. The 2-compartment model of venous return (see Figure 2.4) can be presented as a mathematical model. We can present equations defining how the compartments behave and how they are connected. The pressure in the compartments (right atrial and venous, but exemplified here with the right atrial compartment) is

$$P_{RA} = \frac{V_{RA}}{C_{RA}},$$

where  $P$  is pressure,  $V$  is volume and  $C$  is compliance. Venous return ( $Q_V$ ) is

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

where  $R_V$  is resistance to venous return. The change in volume over time in each compartment is:

$$\dot{V}_V = CO - Q_V, \quad \dot{V}_{RA} = Q_V - CO.$$

To solve this set of differential equations, we define a constant CO (or a Frank-Starling-like cardiac function), a fixed compliance for each vascular compartment, a fixed  $R_V$  and set the starting volumes for each vascular compartment ( $V_V$  and  $V_{RA}$ ). We can now simulate how these volumes will evolve over time, and what their steady-state values will be.

### 3. Methods

#### 3.1.3 Statistical models

A statistical model is a specific type of mathematical model that uses random variation to explain the difference between the values expected from the model, and the values that are actually measured. In medical science, statistical models are usually extremely simplified representations of the actual (or assumed) causal relationship between variables.

For example, we may believe that  $V_T$  affects PPV through a number of intermediate steps. But often, we are not interested in the intermediate effects—they may not even be identifiable from the data we have available. Instead we may assume that if we vary  $V_T$  in a single individual, there is a linear relationship between  $V_T$  and PPV:

$$PPV = \alpha + \beta V_T.$$

If we measure PPV repeatedly, and only change  $V_T$  between measurements, we expect that PPV will vary according to the linear model above. Additionally, we also anticipate some variation in PPV that is not explained by the change in  $V_T$ . This additional variation can be due to a physiological change over time, unrelated to the  $V_T$  change, or due to imprecise measurements. Unexplained variation is traditionally represented by the greek letter  $\epsilon$ , creating the statistical linear model:

$$PPV = \alpha + \beta V_T + \epsilon.$$

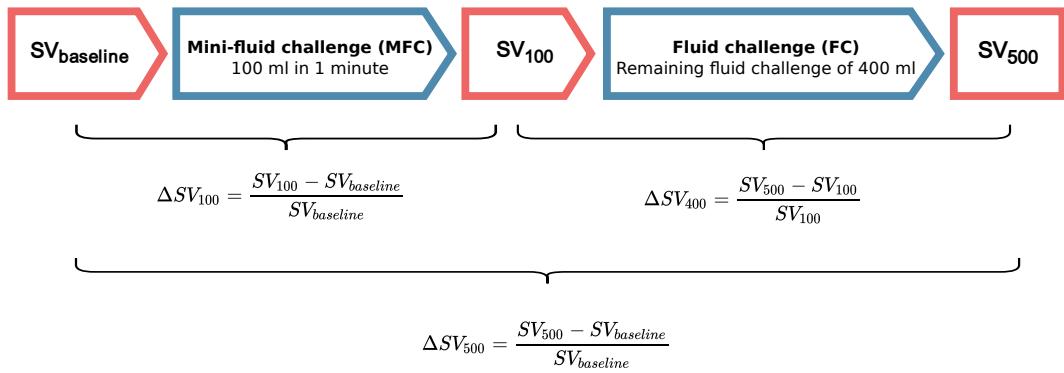
The statistical analysis of such a model quantifies how the independent variables (here  $V_T$ ) and random variation ( $\epsilon$ ) together explain the observed variation in the dependent variable (here PPV), and, importantly, how certain we are of each component's contribution.

## 3.2 Paper 1 – The mini-fluid challenge

Paper 1 is a critique of the method used in most existing studies of the MFC as a predictor of fluid responsiveness. In this section, I will describe this common design and explain why it is problematic.

Most MFC studies [43,68–75] use a design where SV, or a related variable, is measured thrice. First, a baseline measurement before any fluid is given; then, after the MFC (e.g. 100 ml) is given; and, lastly, after the remaining fluid (e.g. 400 ml) is given. The response to the MFC, here referred to as  $\Delta SV_{100}$ , is then calculated as the relative change in SV from baseline to after the MFC. The outcome to predict, is the change from baseline to after the entire volume is given, here referred to as  $\Delta SV_{500}$  (see Figure 3.1).

### 3. Methods



**Figure 3.1:** The design used in most studies of the MFC.

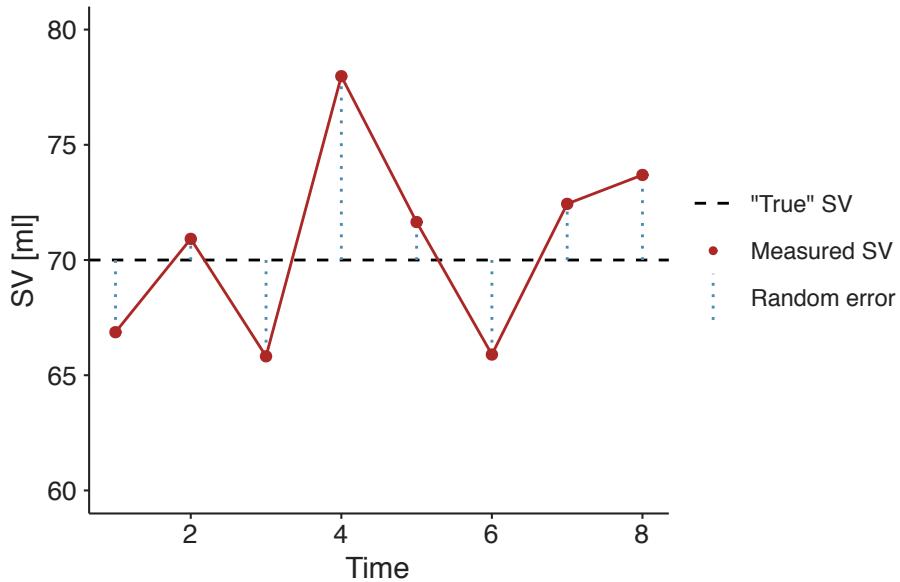
This design has two major problems. First, the response to the MFC is included in the outcome, that the MFC is supposed to predict. This creates a self-fulfilling prophecy: if the MFC response is high enough, the patient can have no response to the remaining fluid, and the conclusion will still be, that the MFC correctly predicted the full fluid response.

The second problem is more subtle. Measurements of SV are influenced by some random variation. This variation is caused by both measurement error and short-term physiological variation in SV. Both are essentially *noise* to the clinician, who wants to evaluate the response to a fluid challenge. The clinician is interested in the underlying, but unknowable average SV, not the short-term variation around this value. I will call this the “true” SV. If a measurement is, by chance, higher than the “true” SV, the next measurement will most likely show a decrease. Likewise, if a measurement is randomly low, the next will likely show be higher than the current. This phenomenon is called *regression towards the mean* (see Figure 3.2).

In the common MFC study design,  $SV_{baseline}$  is included in the calculation of both the predictor ( $\Delta SV_{100}$ ) and the outcome ( $\Delta SV_{500}$ ). Therefore, the random measurement error in  $SV_{baseline}$  is included in both the predictor and the outcome, making them spuriously correlated [76,77]. In Figure 3.3, this spurious correlation is illustrated with a simple computer simulation. It shows a simulated a MFC study, where all subjects have no response to fluid; only random variation in measured SV. A correct analysis should show that the MFC has no predictive ability—since there is nothing to predict—but it is clear that the change over the MFC ( $\Delta SV_{100}$ ) is correlated to the change over the full fluid administration ( $\Delta SV_{500}$ ) (see Figure 3.3b).

The first problem, that the predictor is included in the outcome, can be fixed by predicting the response to only the remaining 400 ml ( $\Delta SV_{400}$ ). This, however, simply reverses the second problem, as the predictor and outcome now share the measurement error at  $SV_{100}$ , creating an inverse spurious correlation (see Figure 3.3c).

### 3. Methods



**Figure 3.2:** An illustration of *regression towards the mean*. An extreme SV measurement is likely followed by a measurement closer to the “true” (mean) SV.

To illustrate the magnitude of the bias, we reanalysed data from the MFC study by Muller et al. [43]. Instead of showing the MFC’s ability to predict the full fluid response, we analysed the MFC’s ability to predict the response to the remaining 400 ml fluid. We did this by extracting data from Fig. 3A in Muller et al., 2011 [43], and used the plotted  $\Delta SV_{100}$  and  $\Delta SV_{500}$  to calculate  $\Delta SV_{400}$ .<sup>1</sup> As noted in the paragraph above, this approach is also biased, but now in the opposite direction.

## 3.3 Paper 2 – Generalized additive models

Paper 2 introduces generalized additive models (GAMs) as a tool for analysing medical time series and waveforms. The paper describes how GAMs can be used for robust estimation of PPV (a method used in Paper 3), and it presents a novel method for decomposing medical waveforms into simpler, repeating components.

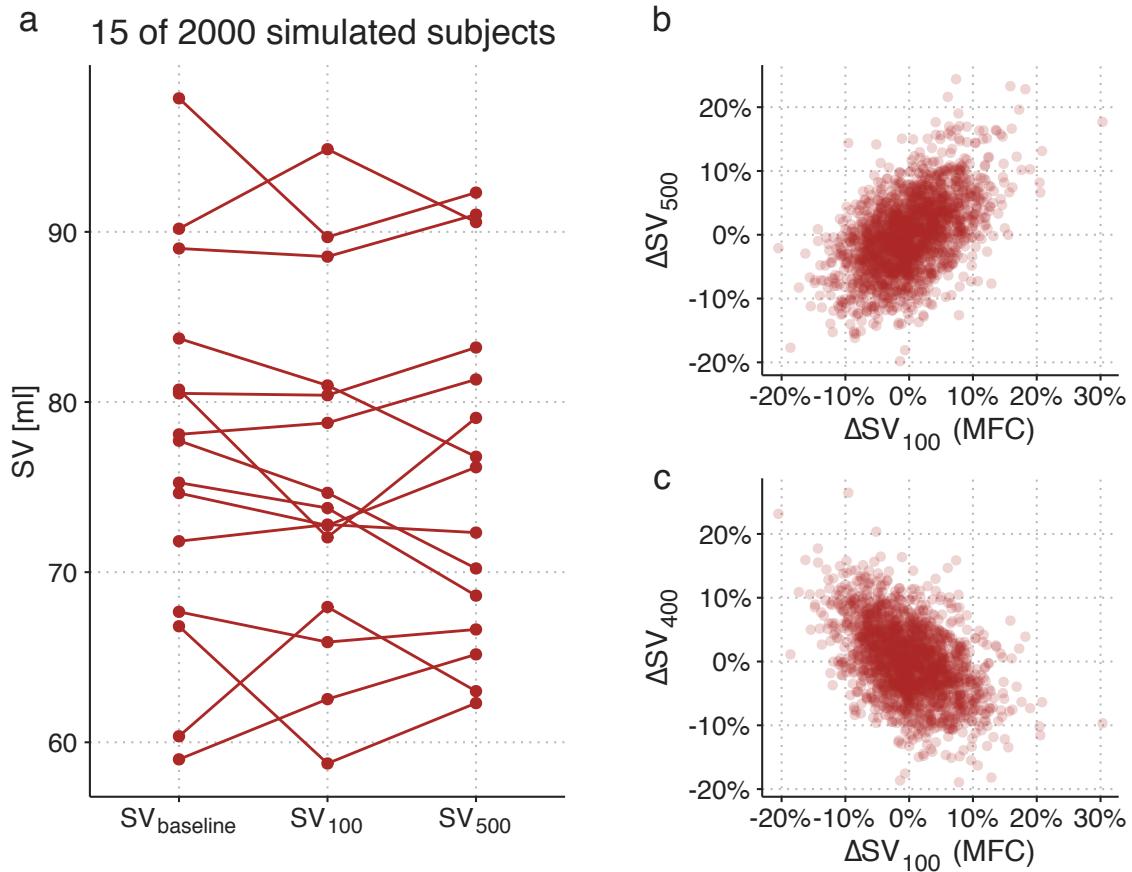
### 3.3.1 An introduction to generalized additive models

A GAM is an extension of the generalized linear model [78]. For the purposes described here, we can, however, disregard the *generalized* part, and simply consider GAMs as multiple regression models that can fit smooth functions instead of only straight lines.

---

<sup>1</sup>In Muller’s study, the fluid response was evaluated with left-ventricular outflow tract velocity time integral (LVOT VTI, a Doppler ultrasonic measure correlated with SV). For simplicity, we simply refer to this as SV.

### 3. Methods



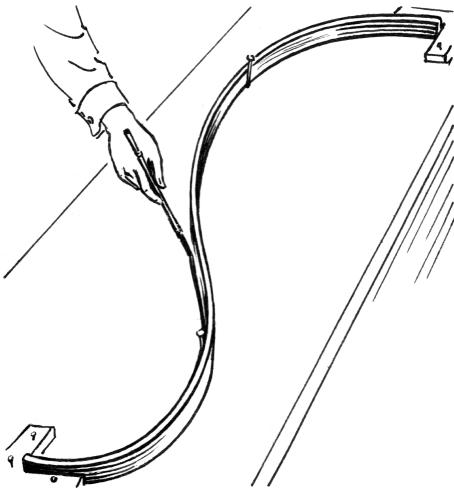
**Figure 3.3:** A simulated MFC study with 2000 subjects. Each subject has a constant “true” SV, and only random measurement variation around this value. **a)** A sample of data from 15 of the simulated subjects. **b)** Relationship between the MFC response and full fluid response. **c)** Relationship between MFC response and the response to the remaining fluid.

These smooth functions are called regression splines (see Figure 3.4). Figure 3.5 illustrates how a linear model and a GAM fit the same simulated data.

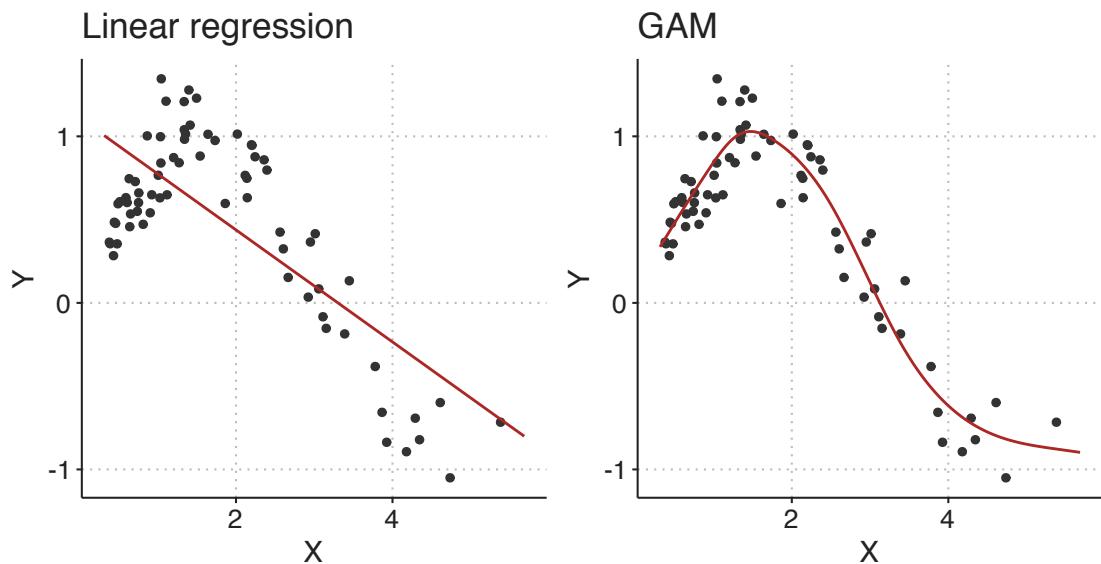
In hemodynamic time series and waveforms, we often have patterns that repeat with each respiratory cycle or heart beat. When we fit such a repeating pattern with GAM, we can constrain the spline to start and end at the same value, slope and curvature (0th, 1st and 2nd derivative), to create a smoothly repeating curve (a cyclic spline). The function that generates the data in Figure 3.5 is  $Y = \sin(X)$ , so we know that it repeats itself at  $X = 2\pi = 6.28$ . Figure 3.6 illustrates the fit of the GAM with this additional, cyclic constraint.

The high flexibility of GAMs could also make them prone to overfitting data. This is countered by penalising the splines by their wigginess (wigginess is the inverse of smoothness).

### 3. Methods



**Figure 3.4:** A regression spline is named after the wooden splines traditionally used by technical drawers for tracing smooth curves. (by Pearson Scott Foresman, public domain).



**Figure 3.5:** A linear model and a generalized additive model (GAM) fitting the same simulated data.

A classic (unpenalised) least squares fit is found by minimising the loss function:

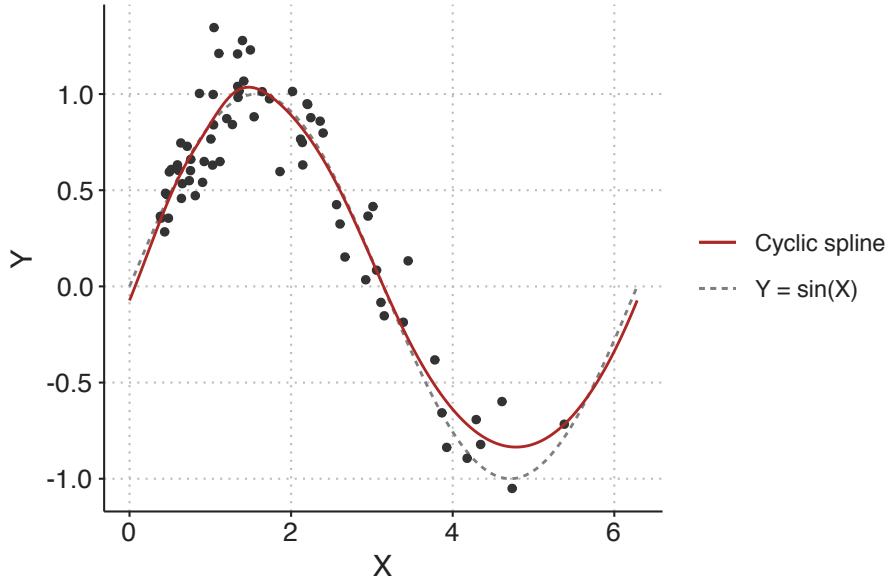
$$\sum_{i=1}^n (Y_i - \hat{Y}_i)^2,$$

where  $Y_i$  are the observed values and  $\hat{Y}_i$  are the values predicted by the model. For fitting penalised splines, a wigginess penalty is added to the loss function:

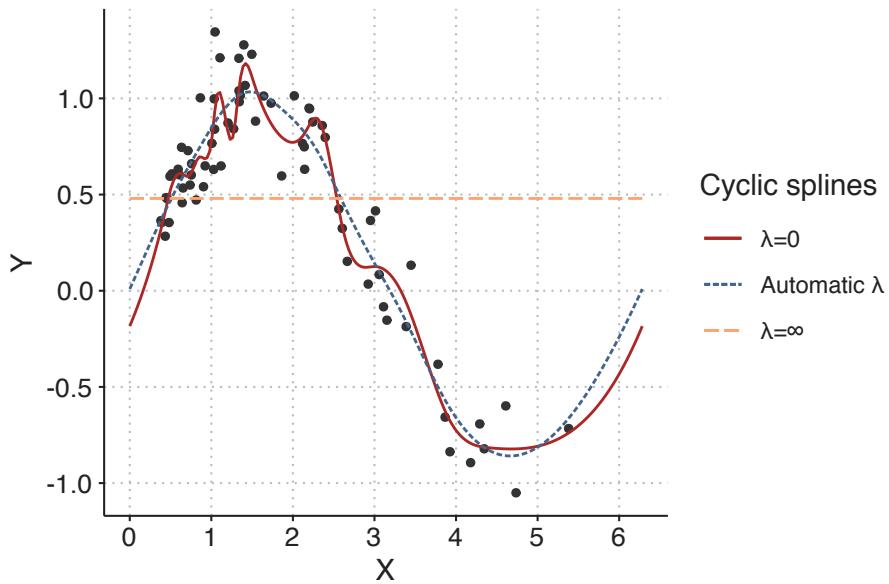
$$\sum_{i=1}^n (Y_i - \hat{Y}_i)^2 + \lambda \int f''(X)^2 dX.$$

$f''(X)$  is the curvature of the spline (for a straight line, the curvature is 0). This is squared

### 3. Methods



**Figure 3.6:** Illustration comparing a cyclic spline fit to the data-generating function ( $Y = \sin(X)$  with added Gaussian noise).



**Figure 3.7:** In a penalised spline, the smoothing parameter,  $\lambda$ , determines the tradeoff between fitting the data and creating a smooth line. *lambda* can be automatically chosen to minimise the risk of overfitting.

to keep it positive and then integrated over the range of the independent variable,  $X$ . The smoothing parameter,  $\lambda$ , is used to control the tradeoff between fitting data with minimal error and preferring a smooth spline less “prone” to overfitting [79]. It is possible, and often recommended, to automatically choose the smoothing parameter that optimise out-of-sample prediction (e.g. through cross-validation) [80] (see Figure 3.7).

### 3. Methods

#### 3.3.2 Estimating PPV using GAMs

Using GAMs for estimating PPV was first proposed by Wyffels et al. [60]. They demonstrated how GAMs could be used to isolate the respiratory variation in PP in patients with atrial fibrillation, from the PPV caused by irregular beat-to-beat intervals. In Paper 2, we demonstrate how to use a simpler version of Wyffel's model to estimate PPV from PP time series recorded on patients in sinus rhythm.

In sinus rhythm, the major source of PPV is respiratory variation in PP (as described in Section 2.9). We expect that there is some smooth relationship between *when in the respiratory cycle a heart beat occurs* and *the PP it produces*. We also expect that the PP produced by a beat in the end of a cycle is very similar to the PP produced by a beat in the beginning of a cycle; i.e., that the effect of the *position in the respiratory cycle* is cyclic. This cyclic respiratory variation in PP can be represented with the following GAM:

$$E(PP) = \alpha + f(pos_{\text{respiratory cycle}}),$$

where  $E(PP)$  is the expected PP,  $\alpha$  is the mean PP and  $f(pos_{\text{respiratory cycle}})$  is the variation in PP around the mean, represented as a cyclic spline that goes from 0 (start of respiratory cycle) to 1 (end of respiratory cycle) and connects smoothly back to 0 (see Figure 3.8b).

In addition to the respiratory variation in PP, we also want the model to allow that PP can change slowly over time:

$$E(PP) = \alpha + f(pos_{\text{respiratory cycle}}) + f(time).$$

Now,  $E(PP)$  is the sum of the repeating effect respiration and a smooth trend in PP over time ( $f(time)$ ) (see Figure 3.8).

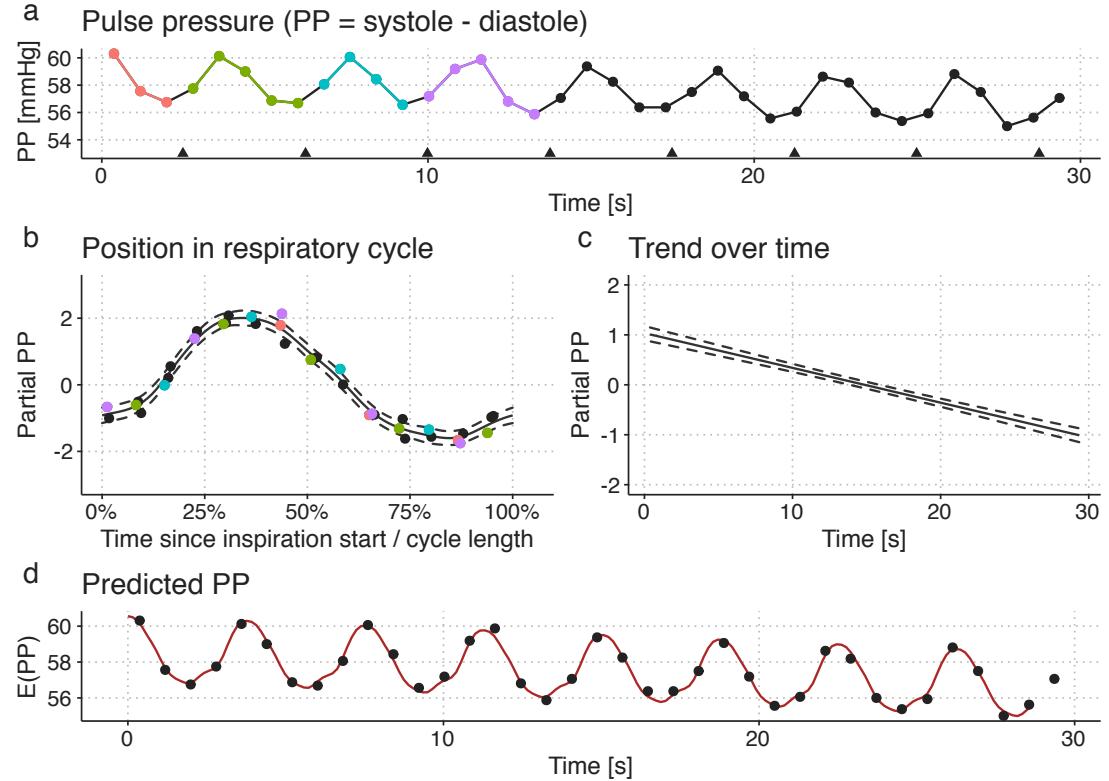
From the fitted GAM, we can calculate PPV as the amplitude of the respiratory effect (Figure Figure 3.8b) divided by the mean PP ( $\alpha$ ):

$$PPV = 100\% \cdot \frac{\max(f(pos_{\text{respiratory cycle}})) - \min(f(pos_{\text{respiratory cycle}}))}{\alpha}$$

#### 3.3.3 Decomposing waveforms

In the PP model above, we reduced the arterial blood pressure waveform to a time series of PP measurements before fitting the GAM. It is also possible to decompose a raw waveform. In addition to the effect of ventilation, we now also have a smooth effect that

### 3. Methods



**Figure 3.8:** Illustration of how a pulse pressure (PP) time series (a) can be modelled as a smooth repeating effect of ventilation (b) and a slow trend over time (c). Panel d shows the models prediction (fit): the sum of the mean PP, the effect of ventilation (b) and the trend over time (c).

repeats with each heartbeat. Furthermore, these two effects may interact. In Paper 2, we exemplify this by decomposing a CVP waveform. The GAM has the following form:

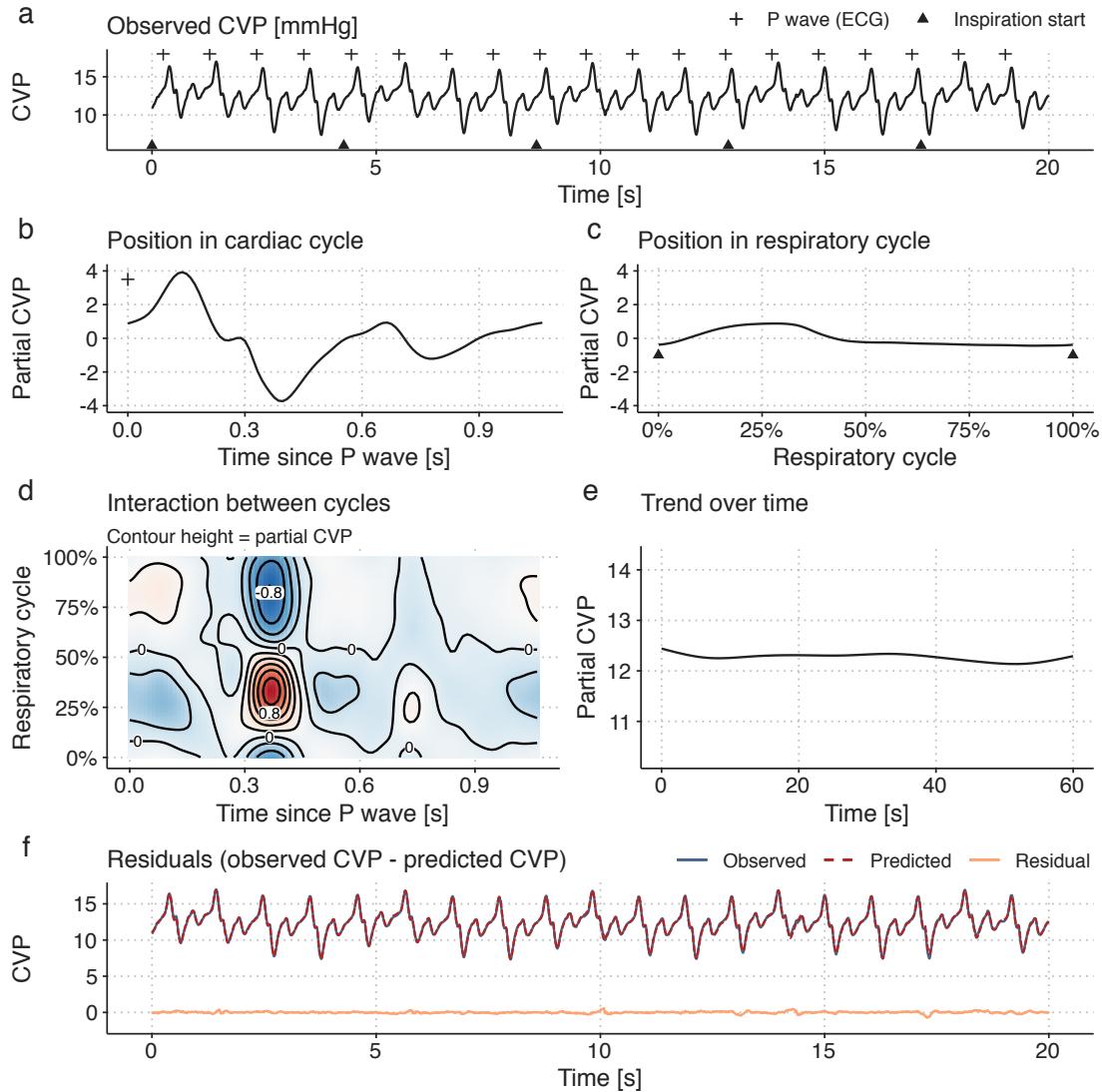
$$E(CVP) = \alpha + f(pos_{\text{cardiac cycle}}) + f(pos_{\text{respiratory cycle}}) \\ + f(pos_{\text{cardiac cycle}}, pos_{\text{respiratory cycle}}) + f(time).$$

Here,  $f(pos_{\text{cardiac cycle}})$  is similar to  $f(pos_{\text{respiratory cycle}})$ , but repeats with every cardiac cycle (starting at the P-wave in the ECG).  $f(pos_{\text{cardiac cycle}}, pos_{\text{respiratory cycle}})$  is a 2-dimensional smooth plane representing the effect of the interaction between the respiratory cycle and the cardiac cycle on CVP (see Figure 3.9).

## 3.4 Paper 3 – Effects of Ventilator Settings on PPV

Paper 3 investigates the effects of  $V_T$  and RR on PPV, and how these ventilator setting affect PPV's ability to predict the SV response to a fluid bolus. We expected that PPV at

### 3. Methods



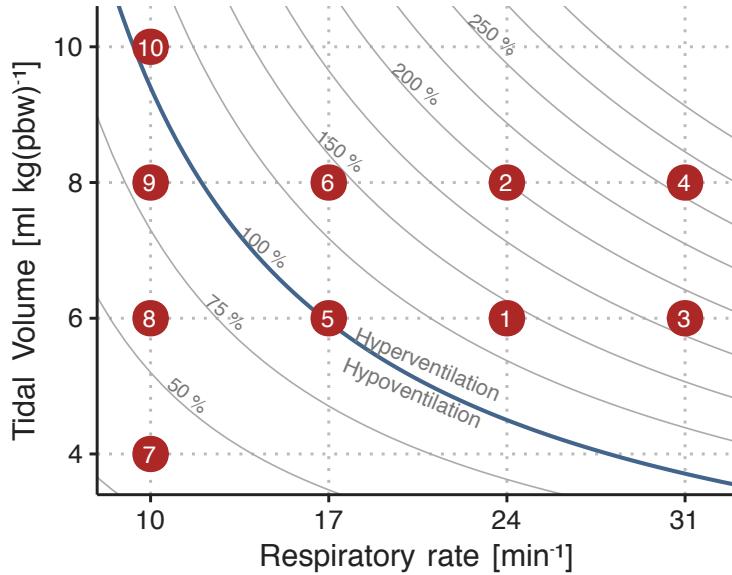
**Figure 3.9:** A central venous pressure (CVP) waveform (a) can be decomposed using a GAM into: b, a repeating effect for each heartbeat; c, a cyclic effect of ventilation; d, an interaction effect between the respiratory- and cardiac cycles; and, e, a slow trend over time (added to the mean CVP). Panel f shows the models prediction and the prediction error (residual): The prediction is the sum of the effects visualised in b to e.

$V_T=10 \text{ ml kg}^{-1}$  and  $\text{RR}=10 \text{ min}^{-1}$  would predict fluid responsiveness with high accuracy, but that the accuracy would fall with lower  $V_T$  and higher RR.

The study was approved by the regional ethics committee (January 2020, case 1-10-72-245-19) and registered on ClinicalTrials.gov (March 2020, NCT04298931). It was conducted at Aarhus University Hospital from May 2020 to June 2021. Patients gave written informed consent to participate.

Paper 3 presents a number of statistical methods including Bland-Altman analysis for method comparison, bootstrapped confidence intervals, mixed-effects models and Bayesian

### 3. Methods



**Figure 3.10:** Example of the ventilation protocol. Dots represent combinations of  $V_T$  and RR. Settings were applied in numerical order. The order of  $RR=17, 24$  and  $31\text{ min}^{-1}$  were drawn at random for each subject. Curved lines represent settings that should result in equal alveolar ventilation, assuming a dead space volume of  $1\text{ ml kg}^{-1}$ .

statistics. In this dissertation, I will only elaborate on Bayesian statistics and explain the Bayesian mixed-effects model presented in the paper.

#### 3.4.1 Protocol

The experiment protocol was initiated after anaesthesia induction, in relation to a planned 250 ml fluid administration, where the patient was considered hemodynamically stable by the treating anesthesiologist. First, the patient was ventilated with ten different ventilator settings for 30 s each. Each setting was a unique combination of  $V_T$  between 4 and 10  $\text{ml kg}^{-1}$  (predicted body weight [81]) and RR between 10 and 31  $\text{min}^{-1}$  (see Figure 3.10). Then, a two-minute period was used to establish a baseline SV, after which a 250 fluid bolus was administered. If SV after the fluid administration was more than 10% higher than the baseline SV, the patient classified as a fluid responder.

#### 3.4.2 PPV calculation

PPV was calculated using two different methods:

$PPV_{Classic}$  was calculated as described in the study introducing low HR/RR as a limitation to the reliability of PPV [53,57]:

$$PPV_{Classic} = \frac{PP_{max} - PP_{min}}{(PP_{max} + PP_{min})/},$$

### 3. Methods

averaged over three consecutive ventilation cycles.

$PPV_{GAM}$  was calculated as described in Paper 2 (see Section 3.3.2).

#### 3.4.3 Stroke volume measurement using uncalibrated pulse contour analysis

In Paper 3, we measured SV using the commercial FloTrac algorithm (fourth generation) built into the cardiac monitor (Hemosphere or EV1000, Edwards Lifesciences, Irvine, California) routinely used during major abdominal surgery at Aarhus University Hospital. The FloTrac estimates beat-to-beat SV with a proprietary algorithm that analyses the arterial blood pressure waveform.

Since change in SV is used to determine a patient's response to a fluid bolus, it is important that the FloTrac estimates SV correctly, and, especially, that changes in SV are measured correctly.

There are several methods for estimating SV or CO. The conversion between these is  $CO = SV \cdot HR$ , and in the following, I will simply refer to all methods as estimating CO. The gold standard for CO measurement is debated, but is generally considered to be the *direct Fick principle* or *thermodilution* ( $CO_{TD}$ ) via a pulmonary artery catheter [82,83]. Thermodilution is more common in clinical practice, and is usually the reference, when new methods for measuring CO are evaluated [84,85].

The first three generations of the FloTrac algorithm were evaluated in a systematic review of 65 studies [86]. The overall *percentage error* was approx. 50%; a slightly simplified interpretation of this statistic is, that if a patient has a *true* CO of  $5 \text{ L min}^{-1}$ , the difference between  $CO_{\text{FloTrac}}$  and  $CO_{TD}$  can be up to  $2.5 \text{ L min}^{-1}$ . The fourth-generation FloTrac (used in Paper 3) aimed to improve performance, but of 6 studies comparing the fourth-generation FloTrac against  $CO_{TD}$ , one showed a percentage error of 35% [87] while the others had a percentage error of  $>50\%$  [88–92].

Even if  $CO_{\text{FloTrac}}$  does not always agree with  $CO_{TD}$ , they could agree on the changes in CO, if the difference between the devices is relatively constant within each patient. This agreement on CO changes (often denoted *trending ability*) is evaluated in most recent studies. These results are, unfortunately, not encouraging either; a 20% increase in  $CO_{TD}$  may be accompanied by a fall in  $CO_{\text{FloTrac}}$  or *vice versa* [87–90,92].

The performance of FloTrac may, however, be underestimated, since it is evaluated on its agreement with  $CO_{TD}$ , which is likely imprecise itself [93,94].

### 3. Methods

#### 3.4.4 Bayesian statistics

The use of Bayesian statistics in Paper 3 is mainly due to the flexibility of Bayesian models, rather than a choice based on the Bayesian *vs* frequentist philosophy of statistics, or a wish to incorporate *prior* knowledge. Still, it is relevant to give a short, general introduction to Bayesian statistics, as many scientists have sparse familiarity with the topic.

Bayesian statistical models can be understood from a generative perspective: we define a model that we believe could generate data similar to what we have observed, if the model parameters are set appropriately. Finding the values for the model parameters, that makes the generated data most similar to the observed data, is called *fitting* the model. We define a model based on our *domain knowledge* about the relation between the predictor- and outcome variables, we specify our *prior* knowledge about the model parameters, and we provide some data we have observed. From these three components, the *model*, the *data* and the *priors*, we can answer the question: *If the data was generated by our proposed model, which sets of parameter values are compatible with the data we have observed?* [95].

Bayesian updating provides the recipe for *fitting* the model, i.e. the act of turning *data* and *prior* knowledge about model parameters into *posterior* knowledge about model parameters (*prior* is before seeing the data, *posterior* is after seeing the data). This is formalised in Bayes theorem:

$$P(\text{parameters}|\text{data}) = \frac{P(\text{data}|\text{parameters})P(\text{parameters})}{P(\text{data})},$$

where,

- *parameters* is a specific set of parameter values,
- $P(\text{parameters})$  is the *prior*: our belief about how plausible the set of parameter values is without seeing the *data*,
- $P(\text{data}|\text{parameters})$  is the *likelihood* distribution: how plausible the data we have observed is given the specific set of parameter values,
- $P(\text{data})$  is the average probability of the data: essentially how likely the data is on average for the entire prior parameter distribution (this is a bit abstract, but it essentially ensures that the posterior distribution will sum to 1).
- $P(\text{parameters}|\text{data})$  is the *posterior* distribution of the model parameters: the plausibility of the parameter values after seeing the data.

### 3. Methods

#### Bayesian updating of a simple model

To gain some intuition about how Bayesian updating works, we will go through fitting a very simple model. We sample subjects from a population and measure their heights. The aim is to estimate the mean and standard deviation (SD) of heights in the population from the heights in the sample.

First, we define a generative model, i.e., a model that—with the right parameter values—will generate data that looks like the data we have collected. In this example, the model is simple: we assume that each subject is drawn from a normal (Gaussian) distribution with a mean,  $\mu$ , and SD,  $\sigma$ :

$$Height_i \sim Normal(\mu, \sigma),$$

where  $Height_i$  is the height of subject  $i$ . To complete the model, we add prior distributions for the two parameters,  $\mu$  and  $\sigma$ :

$$\begin{aligned}\mu &\sim Normal(170, 20) \\ \sigma &\sim Uniform(1, 30).\end{aligned}$$

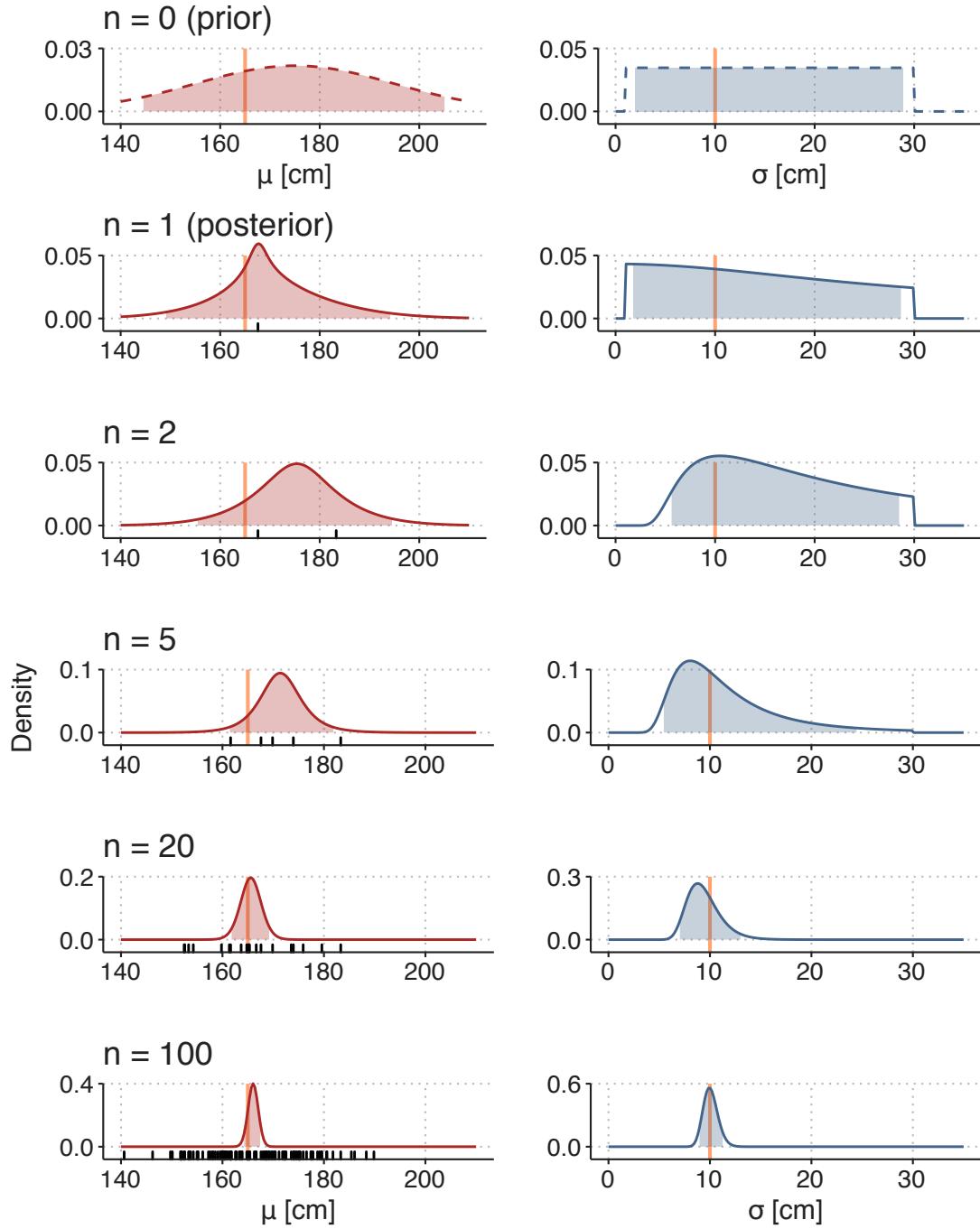
The interpretation of this prior distribution is, that before we see any data, we are quite certain that the mean height in the population ( $\mu$ ) is somewhere between 130 and 210 cm ( $170\text{cm} \pm 2 \cdot 20\text{cm}$ ), and that the SD ( $\sigma$ ) is between 1 and 30 cm, with all values equally probable and values outside this range entirely impossible. These priors are only weakly informative and could be improved without much controversy, but here, their purpose is to illustrate how prior beliefs can be updated through observation of data.

Now, we systematically update our knowledge about the model parameters through observation of more and more data (heights). The top panel-pair in Figure 3.11 displays our prior belief about the model parameters. When we start sampling heights from the population (black ticks in Figure 3.11), we update our beliefs about the parameter values, and the range of compatible parameter values become narrower. In this example, data is simulated, and we know that the *true* mean height ( $\mu$ ) is 165 cm and the *true* SD ( $\sigma$ ) is 10 cm.

In Figure 3.11, the distributions of  $\mu$  and  $\sigma$  are visualised separately. They do, however, constitute a single, multidimensional, distribution, and in some models, parameter values can be highly correlated. Figure 3.12 visualise the two-dimensional distribution of the model parameters.

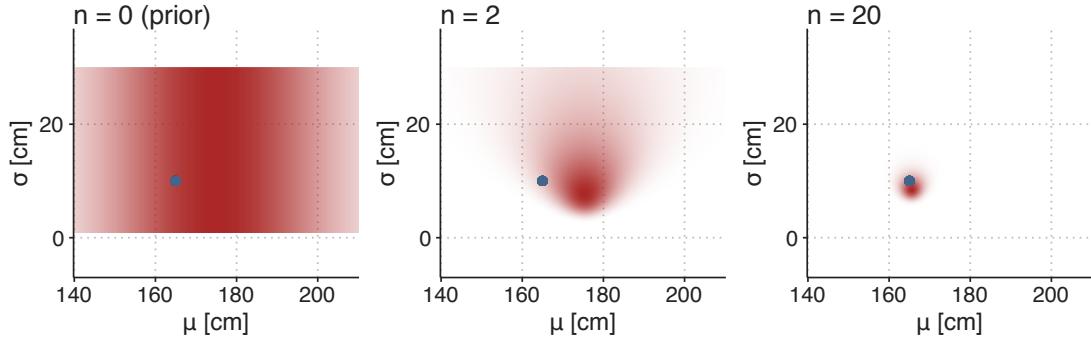
The shaded areas in Figure 3.11 contain 95% of the probability mass for the parameter values. These are traditionally called 95% credible intervals, but I will use the alternative term *compatibility interval* (CI), as they can be interpreted as containing the parameter values that are compatible with the data we observe—this is also a reasonable interpretation of frequentist *confidence intervals* [96].

### 3. Methods



**Figure 3.11:** Illustration of how prior beliefs can be updated through observation of data. See main text for an explanation of the simulation. Orange lines are *true* parameter values. Shaded areas are 95% compatibility intervals. Black ticks are observations.

### 3. Methods



**Figure 3.12:** The prior and posterior distributions are multidimensional distributions with a dimension for each parameter in the model. Color density represent probability density.

#### 3.4.5 The mixed-effects model in Paper 3

A main aim in Paper 3 was to understand how  $V_T$  and RR affect PPV. The experiment had a hierarchical structure, where PPV was measured repeatedly in each subject, with ten different RR and  $V_T$  settings; we also calculated PPV with two different methods ( $PPV_{\text{Classic}}$  and  $PPV_{\text{GAM}}$ ) and wanted to know if  $V_T$  and RR impact PPV differently depending on the method. We used a Bayesian generalized mixed-effects model to find the effects of  $V_T$ , RR and PPV-method on PPV. Below, I will walk through this model.

First, we specified a *likelihood function*. This specifies how likely different PPV values are given a mean,  $\mu$ , and a variance parameter,  $\sigma$ :

$$PPV_{strm} \sim \text{StudentT}(\mu_{strm}, \sigma_{strm}, \text{df} = 4),$$

where  $PPV_{strm}$  is PPV for subject  $s$ ,  $V_T$   $t$ , RR  $r$ , PPV-method  $m$ . We used a Student's T distribution with 4 degrees of freedom rather than a normal distribution, because it puts higher likelihood on extreme observations, and, therefore, is more robust to outliers.

We specified a linear model for  $\mu$ , which, in this model, represents the expected PPV:

$$\log(\mu_{strm}) = \beta_0_m + \beta_1_{tm} + \beta_2_{rm} + \alpha_s.$$

Specifically, the linear model is for  $\log(\mu)$ . Therefore, the exponential of the parameters correspond to relative effects on PPV.  $\exp(\beta_1_{tm})$  and  $\exp(\beta_2_{rm})$  are the effects of each  $V_T$ - and RR setting for each PPV method relative to the expected PPV at  $V_T=10 \text{ ml kg}^{-1}$  and  $RR=10 \text{ min}^{-1}$ .  $\exp(\beta_0_m)$  is the expected PPV at  $V_T=10 \text{ ml kg}^{-1}$  and  $RR=10 \text{ min}^{-1}$  for each method.  $\alpha_s$  is a random effect of each subject. It captures the between-individual variation in PPV.

### 3. Methods

The random effect,  $\alpha$ , was estimated for each subject. We assumed that  $\alpha$  for each subject is drawn from a shared normal distribution with SD  $\sigma_\alpha$ :

$$\alpha_s \sim Normal(0, \sigma_\alpha)$$

This greatly reduces the degrees of freedom associated with estimating a parameter value for each subject. It also reduces the risk of overfitting, as the estimate of  $\alpha$  for subjects with extreme values for PPV is shrunk towards the mean (an effect known as *partial pooling*) [95].

The linear model for  $\sigma$  is similar to the model for  $\mu$ , but without the random effect:

$$\log(\sigma_{trm}) = \gamma_0 m + \gamma_1 t_m + \gamma_2 r_m.$$

It allows the variance of the likelihood function to vary with ventilator setting and PPV-method.

Lastly, priors for all parameters are specified. These are all weakly informative (they allow all parameter values that are even remotely reasonable), and mainly serve to make the parameter estimation more efficient. Arguments for why these values are considered weakly informative are presented in Paper 3's supplementary 1.

**[Prior for SD of subjects]**

$$\sigma_\alpha \sim truncNormal(0, 1.5, trunc_{low} = 0)$$

**[Prior for PPVmethod-specific intercept]**

$$\beta_0 m \sim Normal(2.3, 1), \text{ for PPVmethod } m = (\text{gam, classic})$$

**[Prior for  $\beta$ ]**

$$(\beta_1 t_m, \beta_2 r_m) \sim Normal(0, 2),$$

for ventVT  $t = (8, 6, 4)$ ,

ventRR  $r = (17, 24, 31)$ ,

PPVmethod  $m = (\text{gam, classic})$

**[Prior for  $\gamma$ s]**

$$(\gamma_0 m, \gamma_1 t_m, \gamma_2 r_m) \sim Normal(0, 1.5),$$

for ventVT  $t = (8, 6, 4)$ ,

ventRR  $r = (17, 24, 31)$ ,

PPVmethod  $m = (\text{gam, classic})$

### *3. Methods*

#### **3.4.6 Sample size**

The sample size calculation was done with respect to fluid responsiveness prediction. We expected a fluid response in half the subjects and an AUROC of 0.8. To attain a power of 0.9 with  $\alpha = 0.05$ , 33 patients should be included. To account for uncertainty in the proportion of fluid responders, and to obtain more precise estimates for the effect of ventilator settings on PPV, we planned to include 50 subjects.

*“In statistics, the N’s justify the means.”*

— Unknown

# 4

## Results and discussion

### 4.1 Paper 1 – Most mini-fluid challenge studies overestimate predictive performance

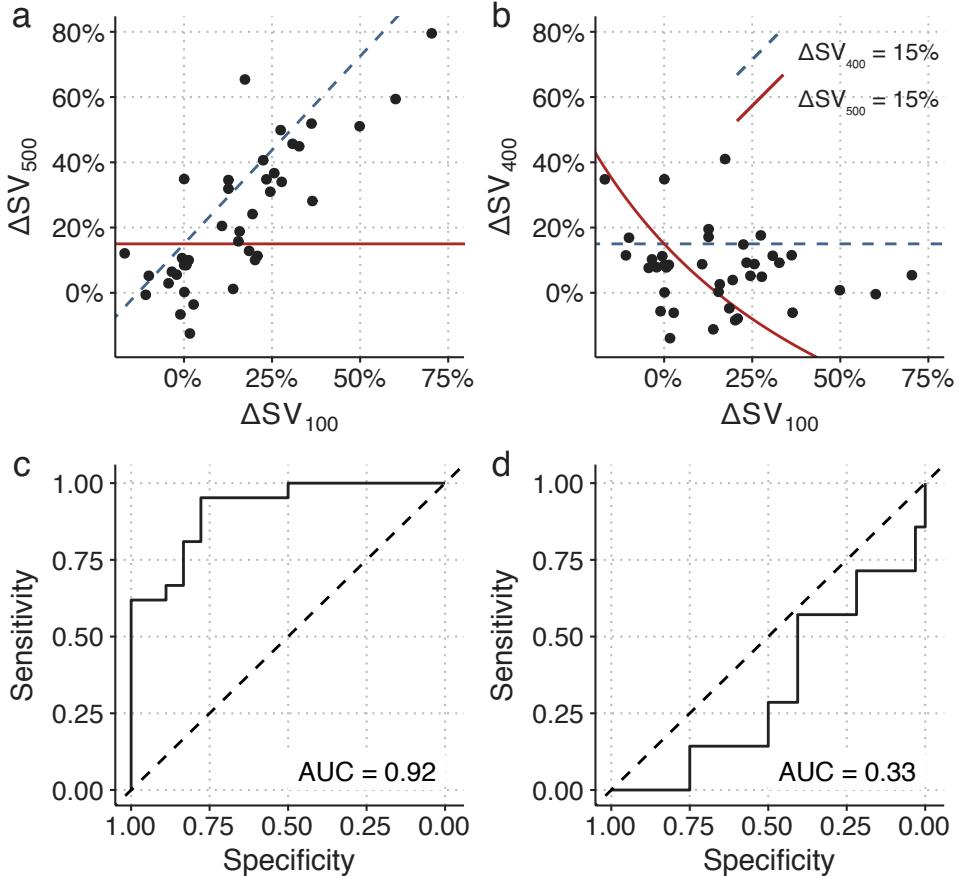
We identify two problems in most MFC studies: First, the predictor (MFC) is included in the outcome (the response to the full fluid challenge). And, second, a shared measurement error between the predictor and outcome will cause a spurious correlation. Both effects lead to an overestimation of the predictive performance of the MFC.

In the reanalysis of data from the study by Muller et al., 2011 [43], we found that the MFC predicted a response to the full fluid challenge ( $\Delta SV_{500} > 15\%$ ) with an AUROC of 0.92, 95%CI [0.83; 1.00] (as reported in the original study), while prediction of the response to the remaining fluid ( $\Delta SV_{400} > 15\%$ ) had an AUROC of 0.33 95% CI [0.11; 0.55]—i.e., worse than flipping a coin. While neither of these reflect the true predictive ability of the MFC, this analysis illustrates the importance of this design choice (see Figure 4.1).

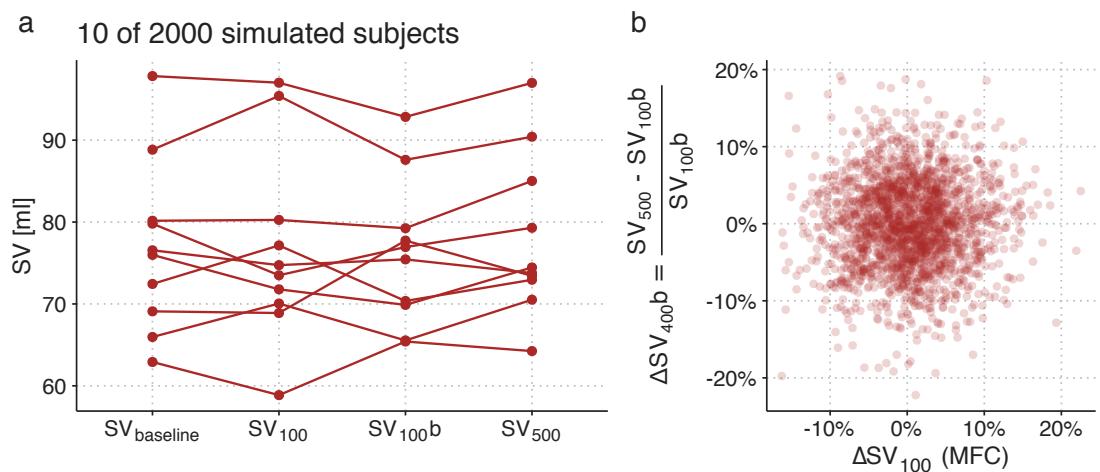
#### 4.1.1 An improved design

The MFC study design used by Guinot et al. [97] avoids including the predictor in the outcome by only predicting the response to the remaining 400 ml fluid, and it avoids the spurious correlation by adding a new baseline SV-measurement before administering the remaining 400 ml fluid. Figure 4.2 illustrates how this design does not cause a spurious correlation between predictor and outcome.

#### 4. Results and discussion



**Figure 4.1:** a) Reconstruction of fig. 3A from Muller et al., 2011 [43]. b) The relationship between the response to  $\Delta SV_{100}$  (MFC) and the response to the remaining 400 ml. c & d) Receiver operating characteristic (ROC) analysis of the ability of  $\Delta SV_{100}$  (MFC) to predict a fluid response  $> 15\%$  for  $\Delta SV_{500}$  and  $\Delta SV_{400}$  respectively.



**Figure 4.2:** Extension of the simulation in Figure 3.3 with a second measurement of SV after the 100 ml fluid. This new baseline measurement removes the spurious correlation between predictor and outcome.

#### *4. Results and discussion*

The design with a new baseline measurement is still not without problems. Essentially, we want to use the MFC response to predict whether a patient will benefit from being given the remaining fluid bolus. However, since all patients receive both fluid boluses, we simply do not know what would happen if the remaining fluid was withheld. The implicit assumption in studies with the problematic design (predicting the full response,  $\Delta SV_{500}$ ) is that SV would return to the baseline level if the remaining fluid was withheld, since this is the level that the final SV measurement is compared against. In the design with a new baseline measurement (used by Guinot et al. [97]), the implicit assumption is that SV would remain at the level of the new baseline ( $SV_{100b}$  in Figure 4.2) if no further fluid challenge was given. But this may not be the case. Imagine that a subject responds to the MFC and retains this new and higher SV after receiving the remaining fluid. This subject would be classified as a non-responder (no benefit), but it could be that in a counterfactual situation, where only the MFC was given, the subject would drop in SV. In that hypothetical situation, the subject may actually have benefitted from the remaining fluid, without directly *responding* to it. By using the design by Guinot et al. [97], we overcome an important statistical problem, but we risk classifying subjects as non-responders, who may have benefited from the remaining fluid challenge. This limitation can be reduced by waiting, e.g., 10 minutes between finishing the MFC and starting the infusion of the remaining fluid. By that time, any large but transient response to the MFC has likely subsided.

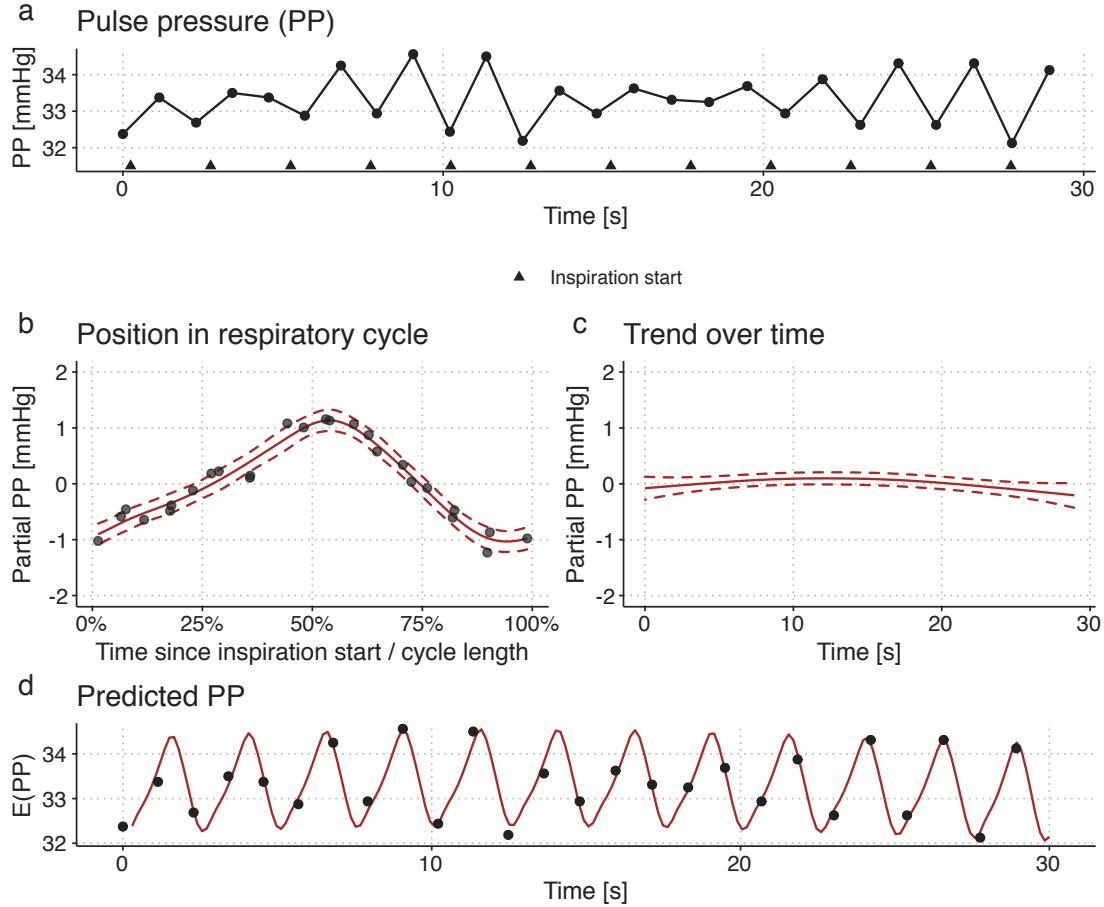
## **4.2 Paper 2 – GAMs can extract PPV at ventilation with low HR/RR and decompose the heart-lung interaction in a CVP waveform**

A main aim with Paper 2 was to introduce the reader to GAMs in a practical manner. The examples in the paper have some intriguing clinical use cases, but we did not attempt to prove any benefit over existing methods. We demonstrated that the information in steady-state recordings of PP and CVP can effectively be represented in a GAM, and that from these fitted models, we can draw physiological inferences that are not directly available in the raw time series or waveform.

### **4.2.1 PPV at ventilation with low HR/RR**

PPV-estimation during low heart-rate-to-respiratory-rate (HR/RR) ventilation is an intriguing example of how a GAM can reveal hidden information. When  $HR/RR < 3.6$ , the *classical* method for PPV estimation is believed to be unreliable [57]. In Paper 2,

#### 4. Results and discussion



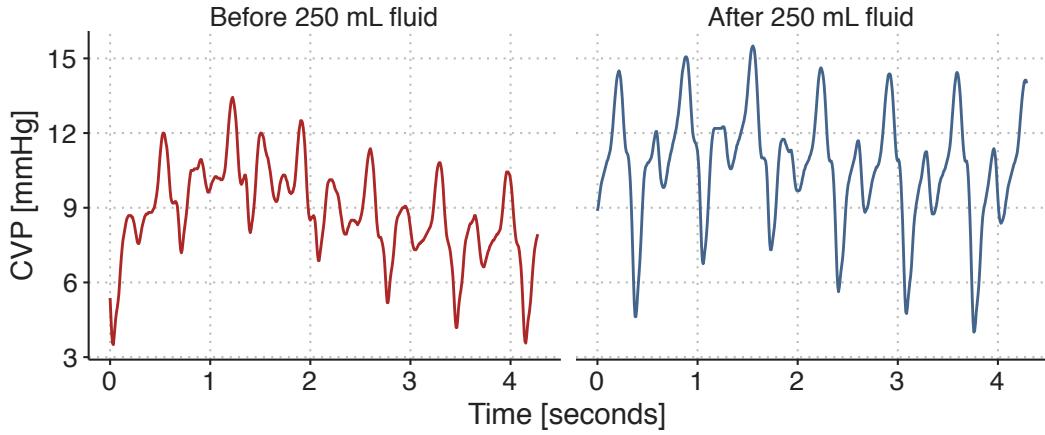
**Figure 4.3:** Illustration of a GAM fitted to a pulse pressure (PP) time series from a patient ventilated with a low HR/RR (2.17 beats per breath). **a)** The observed time series. **b)** A smooth repeating effect of ventilation. **c)** Trend over time. **d)** The models prediction (fit): the sum of the mean PP, the effect of ventilation (**b**) and the trend over time (**c**).

we fit a GAM to a time series of PP recorded while the patient is ventilated with a HR/RR of 2.17—i.e., there are just above two beats per respiratory cycle. When the PP time series is plotted, it seems like PPV is changing (see Figure 4.3a). However, the cyclic effect of ventilation (Figure 4.3b) explains nearly all the variation in PP. From this smooth relationship between a beat’s position in the respiratory cycle and the PP created, we can calculate PPV.

#### 4.2.2 Decomposition of a CVP waveform

The second example in Paper 2 demonstrates how a GAM with just four relatively simple components can represent most of the variation in a steady-state CVP waveform. We recorded a CVP waveform before and after administration of a 250 ml fluid bolus (two short samples of the observed data are presented in Figure 4.4). It is clear that the

#### 4. Results and discussion



**Figure 4.4:** A sample of the CVP waveform recorded before and after administration of 250 ml fluid.

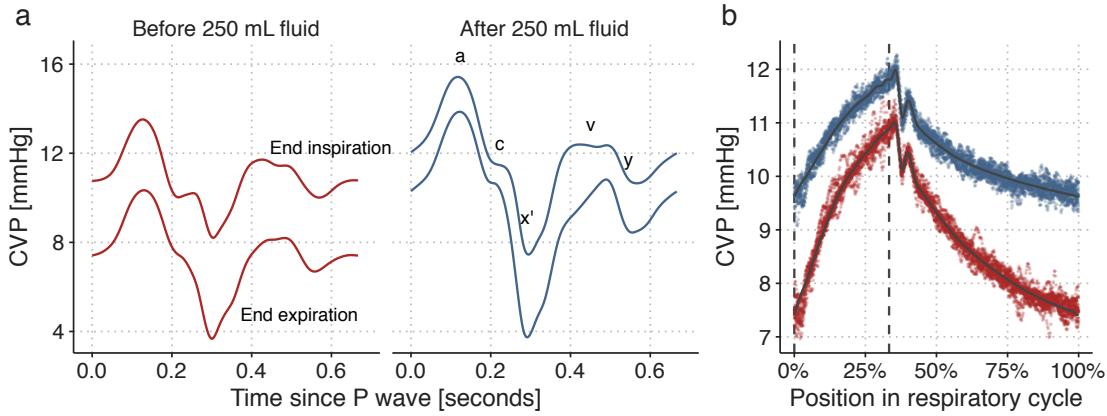
CVP waveform changes with the fluid administration, but it is not easy to precisely describe this change.

By fitting a GAM to the recorded CVP waveform, we can make inferences and visualisations that are not readily available in the raw data. A relatively simple example is the visualisation of how the shape of the CVP waveform changes with the administration of fluid, and how the effect of ventilation on the CVP waveform differs before and after fluid: Figure 4.5a shows the GAM-fit of the CVP through a cardiac cycle at end inspiration and end expiration, both before and after fluid. These curves do not directly exist in the observed waveforms; both because they represent the average cardiac cycle, and because the effect of ventilation is fixed to either end expiration or end inspiration (in the observed waveform, the position in both the cardiac and respiratory cycle change continuously).

The GAM is only fitted to a single subject's CVP, so the conclusions apply only to this single fluid administration in this subject. Figure 4.5a shows that after the fluid bolus, the variation in CVP over a cardiac cycle increased. Conversely, the variation over a respiratory cycle decreased (this is especially clear in Figure 4.5b). The height of the *a* wave (corresponding to the atrial contraction) was highest after fluid, and this height was less variable after fluid. For the *v* wave (the right atrial pressure before the tricuspid valve opens), the interaction between the cardiac and respiratory cycle changed with fluid: the shape of the *v* wave before and after fluid was similar at end inspiration; but at end expiration, the peak of the *v* wave was later and higher, possibly representing ongoing venous return filling the right atrium.

Many hypotheses regarding the shape of a waveform could be investigated and formally tested from GAMs similar to this. For example, it has been proposed that the depth

#### 4. Results and discussion



**Figure 4.5:** Visualisation of a GAM fitted to two 60-second CVP waveforms, before and after a 250 ml fluid bolus. **a)** A cardiac cycle at two fixed positions in the respiratory cycle (end expiration and end inspiration) before and after fluid administration. **b)** The effect of ventilation alone. Coloured dots are partial residuals, illustrating the variation in data not represented in the model. Vertical, dashed, lines show end expiration and end inspiration.

and slope of the  $x'$  descend is related to the strength of the right ventricular contraction (similarly to the ultrasonic measure, *tricuspid annular plane systolic excursion* (TAPSE)) [98]. If this is the case, the respiratory variation in the  $x'$  descend may represent the right ventricular response to changes in afterload, and a variable  $x'$  descend could imply that increases in airway pressure (through *positive end expiratory pressure* (PEEP) or high  $V_T$ ) should be avoided. The respiratory variation in the  $x'$  descend is captured in the cardio-respiratory interaction term (illustrated in Figure 3.9d, approx. 0.4 seconds after the P wave).

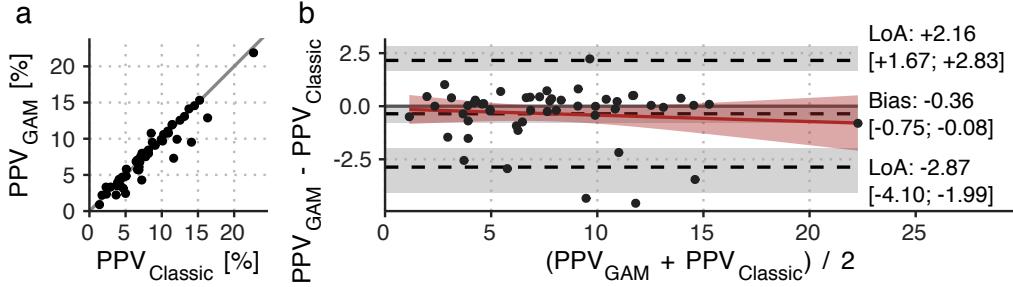
### 4.3 Paper 3 – There is a near-linear relationship between PPV and $V_T$

Paper 3 covers a clinical study of 52 subjects undergoing open abdominal surgery. Of these, 50 received a 250 ml fluid bolus. We had three overall aims for the study: First, we investigated the accuracy of fluid responsiveness prediction for PPV measured at different ventilator settings. Second, we compared a classic method for deriving PPV ( $PPV_{Classic}$ ) [53,57] to the GAM derived PPV presented in Paper 2 ( $PPV_{GAM}$ ). Third, we investigated the effect of  $V_T$  and RR on PPV (*Classic* and *GAM-derived*).

#### 4.3.1 Fluid responsiveness prediction

The fluid responses observed in this study were unfortunately very small. Only ten subjects had an increase in  $SV > 10\%$ , prespecified as a significant fluid response. The

#### 4. Results and discussion



**Figure 4.6:** Comparison of  $PPV_{GAM}$  and  $PPV_{Classic}$  at  $RR=10 \text{ min}^{-1}$ ,  $V_T=10 \text{ ml kg}^{-1}$ . **a)** scatter plot; **b)** Bland-Altman plot. LoA, 95% limits of agreement. Square brackets contain 95%CI.

low number of responders made the estimates of AUROC imprecise. For example, the AUROC for  $PPV_{GAM}$  at  $RR=10 \text{ min}^{-1}$ ,  $V_T=10 \text{ ml kg}^{-1}$  was 0.73 95%CI [0.57; 0.90]. I.e., the data is compatible with both poor and excellent fluid responsiveness prediction. Generally, the point estimate for AUROC fell with higher RR and lower  $V_T$ . The results were similar for  $PPV_{Classic}$ . All ROC analyses are shown in [Paper 3 - Supplementary 3 Fig. S1](#).

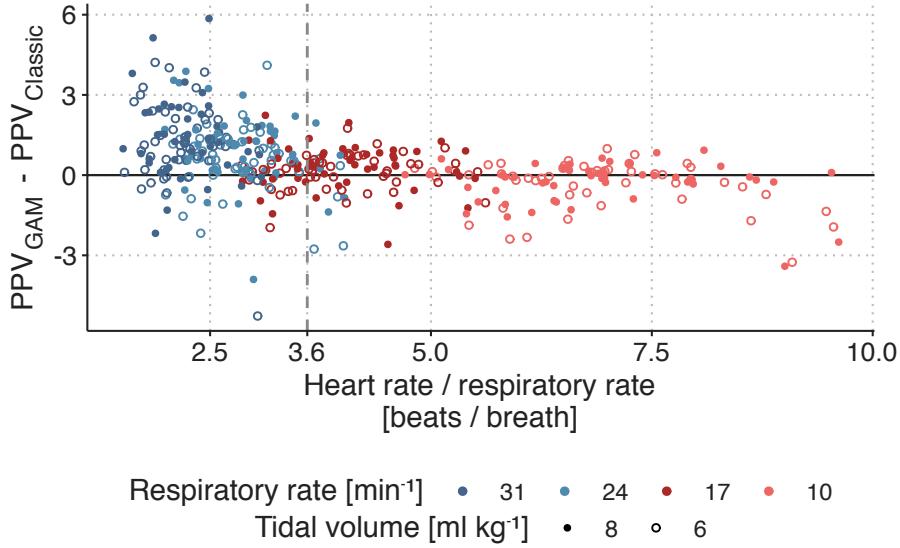
#### 4.3.2 Comparison of $PPV_{Classic}$ and $PPV_{GAM}$

We used Bland-Altman plots to compare  $PPV_{Classic}$  and  $PPV_{GAM}$  for each ventilator setting. The comparison for  $RR=10 \text{ min}^{-1}$ ,  $V_T=10 \text{ ml kg}^{-1}$  showed good agreement between the methods, though three subjects had a  $PPV_{Classic}$  that was three to five %-points higher than their  $PPV_{GAM}$  (see Figure 4.6). These few discrepancies may reflect that  $PPV_{Classic}$  is sensitive to outlying beats, compared to  $PPV_{GAM}$ , which is fitted to all beats and constrained to fit a smooth, cyclic variation.

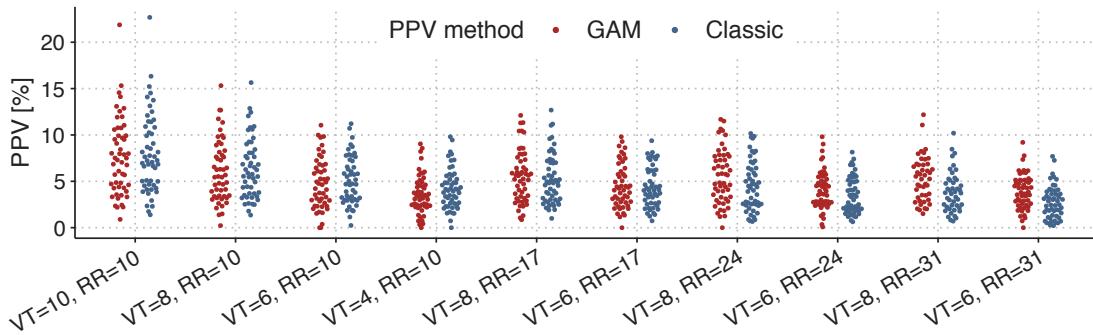
For  $V_T$ 's 8 and 6  $\text{ml kg}^{-1}$ , comparable results were observed for the agreement (see [Paper 3 - Supplementary 3 Fig. S3](#)), but at  $V_T=4 \text{ ml kg}^{-1}$ , agreement was markedly worse with a bias of -0.64, 95%CI [-1.17; -0.31], and 95% limits of agreement (LoA) from -3.71 to +2.43 (while this is similar to the LoA at  $RR=10 \text{ min}^{-1}$ ,  $V_T=10 \text{ ml kg}^{-1}$ , the mean PPV has halved, making the relative agreement worse).

The agreement between  $PPV_{Classic}$  and  $PPV_{GAM}$  was lower with higher RR. Indicating that HR/RR have different effects on  $PPV_{Classic}$  and  $PPV_{GAM}$ . In Figure 4.7, we present the relationship between HR/RR and the difference between  $PPV_{Classic}$  and  $PPV_{GAM}$ . At high HR/RR, there was close agreement between the methods, but with lower HR/RR, precision decreased and  $PPV_{GAM}$  was generally higher than  $PPV_{Classic}$ .

#### 4. Results and discussion



**Figure 4.7:** Difference between  $PPV_{Classic}$  and  $PPV_{GAM}$  by HR/RR



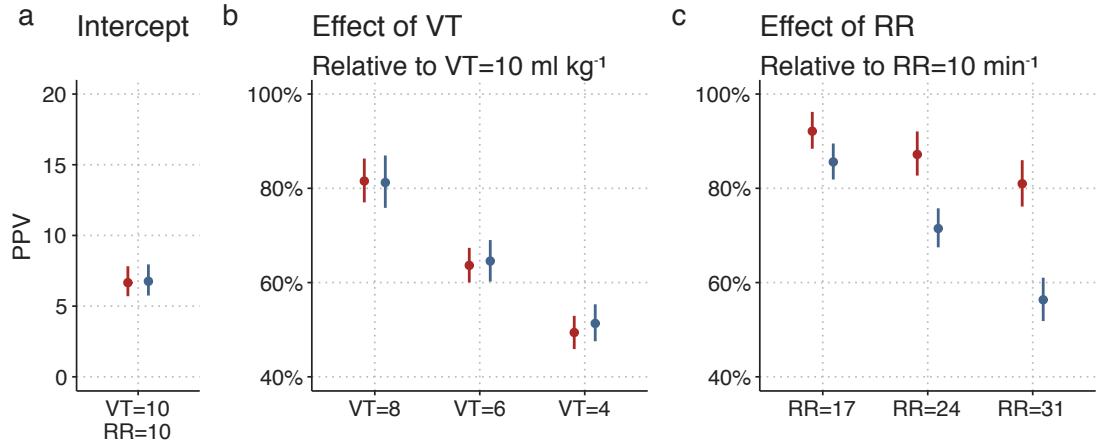
**Figure 4.8:** PPV (GAM and Classic) for all ventilator settings in 52 subjects ( $n = 507$  for both methods).

#### 4.3.3 The effect of ventilator settings on PPV

With 52 subjects and ten ventilator settings for each, we had 520 potential windows for PPV analysis. Of these, 507 were eligible for analysis. Exclusions were due to either 3 extrasystoles within a window ( $n=7$ ) or because the setting with  $RR=31 \text{ min}^{-1}$ ,  $V_T=8 \text{ ml kg}^{-1}$  was omitted when the airway pressure was expected to exceed  $40 \text{ cmH}_2\text{O}$  ( $n=4$ ). For all eligible windows, both  $PPV_{Classic}$  and  $PPV_{GAM}$  were calculated (see Figure 4.8).

Figure 4.9 shows the posterior estimates for the Bayesian mixed model fitted to all 1014 PPV estimates. We see that PPV is nearly proportional to  $V_T$  for both  $PPV_{Classic}$  and  $PPV_{GAM}$ . A direct proportionality implies that PPV at  $V_T=4 \text{ ml kg}^{-1}$  would be 40% of PPV at  $V_T=10 \text{ ml kg}^{-1}$ , but we see in Figure 4.9b that the expected PPV is slightly higher (approx. 50%).

#### 4. Results and discussion



**Figure 4.9:** Posterior distributions (estimate and 95%CI) for parameters of the Bayesian mixed model fitted to PPV values for different ventilator settings. The intercept is the median PPV at  $RR=10 \text{ min}^{-1}$ ,  $V_T=10 \text{ ml kg}^{-1}$  across subjects. The effects of RR and  $V_T$  are multiplied to give the expected relative differences between the PPV at  $RR=10 \text{ min}^{-1}$ ,  $V_T=10 \text{ ml kg}^{-1}$  and the PPV at other ventilator settings within a single subject.

The effect of RR differs markedly between the two PPV methods: at  $RR=31 \text{ min}^{-1}$ ,  $PPV_{GAM}$  is reduced by approx. 20% while  $PPV_{Classic}$  is reduced by more than 40% relative to at  $RR=10 \text{ min}^{-1}$ . This difference matches what we see in Figure 4.7—at high RR,  $PPV_{GAM}$  is generally higher than  $PPV_{Classic}$ .

A probable explanation for the strong effect of RR on  $PPV_{Classic}$  is, that with high RR, we often have low HR/RR. This creates a sampling problem as visualised in Figure 4.3, where  $PPV_{Classic}$  will underestimate the “true” PPV, since only some respiratory cycles will have beats occurring at the maximum and minimum effect of ventilation; simply because there are few beats per cycle. The fall in  $PPV_{GAM}$  with increasing RR may reflect a true physiologic effect of RR on PPV. It could also reflect a physiologic effect of HR/RR is present, though smaller than originally suggested by De Backer et al., 2009 [57]. From this study, we cannot determine if the effect on  $PPV_{GAM}$  is from RR directly or from HR/RR.

We can use the results from this model to account for ventilator settings when interpreting PPV. The optimal PPV threshold for predicting a response to fluid seems to be around 12% at  $V_T=10 \text{ ml kg}^{-1}$  [55]; with this  $V_T$ , RR will usually be around  $10 \text{ min}^{-1}$ . Instead of running studies to investigate the optimal threshold for all combinations of RR and  $V_T$ , we can use the results from our model to calculate a *best guess* for what PPV would be, if the ventilator settings were changed to  $RR=10 \text{ min}^{-1}$ ,  $V_T=10 \text{ ml kg}^{-1}$ . As an example, if a patient is ventilated with  $RR=24 \text{ min}^{-1}$  and  $V_T=6 \text{ ml kg}^{-1}$ , and we measure a PPV

#### 4. Results and discussion

of 8% using the GAM method, we can calculate a *best guess* for PPV at RR=10 min<sup>-1</sup>, V<sub>T</sub>=10 ml kg<sup>-1</sup> with the following formula:

$$PPV_{RR=10, V_T=10} = 8\% \cdot \frac{1}{0.64} \cdot \frac{1}{0.87} \approx 14\%,$$

where 0.64 is the model estimate for the relative effect of ventilation with V<sub>T</sub>=6 ml kg<sup>-1</sup>, and 0.87 is the relative effect of RR=24 min<sup>-1</sup>.

For practical use, we can simplify the formula for correcting *PPV<sub>GAM</sub>* by simply assuming a direct proportionality between *PPV<sub>GAM</sub>* and V<sub>T</sub>. The reason this approximation works is, that while *PPV<sub>GAM</sub>* does not change in exact proportion to V<sub>T</sub>, an increase in V<sub>T</sub> is usually accompanied with lowering RR to ensure normoventilation. Lowering RR happens to increase *PPV<sub>GAM</sub>* by the amount needed to make *PPV<sub>GAM</sub>* change in proportion to the change in V<sub>T</sub>. For example, a patient, who is normoventilated at RR=10 min<sup>-1</sup>, V<sub>T</sub>=10 ml kg<sup>-1</sup>, is also approximately normoventilated at RR=31 min<sup>-1</sup>, V<sub>T</sub>=4 ml kg<sup>-1</sup>. The effects on *PPV<sub>GAM</sub>* of these changes are 81% for the RR change and 50% for the V<sub>T</sub> change. Therefore, the combined effect of the change in ventilator settings is 81% · 50% = 40.5%, corresponding to a *PPV<sub>GAM</sub>* change in proportion to the V<sub>T</sub> change.

##### 4.3.4 Limitations

We included patients in a period where the treating anesthesiologist expected that no hemodynamic intervention was needed in the next 15 minutes, to ensure that we could complete the study protocol. This may explain the low number of fluid responders and the correspondingly low PPV values. We do not know whether the reported effects of ventilator settings on PPV can be generalised to a more fluid responsive population with higher PPV values.

Ventilation with high RR (e.g. 31 min<sup>-1</sup>) is rarely used for surgical anaesthesia. It is unknown whether the results for the present intraoperative patient population apply to e.g. the setting of ARDS, where high RR is used, but where lung mechanics can be markedly affected by the underlying condition.

While our results provide a strong indication that the low HR/RR limitation may be overcome with the GAM method, the generally poor predictive performance of PPV in this study made a direct investigation of this impossible. It is plausible that the GAM method accurately measures PPV at low HR/RR, but this PPV is physiologically different from PPV at high HR/RR and does not reflect fluid responsiveness.

*“Statistics used to be seen as analysing data, the combination of numbers. I’ve come to see that as the least important part. The difficult bit is how you design a study, collect the data, avoid bias and provide an honest representation of what you found.”*

— Douglas Altman

# 5

## Conclusion and future research

With this PhD I have tackled challenges in fluid responsiveness prediction during surgery. This was done through three different types of papers: Paper 1 critiques a specific research method, Paper 2 introduces a statistical method and presents novel uses that may be relevant for fluid responsiveness prediction, and Paper 3 is a clinical experiment.

**Paper 1** unravels a methodological problem associated with a specific method for predicting fluid responsiveness: the MFC. The bias that occurs in these studies due to shared measurement error, is a common pitfall in studies of *change* in any measurement. While the problem is especially clear in the MFC studies, this statistical issue should be kept in mind when designing any study where *change* is an outcome of interest.

**Paper 2** introduces GAMs as a tool for analysing medical time series and waveforms. Two motivated examples are presented. The first is relatively simple and has a direct clinical use: by modelling a PP time series with a GAM, we can derive PPV in a robust manner that seems to work even with few beats per ventilation (low HR/RR). In the second example, we demonstrate that a GAM can be used to decompose a CVP waveform into cardiac and ventilatory components.

**Paper 3** presents a clinical study of the effects of  $V_T$  and RR on PPV. The results indicate that using a GAM derived PPV may enable fluid responsiveness prediction despite low HR/RR, and that it may be valuable to correct PPV for  $V_T$ .

## 5. Conclusion and future research

### 5.1 Future research

Future studies of the MFC should take great care to minimise the problems presented in Paper 1. I am aware of only one MFC study published after the publication of Paper 1 [99]. This study used the same problematic design as previous MFC studies—reportedly to ease comparison with previous MFC studies. They do, however, acknowledge that the design might be problematic.

Also, it may be possible to do an unbiased reanalysis of data from the existing MFC studies with only three SV measurements. I do not have a complete recipe for this—nor the data—but a simple regression model of absolute SV measurements, rather than changes ( $\Delta SV$ ), is probably a good starting point. Example:

$$SV_{500} = \beta_1 \cdot SV_{100} + \beta_2 \cdot SV_{baseline}.$$

This is essentially a second-order autoregressive model. The part that is still unclear to me, is how to turn the fit of this model into a simple clinical tool and how to report its accuracy.

The GAM-method for estimating PPV, presented in Paper 2, should be evaluated as a predictor of fluid responsiveness in future studies. To include patients with a higher average fluid response, the study protocol should be shorter than what was used in Paper 3. A fast protocol, with e.g. a two minute baseline measurement of both SV and PPV, immediately followed by a fluid challenge, would allow evaluation in a more hemodynamically unstable state. This setup could be run in an ICU, with ARDS patients on lung-protective ventilation, to assess the performance of GAM-derived PPV in a setting with low HR/RR.

The data presented in Paper 3 contains four  $V_T$  challenges for each subject (a  $V_T$  challenge analyses the effect on PPV of increasing  $V_T$  from 6 to 8  $\text{ml kg}^{-1}$ ). There is currently a medical Master’s project using this data to investigate the effect of RR on the response to the  $V_T$  challenge.

In Paper 2, we also describe how a GAM can be used to decompose a waveform. A proposed use case of this technique is that the  $x'$  descend of the CVP may represent the right ventricle’s contractility, and that a high respiratory variation in this feature could indicate that the right ventricle is sensitive to changes in afterload. It would be interesting to investigate the association between the slope of the  $x'$  descend and TAPSE (ultrasound derived) before and after an increase in PEEP, and whether the respiratory variation in the  $x'$  descend could predict the hemodynamic effect of the PEEP increase.

## 5. Conclusion and future research

### 5.2 General challenges in intraoperative fluid responsiveness prediction

The accuracy of any fluid responsiveness study hinges on the accuracy of the CO or SV measurement. These measurements are generally not very accurate (as described in Section 3.4.3), and uncertainty of the outcome sets a hard limit to how accurate a predictor can be. It would be highly valuable to do rigorous investigations of the agreement between different methods for measuring CO, in a setting where we can trust the *gold standard*, e.g. using Fick’s principle for CO estimation in a very stable setting or using magnetic resonance velocity mapping [100].

In fluid responsiveness studies, the outcome is generally dichotomised into *fluid responders* (e.g.  $\Delta SV > 15\%$ ) and *non-responders*. This is probably done to make the results simpler to interpret, but the choice of the threshold is somewhat arbitrary, and dichotomisation decreases the power of the analysis [101]. It is possible to have the power of the continuous analysis, but report the prediction accuracy dichotomously, by, e.g., reporting the PPV value where 80% of the prediction interval for  $\Delta SV$  is above 15%.<sup>1</sup> It would also be interesting to start analysing fluid responsiveness prediction studies using absolute measurements of flow rather than changes. Changes are notoriously difficult to work with [102,103], and relative changes (e.g. a 15% increase from baseline) can be especially problematic [104]. If we use a regression model to predict the body weight indexed SV (SVI) after a fluid challenge using PPV and the baseline SVI as predictors, we can still evaluate the predictive ability of PPV, and we can investigate whether the baseline SVI level modifies this predictive ability.

It is important to be aware that the predictive performance, e.g. AUROC, reported from a fluid responsiveness prediction study is not simply a measure of the test, but a measure of the test *in that setting* [105]. In some fluid responsiveness studies there is a wide, or even bimodal, distribution of fluid responses, with changes in SV or CO ranging from  $-10\%$  to  $50\%$ , and the predictor of choice (e.g. PPV) accurately distinguishes fluid responders from non-responders [53,54,106,107]. On the other hand, in studies where most subjects have a fluid response close to the threshold (e.g. a  $\Delta SV$  of 15%), responders and non-responders are more similar, and the AUROC will usually be lower. The authors may then conclude that the investigated predictor did not work, even though the results from the two types of studies may be entirely compatible, with one study essentially investigating the most difficult subset of the other study. This would be clearer, if we compared parameters

---

<sup>1</sup>We did not do this in Paper 3, because the quality and distribution of our fluid response measurement was not good enough to get sufficiently meaningful results for prediction accuracy. We presented the standard analysis, because it was registered as the primary outcome of the study.

## *5. Conclusion and future research*

from continuous regression analyses across studies—rather than comparing dichotomous classification characteristics like AUROC, sensitivity and specificity.

## References

1. Lewins R. Injection of Saline Solutions into the Veins. *BOSTON MEDICAL AND SURGICAL JOURNAL*. 1832;VI(24):373-374.
2. Latta T. 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*. 1832;18(457):274-280. doi:[10.1016/S0140-6736\(02\)80289-6](https://doi.org/10.1016/S0140-6736(02)80289-6)
3. Cosnett JE. The origins of intravenous fluid therapy. *Lancet (London, England)*. 1989;1(8641):768-771. doi:[10.1016/s0140-6736\(89\)92583-x](https://doi.org/10.1016/s0140-6736(89)92583-x)
4. Petroianu GA. On saline infusion, clonus, molecules and forgotten scientists: Who was Dr Julius Sander (1840–1909)? *Journal of Medical Biography*. Published online December 2021:09677720211065357. doi:[10.1177/09677720211065357](https://doi.org/10.1177/09677720211065357)
5. Foëx BA. How the cholera epidemic of 1831 resulted in a new technique for fluid resuscitation. *Emergency Medicine Journal*. 2003;20(4):316-318. doi:[10.1136/emj.20.4.316](https://doi.org/10.1136/emj.20.4.316)
6. Kadet A. House Calls for Hangovers. *Wall Street Journal*. Published online August 2015.
7. Miller TE, Myles PS. Perioperative Fluid Therapy for Major Surgery. *Anesthesiology*. 2019;130(5):825-832. doi:[10.1097/ALN.0000000000002603](https://doi.org/10.1097/ALN.0000000000002603)
8. Monnet X, Teboul JL. My patient has received fluid. How to assess its efficacy and side effects? *Annals of Intensive Care*. 2018;8(1):54. doi:[10.1186/s13613-018-0400-z](https://doi.org/10.1186/s13613-018-0400-z)
9. Bhave G, Neilson EG. Volume Depletion Versus Dehydration: How Understanding the Difference Can Guide Therapy. *American Journal of Kidney Diseases*. 2011;58(2):302-309. doi:[10.1053/j.ajkd.2011.02.395](https://doi.org/10.1053/j.ajkd.2011.02.395)
10. Jacob M, Chappell D, Rehm M. The 'third space'-fact or fiction? *Best Practice & Research Clinical Anaesthesiology*. 2009;23(2):145-157. doi:[10.1016/j.bpa.2009.05.001](https://doi.org/10.1016/j.bpa.2009.05.001)
11. Coe Aj, Revanäs B. Is crystalloid preloading useful in spinal anaesthesia in the elderly? *Anaesthesia*. 1990;45(3):241-243. doi:[10.1111/j.1365-2044.1990.tb14696.x](https://doi.org/10.1111/j.1365-2044.1990.tb14696.x)

## References

12. Smith I, Kranke P, Murat I, et al. Perioperative fasting in adults and children: Guidelines from the European Society of Anaesthesiology. *European Journal of Anaesthesiology / EJA*. 2011;28(8):556-569. doi:[10.1097/EJA.0b013e3283495ba1](https://doi.org/10.1097/EJA.0b013e3283495ba1)
13. Myburgh JA, Mythen MG. Resuscitation Fluids. *New England Journal of Medicine*. 2013;369(13):1243-1251. doi:[10.1056/NEJMra1208627](https://doi.org/10.1056/NEJMra1208627)
14. Hoste EA, Maitland K, Brudney CS, et al. Four phases of intravenous fluid therapy: A conceptual model†. *British journal of anaesthesia*. 2014;113(5):740-747. doi:[10.1093/bja/aeu300](https://doi.org/10.1093/bja/aeu300)
15. Malbrain MLNG, Langer T, Annane D, et al. Intravenous fluid therapy in the perioperative and critical care setting: Executive summary of the International Fluid Academy (IFA). *Annals of Intensive Care*. 2020;10(1):64. doi:[10.1186/s13613-020-00679-3](https://doi.org/10.1186/s13613-020-00679-3)
16. Perner A, Haase N, Guttormsen AB, et al. Hydroxyethyl Starch 130/0.42 versus Ringer's Acetate in Severe Sepsis. *New England Journal of Medicine*. 2012;367(2):124-134. doi:[10.1056/NEJMoa1204242](https://doi.org/10.1056/NEJMoa1204242)
17. Frank O. On the dynamics of cardiac muscle (1959 translation of Zur dynamik des Herzmuskels). *American Heart Journal*. 1895;58(2):282-317. doi:[10.1016/0002-8703\(59\)90345-X](https://doi.org/10.1016/0002-8703(59)90345-X)
18. Patterson SW, Piper H, Starling EH. The regulation of the heart beat. *The Journal of Physiology*. 1914;48(6):465-513. doi:[10.1113/jphysiol.1914.sp001676](https://doi.org/10.1113/jphysiol.1914.sp001676)
19. Zimmer HG. Who Discovered the Frank-Starling Mechanism? *Physiology*. 2002;17(5):181-184. doi:[10.1152/nips.01383.2002](https://doi.org/10.1152/nips.01383.2002)
20. Boron WF, Boulpaep EL. The Heart as a Pump. In: *Medical Physiology*. Third edition. Elsevier Health Sciences; 2016.
21. Magder S. Volume and its relationship to cardiac output and venous return. *Critical Care*. 2016;20(1):1-11. doi:[10.1186/s13054-016-1438-7](https://doi.org/10.1186/s13054-016-1438-7)
22. Persichini R, Lai C, Teboul JL, Adda I, Guérin L, Monnet X. Venous return and mean systemic filling pressure: Physiology and clinical applications. *Critical Care*. 2022;26(1):150. doi:[10.1186/s13054-022-04024-x](https://doi.org/10.1186/s13054-022-04024-x)
23. Guyton AC, Lindsey AW, Abernathy B, Richardson T. Venous Return at Various Right Atrial Pressures and the Normal Venous Return Curve. *American Journal of Physiology-Legacy Content*. 1957;189(3):609-615. doi:[10.1152/ajplegacy.1957.189.3.609](https://doi.org/10.1152/ajplegacy.1957.189.3.609)
24. Permutt S, Riley RL. Hemodynamics of Collapsible Vessels With Tone: The Vascular Waterfall. *Journal of applied physiology (Bethesda, Md : 1985)*. 1963;18:924-932. doi:[10.1152/jappl.1963.18.5.924](https://doi.org/10.1152/jappl.1963.18.5.924)

## References

25. Wong F, Sniderman K, Liu P, Allidina Y, Sherman M, Blendis L. Transjugular intrahepatic portosystemic stent shunt: Effects on hemodynamics and sodium homeostasis in cirrhosis and refractory ascites. *Annals of Internal Medicine*. 1995;122(11):816-822. doi:[10.7326/0003-4819-122-11-199506010-00002](https://doi.org/10.7326/0003-4819-122-11-199506010-00002)
26. Dunn JO, Mythen M, Grocott M. Physiology of oxygen transport. *BJA Education*. 2016;16(10):341-348. doi:[10.1093/bjaed/mkw012](https://doi.org/10.1093/bjaed/mkw012)
27. Boron WF, Boulpaep EL. The Microcirculation. In: *Medical Physiology*. Third edition. Elsevier Health Sciences; 2016.
28. Weinbaum S, Tarbell JM, Damiano ER. The Structure and Function of the Endothelial Glycocalyx Layer. *Annual Review of Biomedical Engineering*. 2007;9(1):121-167. doi:[10.1146/annurev.bioeng.9.060906.151959](https://doi.org/10.1146/annurev.bioeng.9.060906.151959)
29. Milford EM, Reade MC. Resuscitation Fluid Choices to Preserve the Endothelial Glycocalyx. *Critical Care*. 2019;23(1):1-11. doi:[10.1186/s13054-019-2369-x](https://doi.org/10.1186/s13054-019-2369-x)
30. Chappell D, Bruegger D, Potzel J, et al. Hypervolemia increases release of atrial natriuretic peptide and shedding of the endothelial glycocalyx. *Critical Care*. 2014;18(5):1-8. doi:[10.1186/s13054-014-0538-5](https://doi.org/10.1186/s13054-014-0538-5)
31. Damén T, Kolsrud O, Dellgren G, Hesse C, Ricksten SE, Nygren A. Atrial natriuretic peptide does not degrade the endothelial glycocalyx: A secondary analysis of a randomized porcine model. *Acta Anaesthesiologica Scandinavica*. 2021;65(9):1305-1312. doi:[10.1111/aas.13853](https://doi.org/10.1111/aas.13853)
32. Iba T, Levy JH. Derangement of the endothelial glycocalyx in sepsis. *Journal of Thrombosis and Haemostasis*. 2019;17(2):283-294. doi:[10.1111/jth.14371](https://doi.org/10.1111/jth.14371)
33. Hahn RG, Lyons G. The half-life of infusion fluids: An educational review. *European Journal of Anaesthesiology / EJA*. 2016;33(7):475-482. doi:[10.1097/EJA.0000000000000436](https://doi.org/10.1097/EJA.0000000000000436)
34. Myles PS, Bellomo R, Corcoran T, et al. Restrictive versus Liberal Fluid Therapy for Major Abdominal Surgery. *New England Journal of Medicine*. 2018;378(24):2263-2274. doi:[10.1056/NEJMoa1801601](https://doi.org/10.1056/NEJMoa1801601)
35. Brandstrup B. Finding the Right Balance. *New England Journal of Medicine*. 2018;378(24):2335-2336. doi:[10.1056/NEJMe1805615](https://doi.org/10.1056/NEJMe1805615)
36. Pearse RM, Harrison DA, MacDonald N, et al. Effect of a Perioperative, Cardiac Output-Guided Hemodynamic Therapy Algorithm on Outcomes Following Major Gastrointestinal Surgery: A Randomized Clinical Trial and Systematic Review. *JAMA*. 2014;311(21):2181-2190. doi:[10.1001/jama.2014.5305](https://doi.org/10.1001/jama.2014.5305)
37. Jessen MK, Vallentin MF, Holmberg MJ, et al. Goal-directed haemodynamic therapy during general anaesthesia for noncardiac surgery: A systematic review and meta-analysis. *British Journal of Anaesthesia*. 2022;128(3):416-433. doi:[10.1016/j.bja.2021.10.046](https://doi.org/10.1016/j.bja.2021.10.046)

## References

38. Edwards MR, Forbes G, MacDonald N, et al. Optimisation of Perioperative Cardiovascular Management to Improve Surgical Outcome II (OPTIMISE II) trial: Study protocol for a multicentre international trial of cardiac output-guided fluid therapy with low-dose inotrope infusion compared with usual care in patients undergoing major elective gastrointestinal surgery. *BMJ Open*. 2019;9(1):e023455. doi:[10.1136/bmjopen-2018-023455](https://doi.org/10.1136/bmjopen-2018-023455)
39. Marik PE, Cavallazzi R. Does the Central Venous Pressure Predict Fluid Responsiveness? An Updated Meta-Analysis and a Plea for Some Common Sense\*. *Critical Care Medicine*. 2013;41(7):1774-1781. doi:[10.1097/CCM.0b013e31828a25fd](https://doi.org/10.1097/CCM.0b013e31828a25fd)
40. Monnet X, Letierce A, Hamzaoui O, et al. Arterial pressure allows monitoring the changes in cardiac output induced by volume expansion but not by norepinephrine\*. *Critical Care Medicine*. 2011;39(6):1394-1399. doi:[10.1097/CCM.0b013e31820edcf0](https://doi.org/10.1097/CCM.0b013e31820edcf0)
41. De Backer D, Heenen S, Piagnerelli M, Koch M, Vincent JL. Pulse pressure variations to predict fluid responsiveness: Influence of tidal volume. *Intensive Care Medicine*. 2005;31(4):517-523. doi:[10.1007/s00134-005-2586-4](https://doi.org/10.1007/s00134-005-2586-4)
42. Monnet X, Teboul JL. Passive leg raising: Five rules, not a drop of fluid! *Critical Care*. 2015;19(1):1-3. doi:[10.1186/s13054-014-0708-5](https://doi.org/10.1186/s13054-014-0708-5)
43. Muller L, Toumi M, Bousquet PJ, 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(3):541-547. doi:[10.1097/ALN.0b013e318229a500](https://doi.org/10.1097/ALN.0b013e318229a500)
44. 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. *Critical Care*. 2019;23(1):264. doi:[10.1186/s13054-019-2545-z](https://doi.org/10.1186/s13054-019-2545-z)
45. Vistisen ST, Enevoldsen JN, Greisen J, Juhl-Olsen P. What the anaesthesiologist needs to know about heart-lung interactions. *Best Practice & Research Clinical Anaesthesiology*. Published online 2019. doi:[10.1016/j.bpa.2019.05.003](https://doi.org/10.1016/j.bpa.2019.05.003)
46. Michard F. Changes in Arterial Pressure during Mechanical Ventilation. *Anesthesiology*. 2005;103(2):419-428. doi:[10.1097/00000542-200508000-00026](https://doi.org/10.1097/00000542-200508000-00026)
47. Chemla D, Hebert JL, Coirault C, et al. Total arterial compliance estimated by stroke volume-to-aortic pulse pressure ratio in humans. *American Journal of Physiology*. 1998;274(2 Pt 2):H500-H505. doi:[10.1152/ajpheart.1998.274.2.H500](https://doi.org/10.1152/ajpheart.1998.274.2.H500)
48. Perel A, Pizov R, Cotev S. Systolic Blood Pressure Variation is a Sensitive Indicator of Hypovolemia in Ventilated Dogs Subjected to Graded Hemorrhage. *Anesthesiology*. 1987;67(4):498-502. doi:[10.1097/00000542-198710000-00009](https://doi.org/10.1097/00000542-198710000-00009)
49. Pizov R, Ya'ari Y, Perel A. Systolic pressure variation is greater during hemorrhage than during sodium nitroprusside-induced hypotension in ventilated dogs. *Anesthesia and Analgesia*. 1988;67(2):170-174.

## References

50. Coriat P, Vrillon M, Perel A, et al. A Comparison of Systolic Blood Pressure Variations and Echocardiographic Estimates of End-Diastolic Left Ventricular Size in Patients After Aortic Surgery. *Anesthesia & Analgesia*. 1994;78(1):46-53.
51. Rooke GA, Schwid HA, Shapira Y. The Effect of Graded Hemorrhage and Intravascular Volume Replacement on Systolic Pressure Variation in Humans During Mechanical and Spontaneous Ventilation. *Anesthesia & Analgesia*. 1995;80(5):925-932.
52. Tavernier B, Makhotine O, Lebuffe G, Dupont J, Scherpereel P. Systolic Pressure Variation as a Guide to Fluid Therapy in Patients with Sepsis-induced Hypotension. *Anesthesiology*. 1998;89(6):1313-1321. doi:[10.1097/00000542-199812000-00007](https://doi.org/10.1097/00000542-199812000-00007)
53. Michard F, Chemla D, Richard C, et al. Clinical Use of Respiratory Changes in Arterial Pulse Pressure to Monitor the Hemodynamic Effects of PEEP. *American Journal of Respiratory and Critical Care Medicine*. 1999;159(3):935-939. doi:[10.1164/ajrccm.159.3.9805077](https://doi.org/10.1164/ajrccm.159.3.9805077)
54. Michard F, Boussat S, Chemla D, et al. Relation between Respiratory Changes in Arterial Pulse Pressure and Fluid Responsiveness in Septic Patients with Acute Circulatory Failure. *American Journal of Respiratory and Critical Care Medicine*. 2000;162(1):134-138. doi:[10.1164/ajrccm.162.1.9903035](https://doi.org/10.1164/ajrccm.162.1.9903035)
55. 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. *Critical Care Medicine*. 2009;37(9):2642-2647. doi:[10.1097/CCM.0b013e3181a590da](https://doi.org/10.1097/CCM.0b013e3181a590da)
56. Teboul JL, Monnet X, Chemla D, Michard F. Arterial Pulse Pressure Variation with Mechanical Ventilation. *American Journal of Respiratory and Critical Care Medicine*. 2018;199(1):rccm.201801-0088CI. doi:[10.1164/rccm.201801-0088CI](https://doi.org/10.1164/rccm.201801-0088CI)
57. De Backer D, Ph D, Taccone FS, et al. Influence of Respiratory Rate on Stroke Volume Variation in Mechanically Ventilated Patients. *Anesthesiology*. 2009;110:1092-1097. doi:[10.1097/ALN.0b013e31819db2a1](https://doi.org/10.1097/ALN.0b013e31819db2a1)
58. Monnet X, Bleibtreu A, Ferré A, et al. Passive leg-raising and end-expiratory occlusion tests perform better than pulse pressure variation in patients with low respiratory system compliance\*. *Critical Care Medicine*. 2012;40(1):152-157. doi:[10.1097/CCM.0b013e31822f08d7](https://doi.org/10.1097/CCM.0b013e31822f08d7)
59. Michard F, Chemla D, Teboul JL. Applicability of pulse pressure variation: How many shades of grey? *Critical Care*. 2015;19(1):15-17. doi:[10.1186/s13054-015-0869-x](https://doi.org/10.1186/s13054-015-0869-x)
60. Wyffels PAH, De Hert S, Wouters PF. New algorithm to quantify cardiopulmonary interaction in patients with atrial fibrillation: A proof-of-concept study. *British Journal of Anaesthesia*. 2021;126(1):111-119. doi:[10.1016/j.bja.2020.09.039](https://doi.org/10.1016/j.bja.2020.09.039)

## References

61. Magder S. Clinical Usefulness of Respiratory Variations in Arterial Pressure. *American Journal of Respiratory and Critical Care Medicine.* 2004;169(2):151-155. doi:[10.1164/rccm.200211-1360CC](https://doi.org/10.1164/rccm.200211-1360CC)
62. Monnet X, Shi R, Teboul JL. Prediction of fluid responsiveness. What's new? *Annals of Intensive Care.* 2022;12(1):46. doi:[10.1186/s13613-022-01022-8](https://doi.org/10.1186/s13613-022-01022-8)
63. Monnet X, Osman D, Ridel C, Lamia B, Richard C, Teboul JL. Predicting volume responsiveness by using the end-expiratory occlusion in mechanically ventilated intensive care unit patients. *Critical Care Medicine.* 2009;37(3):951-956. doi:[10.1097/CCM.0b013e3181968fe1](https://doi.org/10.1097/CCM.0b013e3181968fe1)
64. Myatra SN, Prabu NR, Divatia JV, Monnet X, Kulkarni AP, Teboul JL. The Changes in Pulse Pressure Variation or Stroke Volume Variation After a "Tidal Volume Challenge" Reliably Predict Fluid Responsiveness During Low Tidal Volume Ventilation\*. *Critical Care Medicine.* 2017;45(3):415-421. doi:[10.1097/CCM.0000000000002183](https://doi.org/10.1097/CCM.0000000000002183)
65. Biais M, Larghi M, Henriot J, de Courson H, Sesay M, Nouette-Gaulain K. End-Expiratory Occlusion Test Predicts Fluid Responsiveness in Patients With Protective Ventilation in the Operating Room. *Anesthesia & Analgesia.* 2017;125(6):1889-1895. doi:[10.1213/ANE.0000000000002322](https://doi.org/10.1213/ANE.0000000000002322)
66. Alvarado Sánchez JI, Caicedo Ruiz JD, Diaztagle Fernández JJ, Amaya Zuñiga WF, Ospina-Tascón GA, Cruz Martínez LE. Predictors of fluid responsiveness in critically ill patients mechanically ventilated at low tidal volumes: Systematic review and meta-analysis. *Annals of Intensive Care.* 2021;11(1):28. doi:[10.1186/s13613-021-00817-5](https://doi.org/10.1186/s13613-021-00817-5)
67. Xu Y, Guo J, Wu Q, Chen J. Efficacy of using tidal volume challenge to improve the reliability of pulse pressure variation reduced in low tidal volume ventilated critically ill patients with decreased respiratory system compliance. *BMC anesthesiology.* 2022;22(1):137. doi:[10.1186/s12871-022-01676-8](https://doi.org/10.1186/s12871-022-01676-8)
68. Smorenberg A, Cherpanath TGV, Geerts BF, et al. A mini-fluid challenge of 150 mL predicts fluid responsiveness using Modelflow R pulse contour cardiac output directly after cardiac surgery. *Journal of Clinical Anesthesia.* 2018;46:17-22. doi:[10.1016/j.jclinane.2017.12.022](https://doi.org/10.1016/j.jclinane.2017.12.022)
69. Biais M, de Courson H, Lanchon R, et al. Mini-fluid Challenge of 100 ml of Crystallloid Predicts Fluid Responsiveness in the Operating Room: *Anesthesiology.* 2017;127(3):450-456. doi:[10.1097/ALN.0000000000001753](https://doi.org/10.1097/ALN.0000000000001753)
70. Wu Y, Zhou S, Zhou Z, Liu B. A 10-second fluid challenge guided by transthoracic echocardiography can predict fluid responsiveness. *Critical Care.* 2014;18(3):R108. doi:[10.1186/cc13891](https://doi.org/10.1186/cc13891)
71. 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 Anestesiologica.* 2019;85(9):981-988. doi:[10.23736/S0375-9393.19.13276-2](https://doi.org/10.23736/S0375-9393.19.13276-2)

## References

72. Mukhtar A, Awad M, Elayashy M, et al. Validity of mini-fluid challenge for predicting fluid responsiveness following liver transplantation. *BMC Anesthesiology*. 2019;19(1):56. doi:[10.1186/s12871-019-0728-4](https://doi.org/10.1186/s12871-019-0728-4)
73. Lee CT, Lee TS, Chiu CT, Teng HC, Cheng HL, Wu CY. 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*. 2020;99(6):e19031. doi:[10.1097/MD.00000000000019031](https://doi.org/10.1097/MD.00000000000019031)
74. Fot EV, Izotova NN, Smetkin AA, Kuzkov VV, Kirov MY. Dynamic Tests to Predict Fluid Responsiveness After Off-Pump Coronary Artery Bypass Grafting. *Journal of Cardiothoracic and Vascular Anesthesia*. 2020;34(4):926-931. doi:[10.1053/j.jvca.2019.09.013](https://doi.org/10.1053/j.jvca.2019.09.013)
75. 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. *European Journal of Anaesthesiology / EJA*. 2021;38(4):422-431. doi:[10.1097/EJA.0000000000001406](https://doi.org/10.1097/EJA.0000000000001406)
76. Archie J. Mathematic Coupling of Data: A Common Source of Error. *Annals of Surgery*. 1981;193(3):296-303.
77. Stratton HH, Feustel PJ, Newell JC. Regression of calculated variables in the presence of shared measurement error. *Journal of Applied Physiology*. 1987;62(5):2083-2093. doi:[10.1152/jappl.1987.62.5.2083](https://doi.org/10.1152/jappl.1987.62.5.2083)
78. Hastie T, Tibshirani R. Generalized Additive Models. *Statistical Science*. 1986;1(3):297-318.
79. Wood SN. *Generalized Additive Models: An Introduction with R*. Second. Chapman and Hall/CRC; 2017. doi:[10.1201/9781315370279](https://doi.org/10.1201/9781315370279)
80. Wood SN. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society Series B (Statistical Methodology)*. 2011;73(1):3-36.
81. Network TARDS. Ventilation with Lower Tidal Volumes as Compared with Traditional Tidal Volumes for Acute Lung Injury and the Acute Respiratory Distress Syndrome. *The New England Journal of Medicine*. Published online 2000:8.
82. Kovacs G, Herve P, Barbera JA, et al. An official European Respiratory Society statement: Pulmonary haemodynamics during exercise. *European Respiratory Journal*. 2017;50(5). doi:[10.1183/13993003.00578-2017](https://doi.org/10.1183/13993003.00578-2017)
83. McKenzie SC, Dunster K, Chan W, et al. Reliability of thermodilution derived cardiac output with different operator characteristics. *Journal of Clinical Monitoring and Computing*. 2018;32(2):227-234. doi:[10.1007/s10877-017-0010-6](https://doi.org/10.1007/s10877-017-0010-6)

## References

84. Joosten A, Desebbe O, Suehiro K, et al. Accuracy and precision of non-invasive cardiac output monitoring devices in perioperative medicine: A systematic review and meta-analysis. *British Journal of Anaesthesia.* 2017;118(3):298-310. doi:[10.1093/bja/aew461](https://doi.org/10.1093/bja/aew461)
85. Montenij LJ, Buhre WF, Jansen JR, Kruitwagen CL, de Waal EE. Methodology of method comparison studies evaluating the validity of cardiac output monitors: A stepwise approach and checklist† †This Article is accompanied by Editorial Aew110. *British Journal of Anaesthesia.* 2016;116(6):750-758. doi:[10.1093/bja/aew094](https://doi.org/10.1093/bja/aew094)
86. Slagt C, Malagon I, Groeneveld ABJ. Systematic review of uncalibrated arterial pressure waveform analysis to determine cardiac output and stroke volume variation. *British Journal of Anaesthesia.* 2014;112(4):626-637. doi:[10.1093/bja/aet429](https://doi.org/10.1093/bja/aet429)
87. Geisen M, Ganter MT, Hartnack S, Dzemali O, Hofer CK, Zollinger A. Accuracy, Precision, and Trending of 4 Pulse Wave Analysis Techniques in the Postoperative Period. *Journal of Cardiothoracic and Vascular Anesthesia.* 2018;32(2):715-722. doi:[10.1053/j.jvca.2017.09.006](https://doi.org/10.1053/j.jvca.2017.09.006)
88. Heijne A, Krijtenburg P, Bremers A, Scheffer GJ, Malagon I, Slagt C. Four different methods of measuring cardiac index during cytoreductive surgery and hyperthermic intraperitoneal chemotherapy. *Korean Journal of Anesthesiology.* 2020;74(2):120-133. doi:[10.4097/kja.20202](https://doi.org/10.4097/kja.20202)
89. Murata Y, Imai T, Takeda C, Mizota T, Kawamoto S. Agreement between continuous cardiac output measured by the fourth-generation FloTrac/Vigileo system and a pulmonary artery catheter in adult liver transplantation. *Scientific Reports.* 2022;12(1):11198. doi:[10.1038/s41598-022-14988-z](https://doi.org/10.1038/s41598-022-14988-z)
90. Eisenried A, Klarwein R, Ihmsen H, et al. Accuracy and Trending Ability of the Fourth-Generation FloTrac/EV1000 System in Patients With Severe Aortic Valve Stenosis Before and After Surgical Valve Replacement. *Journal of Cardiothoracic and Vascular Anesthesia.* 2019;33(5):1230-1236. doi:[10.1053/j.jvca.2018.09.015](https://doi.org/10.1053/j.jvca.2018.09.015)
91. Suehiro K, Tanaka K, Mikawa M, et al. Improved Performance of the Fourth-Generation FloTrac/Vigileo System for Tracking Cardiac Output Changes. *Journal of Cardiothoracic and Vascular Anesthesia.* 2015;29(3):656-662. doi:[10.1053/j.jvca.2014.07.022](https://doi.org/10.1053/j.jvca.2014.07.022)
92. Maeda T, Hamaguchi E, Kubo N, Shimokawa A, Kanazawa H, Ohnishi Y. The accuracy and trending ability of cardiac index measured by the fourth-generation FloTrac/Vigileo system™ and the Fick method in cardiac surgery patients. *Journal of Clinical Monitoring and Computing.* 2019;33(5):767-776. doi:[10.1007/s10877-018-0217-1](https://doi.org/10.1007/s10877-018-0217-1)
93. Narang N, Thibodeau JT, Grodin JL, Garg S, McGuire DK, Drazner MH. Thermodilution Cardiac Index Has Poor Agreement with That Measured by the Direct Fick Method in Low-Output States. *The Journal of Heart and Lung Transplantation.* 2021;40(4, Supplement):S269. doi:[10.1016/j.healun.2021.01.766](https://doi.org/10.1016/j.healun.2021.01.766)

## References

94. Narang N, Thibodeau JT, Parker WF, et al. Comparison of Accuracy of Estimation of Cardiac Output by Thermodilution Versus the Fick Method Using Measured Oxygen Uptake. *The American Journal of Cardiology*. 2022;176:58-65. doi:[10.1016/j.amjcard.2022.04.026](https://doi.org/10.1016/j.amjcard.2022.04.026)
95. McElreath R. *Statistical Rethinking: A Bayesian Course with Examples in R and STAN*. CRC Press; 2020.
96. Gelman A, Greenland S. Are confidence intervals better termed “uncertainty intervals”? *BMJ*. 2019;366:l5381. doi:[10.1136/bmj.l5381](https://doi.org/10.1136/bmj.l5381)
97. Guinot PG, Bernard E, Defrancq F, et al. Mini-fluid challenge predicts fluid responsiveness during spontaneous breathing under spinal anaesthesia: An observational study. *European Journal of Anaesthesiology*. 2015;32(9):645-649. doi:[10.1097/EJA.0000000000000175](https://doi.org/10.1097/EJA.0000000000000175)
98. Raut MS, Maheshwari A. “X” descent of CVP: An indirect measure of RV dysfunction ? *Journal of Anaesthesiology, Clinical Pharmacology*. 2014;30(3):430-431. doi:[10.4103/0970-9185.137289](https://doi.org/10.4103/0970-9185.137289)
99. Abdullah T. Short time low PEEP challenge and mini fluid challenge to evaluate fluid responsiveness in the operating room. *Journal Of Cardio-Vascular-Thoracic Anaesthesia And Intensive Care Society*. Published online 2022. doi:[10.14744/GKDAD.2022.04900](https://doi.org/10.14744/GKDAD.2022.04900)
100. Arheden H, Holmqvist C, Thilen U, et al. Left-to-Right Cardiac Shunts: Comparison of Measurements Obtained with MR Velocity Mapping and with Radionuclide Angiography. *Radiology*. 1999;211(2):453-458. doi:[10.1148/radiology.211.2.r99ma43453](https://doi.org/10.1148/radiology.211.2.r99ma43453)
101. Altman DG, Royston P. **The cost of dichotomising continuous variables**. *BMJ : British Medical Journal*. 2006;332(7549):1080.
102. Tu YK, Gilthorpe MS. Revisiting the relation between change and initial value: A review and evaluation. *Statistics in Medicine*. 2007;26(2):443-457. doi:[10.1002/sim.2538](https://doi.org/10.1002/sim.2538)
103. 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 Anaesthesiologica Scandinavica*. 2022;(66):7-24. doi:[10.1111/aas.13965](https://doi.org/10.1111/aas.13965)
104. Törnqvist L, Vartia P, Vartia YO. How Should Relative Changes be Measured? *The American Statistician*. 1985;39(1):43-46. doi:[10.1080/00031305.1985.10479385](https://doi.org/10.1080/00031305.1985.10479385)
105. Høiseth LØ, Hagemo JS. Predicting fluid responsiveness in whom? A simulated example of patient spectrum influencing the receiver operating characteristics curve. *Journal of Clinical Monitoring and Computing*. 2018;32(2):215-219. doi:[10.1007/s10877-017-0019-x](https://doi.org/10.1007/s10877-017-0019-x)

## *References*

106. Feissel M, Teboul JL, Merlani P, Badie J, Faller JP, Bendjelid K. Plethysmographic dynamic indices predict fluid responsiveness in septic ventilated patients. *Intensive Care Medicine*. 2007;33(6):993-999. doi:[10.1007/s00134-007-0602-6](https://doi.org/10.1007/s00134-007-0602-6)
107. Cannesson M, Attof Y, Rosamel P, et al. Respiratory Variations in Pulse Oximetry Plethysmographic Waveform Amplitude to Predict Fluid Responsiveness in the Operating Room. *Anesthesiology*. 2007;106(6):1105-1111. doi:[10.1097/01.anes.0000267593.72744.20](https://doi.org/10.1097/01.anes.0000267593.72744.20)

# **Appendices**

A

Paper 1



Received: 26 April 2021 | Revised: 2 July 2021 | Accepted: 19 July 2021

DOI: 10.1111/aas.13965

REVIEW

Acta Anaesthesiologica Scandinavica

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

Johannes Enevoldsen<sup>1,2</sup> | Thomas W. L. Scheeren<sup>3</sup> | Jonas M. Berg<sup>1,2</sup> |  
Simon T. Vistisen<sup>1,2</sup>

<sup>1</sup>Department of Clinical Medicine, Aarhus University, Aarhus, Denmark

<sup>2</sup>Department of Anaesthesiology and Intensive Care, Aarhus University Hospital, Aarhus, Denmark

<sup>3</sup>Department 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 ago<sup>1</sup> 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.<sup>2</sup> That 'some fluid' was a fluid challenge of around 500 ml as identified by the FENICE study.<sup>3</sup> 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).<sup>1</sup> Others have since investigated and validated the MFC, and a recent systematic review including seven MFC studies (368 fluid challenges in 324 patients)<sup>1,4–9</sup>

## A. Paper 1

identified a pooled AUROC of 0.91 for the MFC method.<sup>10</sup> 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.<sup>11-15</sup>

In 2018, we published a correspondence debating the way MFC studies were designed.<sup>16</sup> 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 ( $\Delta SV_{100}$ ) and the response to the full fluid challenge ( $\Delta SV_{500}$ ) both include the haemodynamic variable measured at baseline, that is before the MFC. Specifically,  $\Delta SV_{100}$  and  $\Delta SV_{500}$  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.<sup>1</sup> 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 ( $\Delta SV_{100}$ ) is also a part of the outcome we try to predict ( $\Delta SV_{500}$ ).

### 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).<sup>17,18</sup> 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

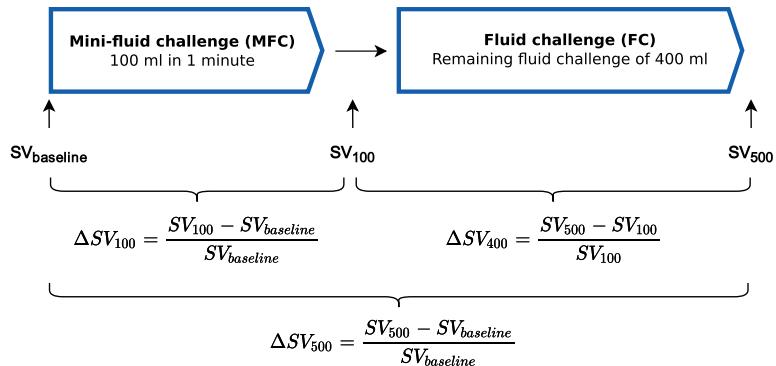
$$\Delta SV_{500} = \frac{SV_{500} + \epsilon_{SV,500}}{SV_{\text{baseline}} + \epsilon_{SV,\text{baseline}}} - 1.$$

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 ( $\Delta SV_{100}$ ) and the outcome ( $\Delta SV_{500}$ ), making them spuriously correlated, and therefore more likely to agree.<sup>19</sup>

### 2.2 | The predictor ( $\Delta SV_{100}$ ) is also a part of the outcome we try to predict ( $\Delta SV_{500}$ )

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 ( $\Delta SV_{100}$ ) to predict  $\Delta SV_{400}$  (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 ( $\Delta SV_{100}$ ) and outcome ( $\Delta SV_{400}$ ) 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.<sup>20,21</sup>

### 2.3 | Secondary analysis of an existing study

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

$\Delta SV_{400}$  is defined as

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

If  $\Delta SV_{100}$  and  $\Delta SV_{500}$  are known, we can calculate  $\Delta SV_{400}$ :

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

Therefore,

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

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

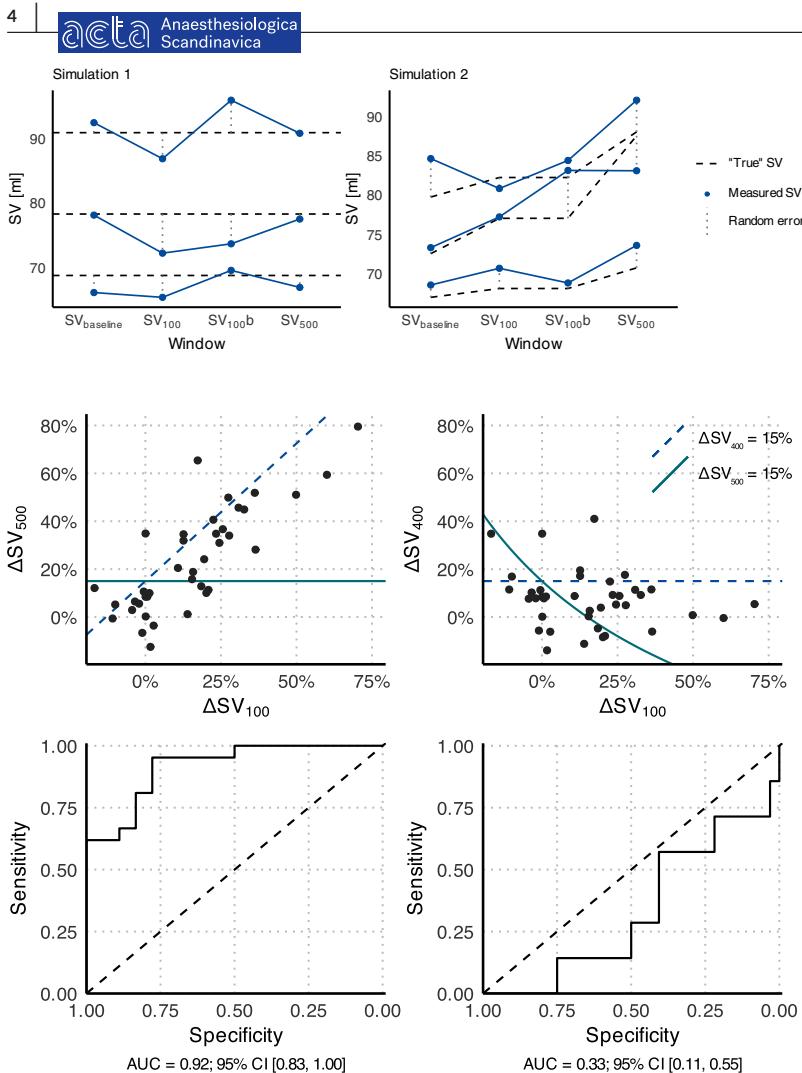
### 2.4 | Simulations

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*,<sup>22–24</sup> 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).<sup>25</sup> From these three simulated measurements of a 'constant' SV (but with random measurement error added), we calculated  $\Delta SV_{100}$ ,  $\Delta SV_{400}$  and  $\Delta SV_{500}$ . 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  $\Delta SV_{100}$  predicted an increase in SV with either  $\Delta SV_{500}$ ,  $\Delta SV_{400}$  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).<sup>1</sup> Upper panels: Scatter plots of the relation between  $\Delta SV_{500}$  and  $\Delta SV_{100}$  (left) and the relation between  $\Delta SV_{400}$  (derived) and  $\Delta SV_{100}$  (right). The full line represents the level at which  $\Delta SV_{500}$  is 15% and the dashed line represents the level at which  $\Delta SV_{400}$  is 15%. Lower panels: Corresponding ROC classification curves of  $\Delta SV_{100}$  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  $SV_{\text{baseline}}$  to  $SV_{500}$  (the 'true'  $\Delta SV_{500}$ ). The simulated fluid response was drawn from a normal distribution (mean change = 15%, SD = 10%). To keep the simulation simple, the 'true'  $\Delta SV_{100}$  was defined as 30% of this 'true'  $\Delta SV_{500}$ :

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

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

'True'  $SV_{100}$  = '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 ( $\Delta SV_{400b}$ ). 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  $\Delta SV_{100}$ 's ability to predict  $\Delta SV_{500} > 15\%$  (left panels) and  $\Delta SV_{100}$ '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 ( $\Delta SV_{400}$ ).

### 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  $\Delta SV_{100}$ ) 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  $\Delta SV_{100}$  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,  $\Delta SV_{100}$  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 ( $SV_{100b}$ ),  $\Delta SV_{100}$  predicted  $\Delta SV_{400b} > 15\%$  with an AUROC of 0.65.

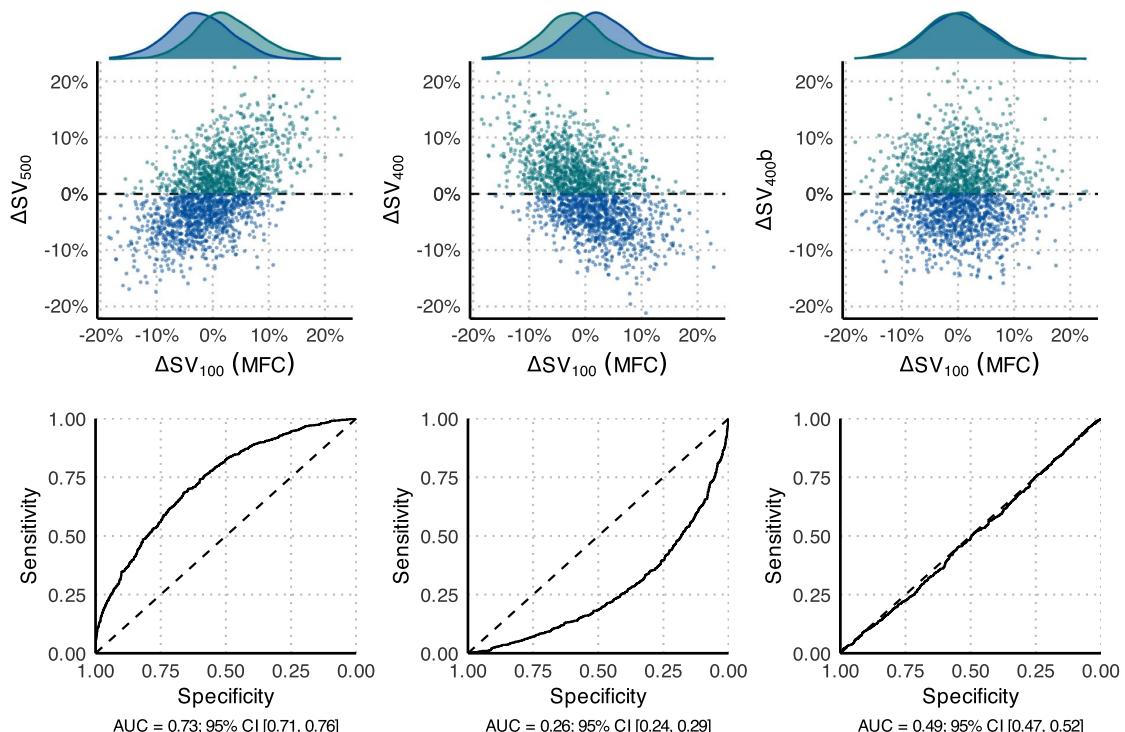
## 4 | DISCUSSION

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

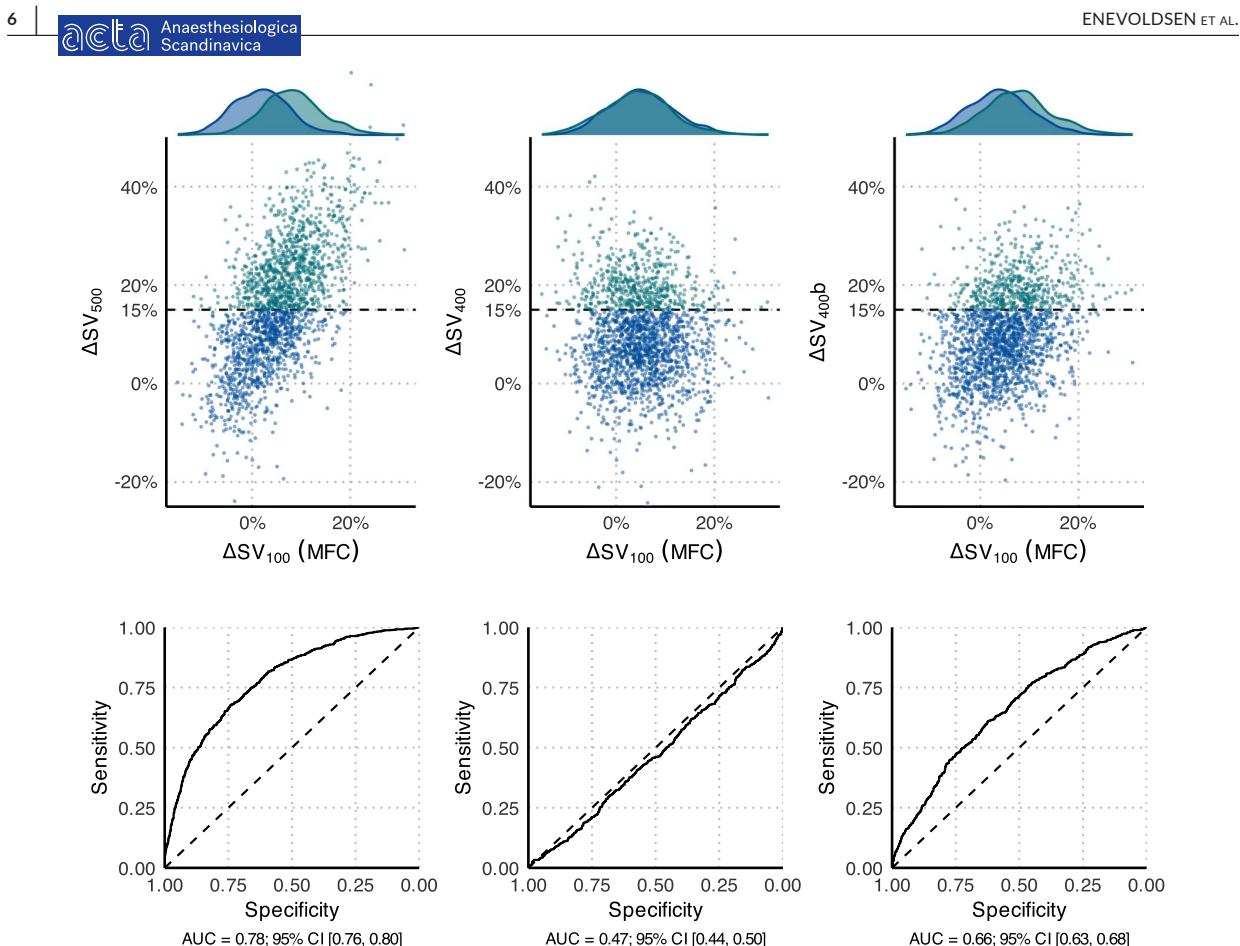
with such problematic designs may overestimate the true classification accuracy of an MFC. This should be considered before adopting the MFC into clinical practice. Still, there are aspects of the above simulations that are worth discussing, and other study designs that should be considered in the search 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’  $\Delta SV_{100}$  is 30% of ‘true’  $\Delta SV_{500}$ ). 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 ( $\Delta SV_{500}$ ), 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  $\Delta SV_{100}$  predicting fluid responsiveness ( $\Delta SV_{500}$ ,  $\Delta SV_{400}$  and  $\Delta SV_{400b} > 0\%$ ). The changes in stroke volume ( $\Delta SV$ ) 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  $\Delta SV_{100}$  predicting fluid responsiveness ( $\Delta SV_{500}$ ,  $\Delta SV_{400}$  and  $\Delta SV_{400}b > 15\%$ ). 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  $\Delta SV_{400}$  is a more clinically meaningful outcome to predict, we discourage using  $\Delta SV_{400}$  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

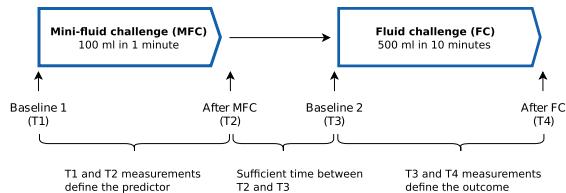
In one study, authors used different monitoring modalities for predictor and outcome variables: changes in pulse pressure variation

( $\Delta PPV$ ) predicting fluid responsiveness (defined as change in cardiac output).<sup>8</sup> 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.<sup>5</sup> 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.<sup>5</sup> 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,<sup>5</sup> 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.<sup>26,27</sup>

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.<sup>10</sup> 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.<sup>28</sup> 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.<sup>29</sup> 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.<sup>30,31</sup> Because of the different infusion rates and durations, the MFC is not simply a 'mini' version of the full fluid challenge. 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.  $\Delta SV_{500}$ ) 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.<sup>32,33</sup> 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.<sup>34</sup> 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.

## A. Paper 1

8

### AUTHOR CONTRIBUTIONS

JE: Conception, manuscript preparation, simulation, artwork and revision. TWLS: 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 CO<sub>2</sub> 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, Elayash 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 fLIuid 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.
- Squara P, Scheeren TWL, Aya HD, et al. Metrology part 1: definition of quality criteria. *J Clin Monit Comput*. 2021;35:17-25.
- Squara 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. *Sociol 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



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

Johannes Enevoldsen<sup>1,2</sup> · Gavin L. Simpson<sup>3</sup> · Simon T. Vistisen<sup>1,2</sup>

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

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

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

<sup>3</sup> Department of Animal Science, Aarhus University, Tjele, Denmark

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.

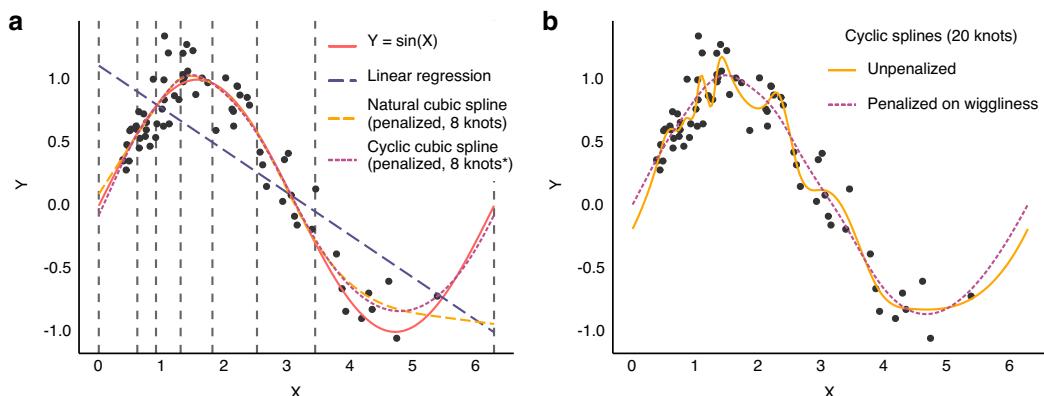
Hastie and Tibshirani introduced GAMs in 1986, as extensions of generalized linear models [5]. Instead of fitting straight lines, GAMs can fit any smooth function. In the basic form of a GAM, a smooth function is fitted for each independent variable in the model. These functions are added together to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

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

where  $\alpha$  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 wiggliness (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 wiggliness 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].

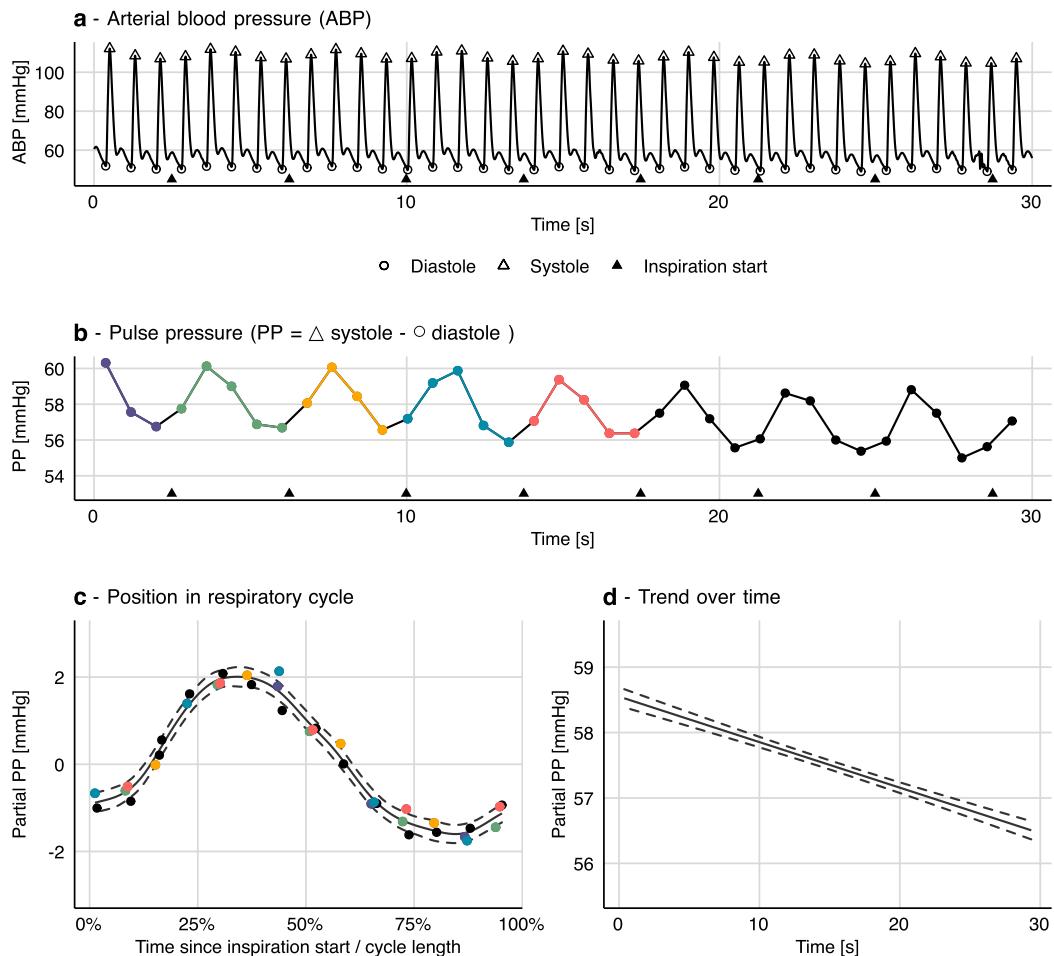
### 2.2 Example 1: Pulse pressure

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

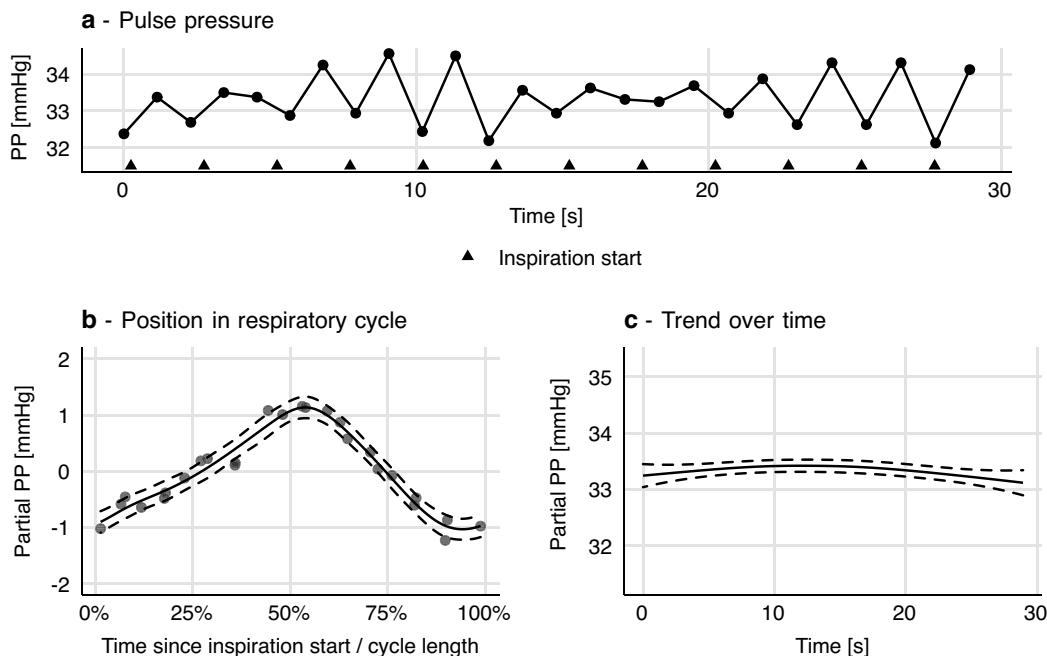
**b** and **c**. **d** The second smooth represents the trend in PP over time with the model constant ( $\alpha$ ) added. The sum of these two smooths (**b** and **c**) gives the model prediction. Residuals of the model ( $\epsilon$ ) are the vertical distance from the smooth to the points in panel **c** (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  $\alpha$  is the mean PP over the entire sample.  $\epsilon$  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,<sup>1</sup> 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.

<sup>1</sup> 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  $\alpha$  is the mean PP, this is equivalent to the classic formula for PPV:

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

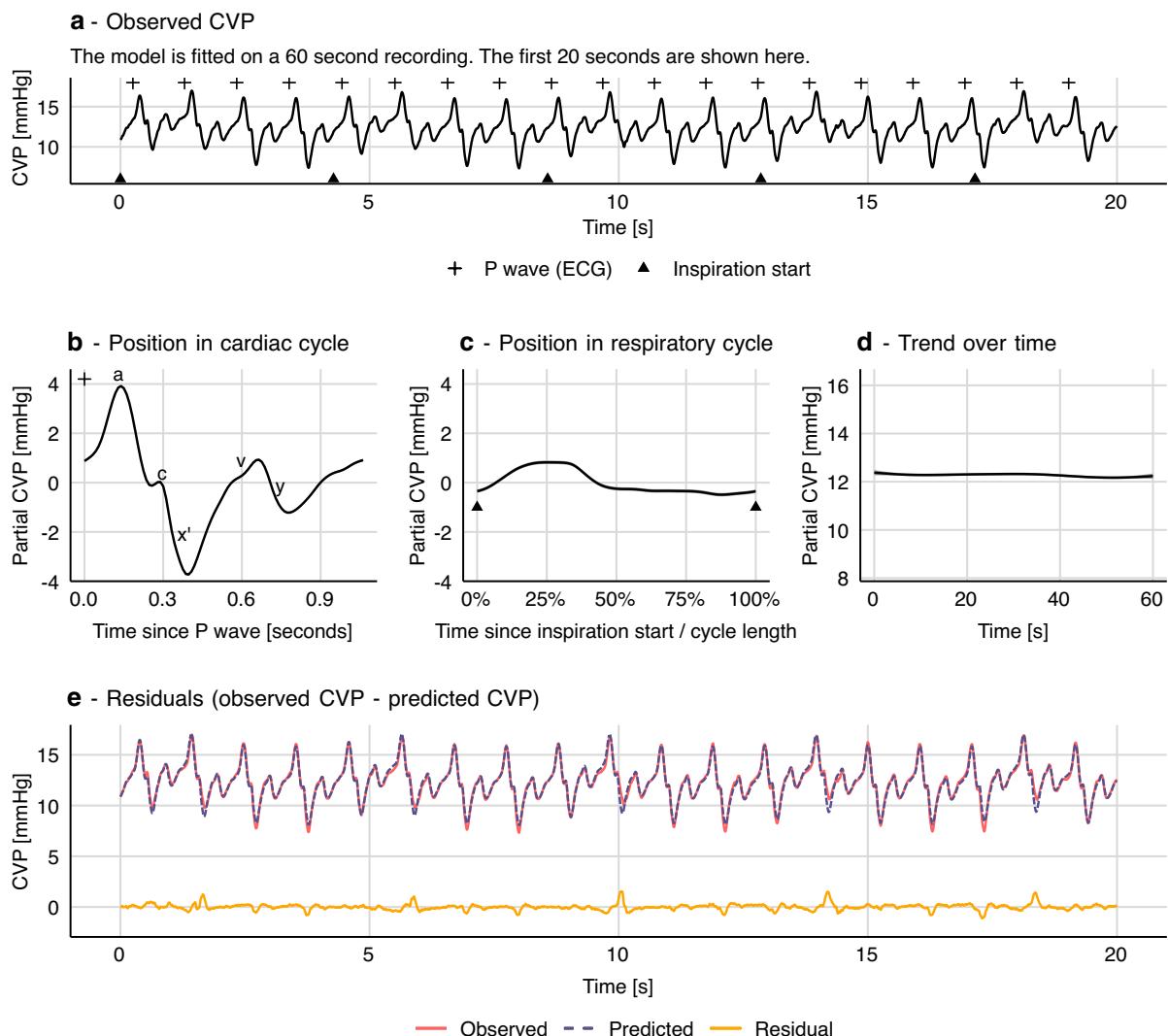
Calculation of a confidence interval for PPV is described in Online Resource 1.

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].

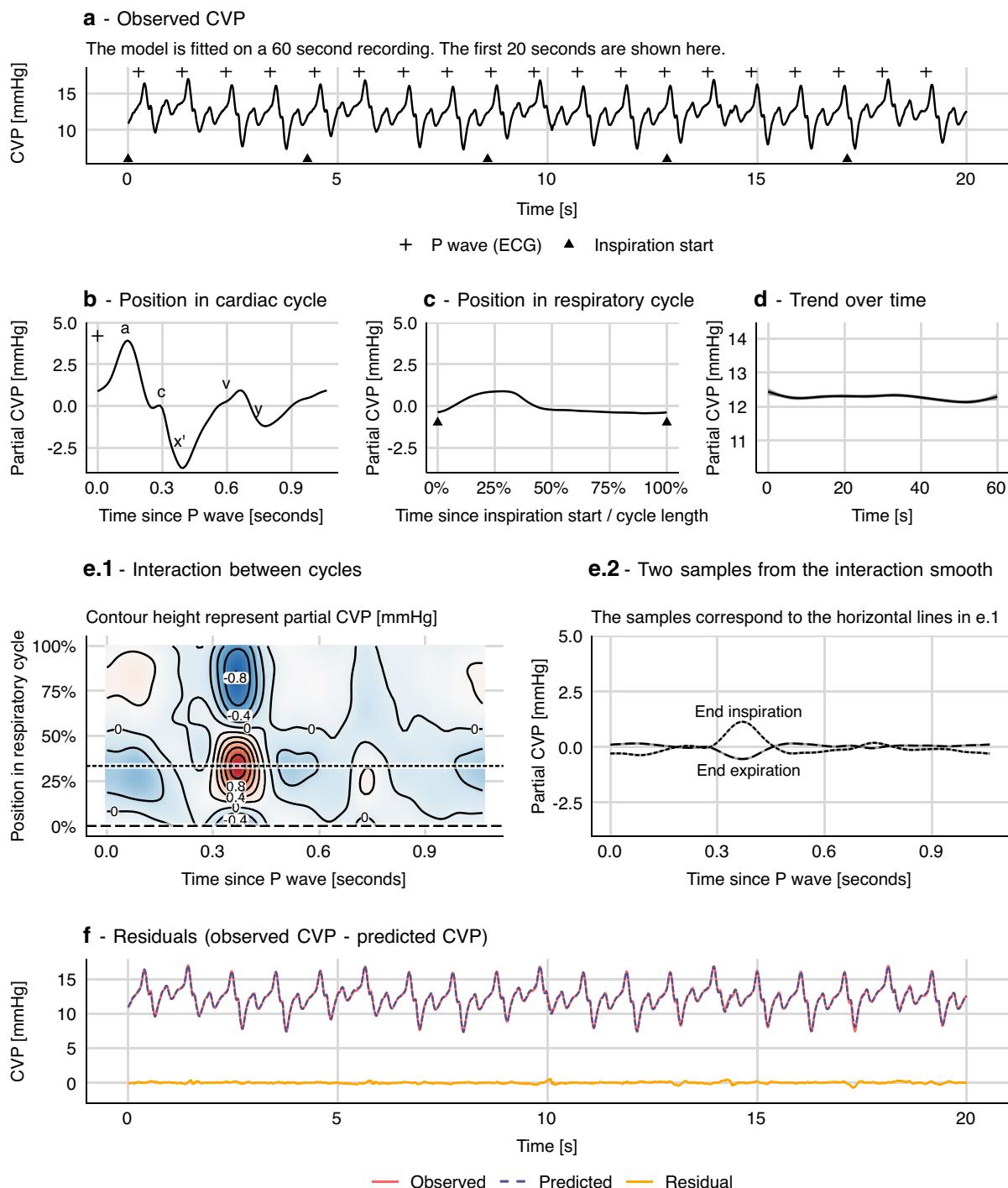
### 2.3.2 How the CVP waveform changes after a fluid bolus

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:

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

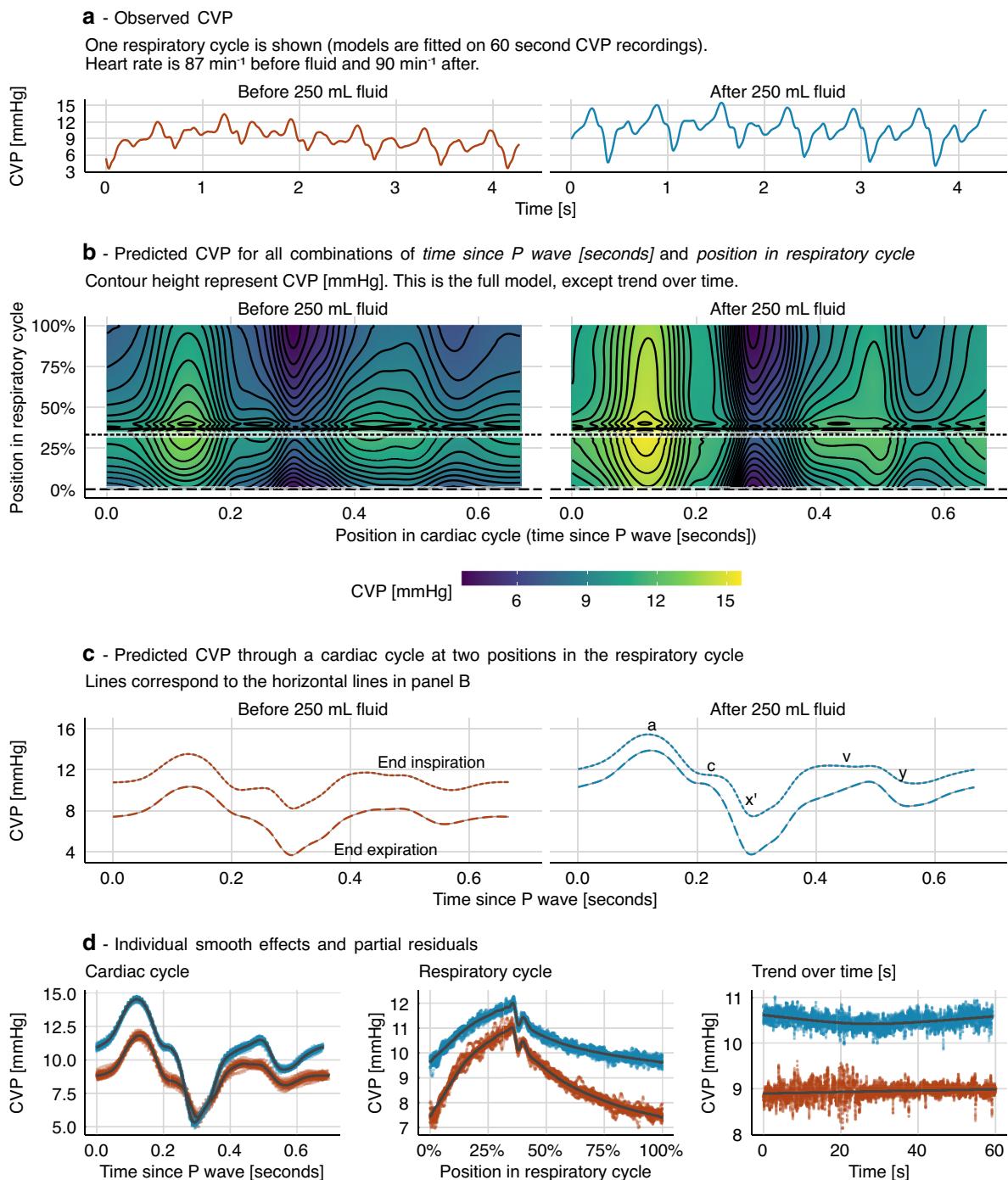
where  $\beta_s$  is an additional constant, that is zero for the pre-fluid section, and  $f_s$  is a spline for each section of data (before or after 250 ml fluid). This model also extends the previous model (Fig. 5) by using an *adaptive* smoothing spline to estimate  $f_s(pos_{ventilation})$ . An 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 expiration<sup>2</sup> 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.

<sup>2</sup> 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.



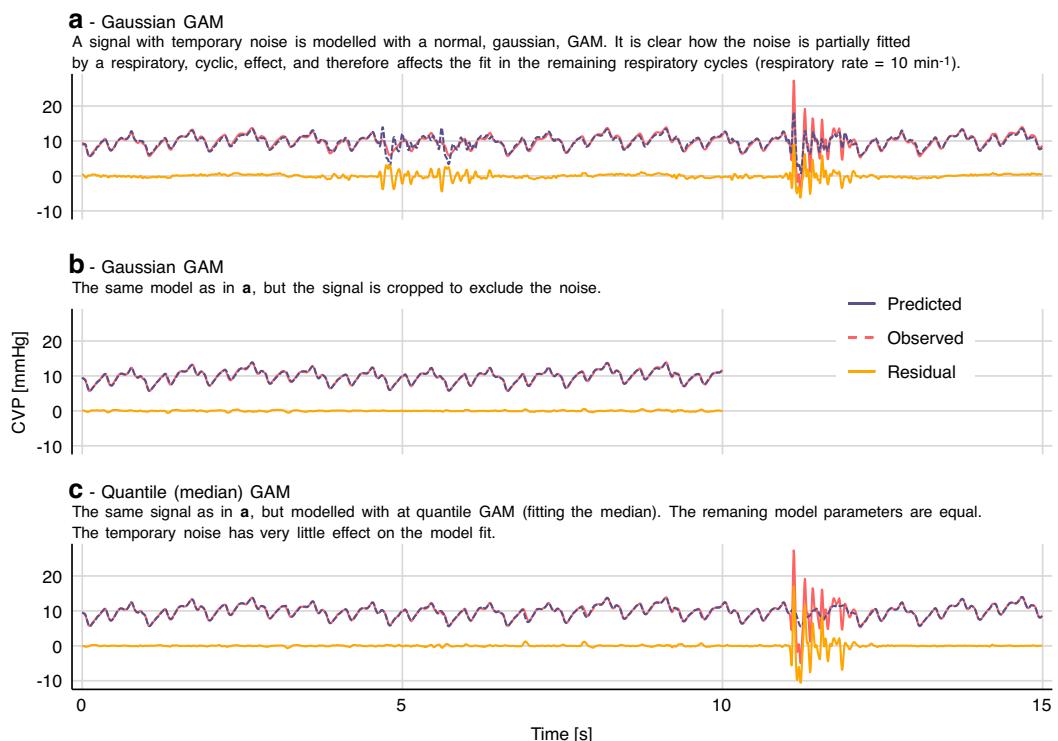
**Fig. 5** How a generalized additive model (GAM) can be fitted to a CVP waveform. **a** Each sample from a 125 Hz CVP waveform is represented with three predictor variables: position in cardiac cycle, position in respiratory cycle and time (seconds since sample start). A GAM is fitted giving the smooth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

stant ( $\alpha$ ) 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 ( $\alpha$  and  $\beta_s$ ) are included in all predictions (b, c and d)



**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 Karippacherril for an effective collaborative effort on extending his open source software, VitalSignCapture, to support exporting real time data from Dräger 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

- 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>.
- 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>.
- 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>.
- 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>.
- Hastie T, Tibshirani R. Generalized additive models. Stat Sci Inst Math Stat. 1986;1:297–318.
- Hastie T, Tibshirani R, Friedman JH. The elements of statistical learning: data mining, inference, and prediction. New York: Springer; 2009.
- 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.
- R Core Team. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2021.
- Simpson GL. *gratia*: graceful ggplot-based graphics and other functions for GAMs fitted using mgcv; 2022. <https://gavinsimpson.github.io/gratia/>.
- 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>.
- 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>.
- Karippacheril JG, Ho TY. Data acquisition from S/5 GE datex anesthesia monitor using VSCapture: an open source.NET/

## B. Paper 2

- 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/10.1097/00000542-200508000-00026>.
  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/10.1097/00000441-190707000-00002>.
  19. Constant J. The X prime descent in jugular contour nomenclature and recognition. *Am Heart J.* 1974;88:372–9. [https://doi.org/10.1016/0000-2870\(74\)90474-8](https://doi.org/10.1016/0000-2870(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/10.1177/088506660001500505>.
  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. Meij LPB, van Houte J, Conjaerts BCM, Bindels AJGH, Bouwmans 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/10.1201/97815370279>.
  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.

C

Paper 3

*C. Paper 3*

Title:

**The Effects of Respiratory Rate and Tidal Volume on Pulse Pressure Variation in Healthy Lungs—A Generalized Additive Model Approach May Help Overcome Limitations**

Johannes Enevoldsen<sup>1,2</sup> (ORCID: 0000-0002-9190-6566) \*

Birgitte Brandsborg<sup>2</sup> (ORCID: 0000-0003-2935-3152)

Peter Juhl-Olsen<sup>1,3</sup> (ORCID 0000-0003-1951-9091)

Stephen Edward Rees<sup>4</sup> (ORCID: 0000-0003-4420-0605)

Henriette Vind Thaysen<sup>5</sup> (ORCID: 0000-0001-6225-1559)

Thomas W.L. Scheeren<sup>6,7</sup> (ORCID: 0000-0002-9184-4190)

Simon Tilma Vistisen<sup>1,2</sup> (ORCID: 0000-0002-1297-1459)

<sup>1</sup> Department of Clinical Medicine, Aarhus University, Aarhus, Denmark

<sup>2</sup> Department of Anaesthesiology, Aarhus University Hospital, Aarhus, Denmark

<sup>3</sup> Department of Cardiothoracic- and Vascular Surgery, Anaesthesia Section, Aarhus University Hospital, Aarhus, Denmark

<sup>4</sup> Department of Health Science and Technology, Aalborg University, Denmark.

<sup>5</sup> Department of Surgery, Aarhus University Hospital, Aarhus, Denmark

<sup>6</sup> Department of Anesthesiology, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands

<sup>7</sup> Edwards Lifesciences

\*Corresponding author:

Johannes Enevoldsen, M.D.

Aarhus University Hospital, Department of Anaesthesiology,

Department of Clinical Medicine, Aarhus University

Palle Juul-Jensens Boulevard 99

8200 Aarhus N, Denmark

*C. Paper 3*

Email: enevoldsen@clin.au.dk

Short running title: Effects of Ventilator Setting on PPV

Keywords: cardiopulmonary interaction; dynamic filling variable; fluid responsiveness; haemodynamic; hemodynamic monitoring; heart-lung interaction; mechanical ventilation; pulse pressure variation; statistical modelling

Abstract word count: 246/250

Body word count: 2997/3000

## Abstract

Background:

Pulse pressure variation (PPV) is a well-established method for predicting fluid responsiveness in mechanically ventilated patients. The predictive accuracy is, however, disputed for ventilation with low tidal volume (VT) or low heart-rate-to-respiratory-rate ratio (HR/RR). We investigated the effects of VT and RR on PPV and on PPV's ability to predict fluid responsiveness.

Methods:

We included patients scheduled for open abdominal surgery. Prior to a 250 ml fluid bolus, we ventilated patients with combinations of VT from 4 to 10 ml kg<sup>-1</sup> and RR from 10 to 31 min<sup>-1</sup>. For each of 10 RR-VT combinations, PPV was derived using both a classic approach and a generalized additive model (GAM) approach. The stroke volume (SV) response to fluid was evaluated using uncalibrated pulse contour analysis. An SV increase > 10% defined fluid responsiveness.

Results:

Fifty of 52 included patients received a fluid bolus. Ten were fluid responders. For all ventilator settings, fluid responsiveness prediction with PPV was inconclusive with point estimates for the area under the receiver operating characteristics curve between 0.62 and 0.82. Both PPV measures were nearly proportional to VT. Higher RR was associated with lower PPV. Classically derived PPV was affected more by RR than GAM-derived PPV.

Conclusions:

Correcting PPV for VT might improve the predictive utility of PPV. Low HR/RR has limited effect on GAM-derived PPV, indicating that the low HR/RR limitation is related to how PPV is calculated. We did not demonstrate any benefit from GAM-derived PPV in predicting fluid responsiveness.

## Background

Ventilator-induced pulse pressure variation (PPV) is a well-established and accurate method for predicting fluid responsiveness.<sup>1,2</sup> Despite this accuracy, there are important limitations to its clinical use, including ventilation with low tidal volumes (VT) and low heart-rate-to-respiratory-rate ratio (HR/RR).<sup>3–5</sup> These limitations are frequently discussed, though their physiological basis is incompletely understood.<sup>3</sup>

Low VT ventilation was presented as a limitation by De Backer and colleagues in 2005, where it was shown that PPV only reliably predicted fluid responsiveness in patients ventilated with a  $VT > 8 \text{ ml kg}^{-1}$ .<sup>5</sup> However, low VT was highly associated with a diagnosis of acute respiratory distress syndrome (ARDS), making it difficult to isolate the effect of VT. The study results may also have been affected by RR,<sup>4</sup> respiratory system compliance,<sup>6</sup> and other aspects of the underlying lung disease. There are clinical studies with low VT ventilation where PPV predicts fluid responsiveness well, though the predictive performance varies substantially.<sup>7</sup> It has been shown that PPV is highly correlated between different VT settings in the same patient,<sup>8</sup> and adjusting PPV by tidal volume, respiratory driving pressure or changes in pleural pressure, has been suggested.<sup>9–11</sup> The effect of VT on PPV is, however, still unclear. No studies have simultaneously investigated low VT and low HR/RR limitations in the same patients to decouple these potentially interacting effects.

Low HR/RR was presented as a limitation by De Backer and colleagues in 2009, where it was shown in 17 patients that a  $HR/RR < 3.6$  hindered accurate fluid responsiveness prediction using PPV.<sup>4</sup> The authors suggested that this was caused by a negative interference between the cyclic swings in the right and left ventricular stroke volume, respectively. We speculate that the low HR/RR limitation may, at least partially, result from a sampling problem specific to the classic PPV algorithm. In the HR/RR study,<sup>4</sup> PPV was calculated for individual respiratory cycles.<sup>4,12</sup> When there are few beats per respiratory cycle (low HR/RR), the beats may not lie close to the maximal and minimal possible PP in each cycle, giving an underestimation of PPV.<sup>13</sup> This limitation might be overcome by estimating PPV from a generalized additive model (GAM) of PP.<sup>14,15</sup>

We investigated the following research questions:

1. How does altering VT and RR affect PPV's ability to predict fluid responsiveness?
2. What is the agreement between PPV derived with a GAM and with the classic approach?
3. How does altering VT and RR affect PPV?

*C. Paper 3*

4. Does HR/RR have different effects on GAM-derived PPV compared to PPV calculated with the classic approach?

## Methods

This prospective study was conducted at Aarhus University Hospital, Denmark, after approval by the regional ethics committee (January 2020, case nr.: 1-10-72-245-19) and registration on ClinicalTrials.gov (March 2020, identifier: NCT04298931). Patients gave written informed consent prior to participation.

### Study population

We included adults ( $\geq 18$  years) scheduled for elective open abdominal surgery with hemodynamic monitoring using a FloTrac (4th generation) based device (EV1000 or HemoSphere, Edwards Lifesciences, Irvine, California).

Exclusion criteria were: irregular heart rhythm (e.g. atrial fibrillation), known left ventricular ejection fraction  $\leq 40\%$ , known right ventricular dysfunction (reported qualitatively or Tricuspid Annular Plane Systolic Excursion (TAPSE)  $< 17$  mm), and pregnancy.

The cohort constitutes a convenience sample.

### Protocol

Anaesthesia was induced with propofol and maintained with sevoflurane on a Dräger Perseus A500 (Dräger, Lübeck, Germany) anaesthesia machine; remifentanil was used for analgesia, and rocuronium for muscle relaxation. A thoracic epidural was placed, and tested with 3 ml lidocaine 2% with adrenaline before induction. Arterial- and central venous pressure transducers were zeroed to atmospheric pressure and levelled at the right atrium. Patients were ventilated with pressure regulated volume control (VC-CMV + Autoflow®) with inspiration-expiration-ratio of 1:2.

Patients were observed in the context of a scheduled fluid administration (250 ml albumin or acetated Ringer's solution), where other hemodynamic perturbations were not expected (e.g. due to surgery or medication). Before the fluid administration, a ventilation protocol was initiated comprised by a series of VT- and RR combinations (10 combinations of VT: 4, 6, 8, and 10 ml kg $^{-1}$  (predicted body weight $^{16}$ ) and RR: 10, 17, 24, and 31 min $^{-1}$ ; see Fig. 1. Each setting was used for 30 seconds. For each RR, VT was applied from the lowest to highest. The order of the RR settings 17, 24 and 31 min $^{-1}$  was randomised, while the four settings with RR of 10 min $^{-1}$  were always applied last to minimise effects of potential lung recruitment from ventilation with VT=10 ml kg $^{-1}$ . The maximal allowed airway pressure was 40 cmH $_2$ O.

## C. Paper 3

Afterwards, the ventilator was reset to pre-protocol settings. Two to four minutes after the ventilation protocol, 250 ml of fluids were administered over two minutes. The observation window ended two minutes after the completion of the fluid administration.

### Data recording

We used VitalRecorder<sup>17</sup> to record data from the bedside monitor (Philips IntelliVue MX550, Eindhoven, the Netherlands) and the haemodynamic monitor (Edwards Hemosphere or EV1000), and VSCapture<sup>18</sup> to record data from the ventilator. Recordings were synchronised before analysis.

### Pulse pressure variation

Recordings were divided into ten 30-second windows: one for each protocolised ventilator setting. Individual heart beats were detected from the arterial waveform, and starttime (diastole) and pulse pressure (systolic pressure - diastolic pressure) were recorded. A beat was marked as an extrasystole and excluded if the time since the previous beat was less than 90% of the median of the 10 preceding beat intervals. The following (post-ectopic) beat was also excluded. Additionally, outlier beats defined by a pulse pressure more than  $\pm 25\%$  from the median of the 10 nearest beats were excluded. Thirty-second windows containing  $\geq 3$  extrasystoles were excluded.

Two PPV calculations were performed for each 30-second window: the classic calculation ( $PPV_{Classic}$ ) and PPV estimated using a GAM (see Fig. 2). With  $PPV_{Classic}$ , we aimed to match the calculation of PPV described by De Backer and colleagues:<sup>4,5</sup> Consecutive pairs of maximum and minimum PP were selected so each maximum is within one respiratory length of the previous minimum, and each minimum is within one respiratory length of the previous maximum. For each maximum-minimum pair, PPV was calculated as

$$PPV = 100 \cdot \frac{(PP_{max} - PP_{min})}{(PP_{max} + PP_{min})/2}.$$

$PPV_{Classic}$  was defined as the mean PPV of the last three maximum-minimum pairs during a protocol ventilator setting (see Fig. 2c).

The GAM-derived PPV ( $PPV_{GAM}$ ) was calculated as described previously (see Fig. 2b).<sup>14</sup> This method quantifies the respiratory variation in PP by decomposing the series of PP measurements into a repeating respiratory component, and a slower trend over time. The

## C. Paper 3

respiratory component's peak-to-peak amplitude divided by the mean pulse pressure constitutes  $\text{PPV}_{\text{GAM}}$ .

### Fluid responsiveness

Stroke volume (SV) was estimated with pulse contour analysis (FloTrac). Each patient's SV response to the 250 ml fluid challenge was calculated as:

$$\Delta SV = 100 \cdot \frac{SV_{\text{post}} - SV_{\text{pre}}}{SV_{\text{pre}}},$$

where  $SV_{\text{pre}}$  and  $SV_{\text{post}}$  is the median of two minutes of SV measurements (6 samples), before and after fluid administration. A  $\Delta SV > 10\%$  was considered a positive fluid response (prespecified).

### Statistics

Data was analysed with R 4.1.0, *tidyverse*, *pROC*, *boot* and *brms*.<sup>19-23</sup> Data and code for this section is available at <https://doi.org/10.5281/zenodo.6984311>.

#### Sample size calculation

The study was powered with respect to fluid responsiveness prediction. We expected that 50% of patients would be fluid responders, so to reach a power of 0.9 with  $\alpha=0.05$ , 33 patients should be included. We decided to include 50 patients to account for uncertainty in the number of fluid responders and to increase precision of the mixed-effects model estimates.

#### Fluid responsiveness prediction

Fluid responsiveness prediction, with  $\text{PPV}_{\text{GAM}}$  or  $\text{PPV}_{\text{Classic}}$  as predictors, was evaluated using receiver operating characteristic (ROC) analysis. Confidence intervals (95%) for area under the ROC curve (AUROC) were calculated using the DeLong-method.<sup>24</sup>

#### Comparison of $\text{PPV}_{\text{Classic}}$ and $\text{PPV}_{\text{GAM}}$

We used Bland-Altman analysis to compare  $\text{PPV}_{\text{Classic}}$  with  $\text{PPV}_{\text{GAM}}$  at each ventilator setting.<sup>25</sup> Limits of agreement (95% LoA) were calculated as  $\text{mean}(\text{PPV}_{\text{GAM}} - \text{PPV}_{\text{Classic}}) \pm 1.96 \cdot \text{SD}(\text{PPV}_{\text{GAM}} - \text{PPV}_{\text{Classic}})$ , where SD is the standard deviation. Confidence intervals

### C. Paper 3

(95%) for bias and LoA were calculated from a nonparametric bootstrap with 4000 resamples.

#### The effect of tidal volume and respiratory rate on pulse pressure variation

To investigate the effect of each ventilator setting on PPV, we fitted a Bayesian mixed-effects model with a patient-specific intercept and a separate variance for each ventilator setting.

In brief, the model describes the relative effects of VT and RR on PPVs measured within a patient. The relative effects are fitted for each of the two PPV methods (GAM and Classic). While the estimated relative differences are the same for all patients (fixed effects), the absolute PPV values can differ between patients (random effect). The physiological understanding of this is that patients have individual Frank-Starling curve operating positions, and that each ventilator setting produces a PPV conditional on the patients position on the Frank-Starling curve.

To increase robustness to outlying values, a Student T likelihood distribution with 4 degrees of freedom was used. The link function was the logarithm, and, consequently, the exponential of the model coefficients are the relative effects on PPV.

A formal model formulation with prior specification and rationale is presented in Supplementary 1.

The model was sampled using Stan, via the R interface *brms*.<sup>23, 26</sup>

Posterior distributions were summarised as median and 95% interval (2.5th to 97.5th percentile). This interval gives a range of values for each parameter that are compatible with the observed data, similar to a confidence interval.<sup>27</sup>

To compare residual standard deviation across different ventilator settings, we calculated the coefficient of variation (CV): the residual standard deviation divided by the expected value of PPV for each ventilator setting.

## Results

From May 2020 to June 2021, we included 52 patients having open abdominal surgery under general anaesthesia. Of these, 50 had a successful measurement of the response to the 250 ml fluid challenge. Ten patients were fluid responders ( $\Delta SV > 10\%$ ). Patient characteristics, vasopressor use and average SV response to fluid are shown in Table 1.

Four patients reached (or were close to) the maximal allowed airway pressure (40 cmH<sub>2</sub>O) at RR=31 min<sup>-1</sup>, VT=6 ml kg<sup>-1</sup>, and did not have the setting: RR=31 min<sup>-1</sup>, VT=8 ml kg<sup>-1</sup> applied. In nine windows (ventilator settings), there were  $\geq 3$  extrasystoles, leaving 507 of 520 potential windows available for analysis. Supplementary 2 shows included and excluded beats and PPV calculation for all ten ventilator settings in all patients.

### Fluid responsiveness prediction

Figure 3 shows scatter plots of PPV<sub>GAM</sub> and the corresponding fluid response ( $\Delta SV$ ) for all ventilator settings (a similar figure for PPV<sub>Classic</sub> is available in Supplementary 3 Fig. S1). The capacity of PPV<sub>GAM</sub> and PPV<sub>Classic</sub> to classify fluid responsiveness ( $\Delta SV > 10\%$ ) is presented as ROC curves in Supplementary 3 Fig. S2. At the ventilator setting RR=10 min<sup>-1</sup>, VT=10 ml kg<sup>-1</sup>, PPV<sub>GAM</sub> had an area under the ROC curve (AUC) of 0.73 [0.57;0.90], while the AUC for PPV<sub>Classic</sub> was 0.74 [0.57;0.92]. At RR=31 min<sup>-1</sup>, VT=6 ml kg<sup>-1</sup>, PPV<sub>GAM</sub> had an AUC of 0.65 [0.45;0.85], while the AUC for PPV<sub>Classic</sub> was 0.62 [0.40;0.84]. Supplementary 3 Table S1 presents AUC, optimal PPV threshold, sensitivity and specificity for fluid responsiveness discrimination for all ten ventilator settings.

### Comparison of PPV<sub>Classic</sub> and PPV<sub>GAM</sub>

At ventilator setting RR=10 min<sup>-1</sup>, VT=10 ml kg<sup>-1</sup>, PPV<sub>GAM</sub> was, on average, slightly lower than PPV<sub>Classic</sub>: mean difference (bias) = -0.36, 95% CI [-0.75; -0.08]; limits of agreement (95%) were -2.87 to 2.16. Bland-Altman plots comparing PPV<sub>Classic</sub> and PPV<sub>GAM</sub> for all ten ventilator settings are presented in Supplementary 3 Fig. S3.

The relationship between PPV and HR/RR is presented in Fig. 4. At HR/RR below 3.6, PPV<sub>GAM</sub> were generally higher than PPV<sub>Classic</sub>: bias = 0.93, 95% CI [0.76; 1.11]; limits of agreement (95%) were -1.73 to 3.59. At HR/RR above 3.6, PPV<sub>GAM</sub> and PPV<sub>Classic</sub> gave very similar values: bias = -0.09 [-0.23; 0.03]; limits of agreement (95%) was -1.85 to 1.67.

## The effect of tidal volume and respiratory rate on pulse pressure variation

Model parameters are visualised in Fig. 5. Estimates of the effects of VT=10, 8 or 6 ml kg<sup>-1</sup> were very close to a direct proportionality between VT and PPV for both PPV<sub>classic</sub> and PPV<sub>GAM</sub>. Relative to PPV<sub>GAM</sub> at VT=10 ml kg<sup>-1</sup>, PPV<sub>GAM</sub> at VT=8 ml kg<sup>-1</sup> was 81 [77; 86]% and PPV<sub>GAM</sub> at VT=6 ml kg<sup>-1</sup> was 64 [60; 67]%. At VT=4 ml kg<sup>-1</sup>, PPV<sub>GAM</sub> was 49 [46; 53]% and not compatible with the 40% expected from a direct proportionality between VT and PPV. The effect of VT on PPV<sub>classic</sub> was similar.

Higher RR was associated with lower PPV, and the effect was most pronounced for PPV<sub>classic</sub>: at RR=31 min<sup>-1</sup>, PPV<sub>classic</sub> was 56 [52; 61]% and PPV<sub>GAM</sub> was 81 [76; 86]%, both relative to at RR=10 min<sup>-1</sup>.

The residual variation is visualised in Supplementary 3 Fig. S4. The relative variation of the observations around the the model predictions (CV of the residuals) was similar between PPV<sub>Classic</sub> and PPV<sub>GAM</sub>, except for at RR=31 min<sup>-1</sup>, where PPV<sub>classic</sub> had a higher uncertainty (the difference in CV is 16 [3; 30]%-points). The CV was lowest at RR=17 min<sup>-1</sup> for both PPV<sub>classic</sub> and PPV<sub>GAM</sub>.

## Discussion

This study had two overall aims. First, we sought to describe VT's and RR's impact on PPV's ability to predict fluid responsiveness. Second, we investigated VT's and RR's impact on PPV directly.

### Fluid responsiveness prediction

Unfortunately, not much can be derived about PPV's ability to predict fluid responsiveness from this study, mainly due to the low number of responders. Point estimates for both  $\text{PPV}_{\text{GAM}}$  and  $\text{PPV}_{\text{Classic}}$  showed fluid responsiveness prediction with mediocre/poor accuracy, even when patients were ventilated at  $\text{VT}=10 \text{ ml kg}^{-1}$  and  $\text{RR}=10 \text{ min}^{-1}$ . Most confidence intervals were compatible with AUCs from 0.6 to 0.9 (poor to excellent classification).

### The effects of VT and RR on PPV

Based on basic Bland-Altman analysis (Supplementary 3 Fig. S3), we demonstrated that PPV derived from GAM and the classic approach are very similar at  $\text{VT} \geq 6 \text{ ml kg}^{-1}$  and  $\text{RR} \leq 17 \text{ min}^{-1}$ .

The mixed-effects model demonstrated that within the same patient, PPV was nearly proportional to VT across various levels of VT and RR. The VT effect was similar for  $\text{PPV}_{\text{GAM}}$  and  $\text{PPV}_{\text{Classic}}$ . On the other hand, high RR reduced  $\text{PPV}_{\text{Classic}}$  markedly more than it reduced  $\text{PPV}_{\text{GAM}}$ . This difference probably reflects a sampling effect rather than a physiological effect. At high RR,  $\text{PPV}_{\text{Classic}}$  is calculated from very few beats per respiratory cycle, increasing the risk of a falsely low PPV. This result is in accordance with the original study investigating the HR/RR ratio effect, which found that PPV (corresponding to  $\text{PPV}_{\text{Classic}}$  in the present paper) dropped markedly when the HR/RR was  $< 3.6$ .<sup>4</sup> Conversely,  $\text{PPV}_{\text{GAM}}$  works well in low HR/RR situations, but can underestimate PPV when HR/RR is an exact low integer ratio (e.g. exactly 2 beats per ventilation). This difference between  $\text{PPV}_{\text{Classic}}$  and  $\text{PPV}_{\text{GAM}}$  at low HR/RR is also illustrated in Fig. 4.

The model allows us to account for ventilator settings in the interpretation of PPV. As an example, consider a patient ventilated at a VT of  $6 \text{ ml kg}^{-1}$  and RR of  $24 \text{ min}^{-1}$ . We estimate a PPV of 8% with the GAM method. The best guess of what PPV would be if VT was changed to  $10 \text{ ml kg}^{-1}$  and RR to  $10 \text{ min}^{-1}$  (a setting where an optimal PPV threshold seems established<sup>2</sup>), is then

### C. Paper 3

$$PPV_{RR=10,VT=10} = 8\% \cdot \frac{1}{0.64} \cdot \frac{1}{0.87} \approx 14\%$$

For  $PPV_{GAM}$ , a pragmatic bedside approximation would be to consider PPV directly proportional to VT and disregard the effect of RR:

$$PPV_{RR=10,VT=10} = PPV \cdot \frac{10}{VT},$$

where PPV is the current PPV and VT is the current VT in  $\text{ml kg(pbw)}^{-1}$ . This approximation works because for reciprocal changes in RR and VT (approximately maintaining minute ventilation), the overcorrection from considering the effect of VT as proportional, closely matches the effect of the RR change.

The CV was similar for  $PPV_{GAM}$  and  $PPV_{Classic}$ , except when RR was high. Here  $PPV_{Classic}$  had a significantly higher CV. This is in accordance with the sampling problem affecting  $PPV_{Classic}$  described above.

## Limitations

The study had relatively few fluid responders (20%). The aim of including patients when no interventions were expected may have resulted in a largely fluid-optimised population. Also, the uncalibrated pulse contour analysis estimate of CO is clinically acceptable and probably one of the most used CO modalities in GDT protocols,<sup>29</sup> but it is not the gold standard for measuring CO.

Regardless of the cause, the poor predictive performance of PPV precludes a meaningful investigation of the hypothesised advantage of using  $PPV_{GAM}$  to predict fluid responsiveness at low HR/RR.

In accordance with fluid responses, PPV values were relatively low. We do not know whether the relative effects of RR and VT found in this study also apply to patients with higher PPV. Also, the results may not generalise to a population with different HR and lung compliance, such as ICU patients with ARDS.

## Conclusion

We demonstrate that the current understanding of ventilator settings' impact on PPV is insufficient. The limitation associated with low HR/RR seems to be predominantly related to a specific method of deriving PPV rather than a physiological limitation. At high RRs PPV should be estimated over multiple respiratory cycles to avoid a basic sampling problem.

*C. Paper 3*

Also, PPV is nearly proportional to VT, suggesting that correcting PPV for VT might make the optimal threshold less dependent on VT, improving the utility of PPV. However, it was not possible to demonstrate whether PPV based on GAM modelling would result in better prediction of fluid responsiveness compared with the classical method of deriving PPV.

## Authors' contributions

Study conception/design: all authors

Patient recruitment and data collection: JE, BB, HVT

Data analysis: JE, STV

Data interpretation: all authors

Writing the first draft of the paper: JE, STV

Critical revision the paper for important intellectual content and approval the final version: all authors

## Acknowledgements

The authors thank J. Karippacheril, Cleveland Clinic, Abu Dhabi for collaborating on extending his open source software, VitalSignCapture, to export real time data from Dräger ventilators.

## Declaration of interests

STV is associate editor of Journal of Clinical Monitoring and Computing. TWLS received research grants and honoraria from Edwards Lifesciences (Irvine, CA, USA) and Masimo Inc. (Irvine, CA, USA) for consulting and lecturing (all payments made to institution). TWLS is chair of the Cardiovascular Dynamics section of the European Society of Intensive Care Medicine (ESICM), and associate editor of the Journal of Clinical Monitoring and Computing and the journal Anesthesia & Analgesia. TWLS is currently working as Senior Medical Director for Edwards Lifesciences (Unterschleißheim, Germany).

## Funding

This work was supported by Aarhus University and Holger & Ruth Hesse's Mindefond.

## References

1. Monnet X, Marik PE, Teboul JL. Prediction of fluid responsiveness: an update. *Ann. Intensive Care.* 2016; **6**: 1–11
2. 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
3. Michard F, Chemla D, Teboul J-L. Applicability of pulse pressure variation: how many shades of grey? *Crit. Care.* 2015; **19**: 15–7
4. De Backer D, Ph D, Taccone FS, et al. Influence of Respiratory Rate on Stroke Volume Variation in Mechanically Ventilated Patients. *Anesthesiology.* 2009; **110**: 1092–7
5. De Backer D, Heenen S, Piagnerelli M, Koch M, Vincent JL. Pulse pressure variations to predict fluid responsiveness: Influence of tidal volume. *Intensive Care Med.* 2005; **31**: 517–23
6. Monnet X, Bleibtreu A, Ferré A, et al. Passive leg-raising and end-expiratory occlusion tests perform better than pulse pressure variation in Patients with low respiratory system compliance. *Crit. Care Med.* 2012; **40**: 152–7
7. Alvarado Sánchez JI, Caicedo Ruiz JD, Diaztagle Fernández JJ, Ospina-Tascón GA, Cruz Martínez LE. Use of Pulse Pressure Variation as Predictor of Fluid Responsiveness in Patients Ventilated With Low Tidal Volume: A Systematic Review and Meta-Analysis. *Clin. Med. Insights Circ. Respir. Pulm. Med.* 2020; **14**: 1–10
8. Liu Y, Lou J sheng, Mi W dong, et al. Pulse pressure variation shows a direct linear correlation with tidal volume in anesthetized healthy patients. *BMC Anesthesiol.* 2016; **16**: 1–7
9. Vallée F, Richard JCM, Mari A, et al. Pulse pressure variations adjusted by alveolar driving pressure to assess fluid responsiveness. *Intensive Care Med.* 2009; **35**: 1004–10
10. Vistisen ST, Koefoed-Nielsen J, Larsson A. Should dynamic parameters for prediction of fluid responsiveness be indexed to the tidal volume? *Acta Anaesthesiol. Scand.* 2010; **54**: 191–8
11. Liu Y, Wei L, Li G, Yu X, Li G, Li Y. Pulse Pressure Variation Adjusted by Respiratory Changes in Pleural Pressure, Rather Than by Tidal Volume, Reliably Predicts Fluid Responsiveness in Patients With Acute Respiratory Distress Syndrome\*: *Crit. Care Med.* 2016; **44**: 342–51
12. Michard F, Chemla D, Richard C, et al. Clinical Use of Respiratory Changes in Arterial Pulse Pressure to Monitor the Hemodynamic Effects of PEEP. *Am. J. Respir. Crit. Care Med.* 1999; **159**: 935–9
13. Kim HK, Pinsky MR. Effect of tidal volume, sampling duration, and cardiac contractility on pulse pressure and stroke volume variation during positive-pressure ventilation. *Crit. Care Med.* 2008; **36**: 2858–62
14. Enevoldsen J, Simpson GL, Vistisen ST. Using generalized additive models to decompose time series and waveforms, and dissect heart-lung interaction physiology. *J. Clin. Monit. Comput.* 2022; Available from: <https://doi.org/10.1007/s10877-022-00873-7>
15. 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
16. The Acute Respiratory Distress Syndrome Network. Ventilation with Lower Tidal Volumes as Compared with Traditional Tidal Volumes for Acute Lung Injury and the Acute Respiratory Distress Syndrome. *N. Engl. J. Med.* 2000; **8** Available from
17. 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

C. Paper 3

18. Karippacheril JG, Ho TY. Data acquisition from S/5 GE Datex anesthesia monitor using VSCapture: An open source.NET/Mono tool. *J. Anaesthesiol. Clin. Pharmacol.* 2013; **29**: 423–4
19. R Core Team. R: A Language and Environment for Statistical Computing [Internet]. 2021. Available from: <https://www.R-project.org/>
20. Wickham H, Averick M, Bryan J, et al. Welcome to the Tidyverse. *J. Open Source Softw.* 2019; **4**: 1686
21. 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
22. Canty A, Ripley BD. boot: Bootstrap R (S-Plus) Functions [Internet]. 2021. Available from: <https://CRAN.R-project.org/package=boot>
23. Bürkner P-C. brms: An R Package for Bayesian Multilevel Models Using Stan. *J. Stat. Softw.* 2017; **80**: 1–28
24. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics.* 1988; **44**: 837–45
25. Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat. Methods Med. Res.* 1999; **8**: 135–60
26. Carpenter B, Gelman A, Hoffman MD, et al. Stan: A Probabilistic Programming Language. *J. Stat. Softw.* 2017; **76**: 1–32
27. Gelman A, Greenland S. Are confidence intervals better termed “uncertainty intervals”? *BMJ.* 2019; **366**: I5381
28. Vistisen ST, Berg JM, Boekel MF, et al. Using extra systoles and the micro-fluid challenge to predict fluid responsiveness during cardiac surgery. *J. Clin. Monit. Comput.* 2019; **33**: 777–86
29. Jessen MK, Vallentin MF, Holmberg MJ, et al. Goal-directed haemodynamic therapy during general anaesthesia for noncardiac surgery: a systematic review and meta-analysis. *Br. J. Anaesth.* 2022; **128**: 416–33

## Tables

Table 1: Characteristics of patients included in the study.<sup>1</sup> Median (IQR); n (%).<sup>2</sup> Two patients received no fluid challenge.<sup>3</sup> Surgical procedures are counted in the first matching category. ASA score, American Society of Anesthesiologists physical status classification system. SV, Stroke volume estimate from FloTrac® pulse contour analysis. MAP, Mean arterial pressure. HIPEC, Hyperthermic intraperitoneal chemotherapy. APE, Abdomino-perineal excision. VRAM, Vertical rectus abdominis myocutaneous flap. pbw, predicted body weight.

Variable	$\Delta SV \leq 10\%$ , N = 40 <sup>1</sup>	$\Delta SV > 10\%$ , N = 10 <sup>1</sup>	Total, N = 52 <sup>2</sup>
Age	65 (57, 73)	61 (55, 69)	64 (57, 72)
Sex			
Female	21 (52%)	4 (40%)	26 (50%)
Male	19 (48%)	6 (60%)	26 (50%)
Height [cm]	173 (166, 179)	180 (165, 184)	173 (165, 180)
Weight [kg]	78 (70, 82)	78 (62, 87)	79 (70, 84)
Body mass index	24.8 (23.4, 28.2)	22.8 (21.6, 25.0)	24.7 (23.3, 28.2)
Predicted body weight (pbw, kg)	67 (58, 74)	75 (57, 79)	68 (57, 75)
Known hypertension	17 (42%)	1 (10%)	19 (37%)
ASA score			
1	14 (35%)	6 (60%)	21 (40%)
2	22 (55%)	3 (30%)	25 (48%)
3	3 (7.5%)	1 (10%)	5 (9.6%)
4	1 (2.5%)	0 (0%)	1 (1.9%)
Surgical procedure <sup>3</sup>			
HIPEC	27 (68%)	7 (70%)	36 (69%)
APE and/or VRAM	7 (18%)	1 (10%)	8 (15%)

C. Paper 3

Colon resection	2 (5.0%)	0 (0%)	2 (3.8%)
Other	4 (10%)	2 (20%)	6 (12%)
Fluid type for fluid challenge			
Acetated Ringer's solution	30 (75%)	6 (60%)	36 (69%)
Human albumin	10 (25%)	4 (40%)	14 (27%)
No fluid challenge	0 (0%)	0 (0%)	2 (3.8%)
Any vasopressor during study protocol	36 (90%)	10 (100%)	47 (90%)
Noradrenaline rate [ $\mu\text{g kg}^{-1} \text{ min}^{-1}$ ]			
0	14 (35%)	6 (60%)	21 (40%)
<0.1	19 (48%)	4 (40%)	23 (44%)
$\geq 0.1$	7 (18%)	0 (0%)	8 (15%)
Dopamine rate [ $\mu\text{g kg}^{-1} \text{ min}^{-1}$ ]			
0	22 (55%)	4 (40%)	28 (54%)
<5	17 (42%)	5 (50%)	22 (42%)
$\geq 5$	1 (2.5%)	1 (10%)	2 (3.8%)
Pre-intervention ventilation			
Tidal volume [ml $\text{kg}^{-1}$ (pbw)]	7.2 (6.9, 7.8)	6.8 (6.4, 8.0)	7.2 (6.9, 7.8)
Respiratory rate [ $\text{min}^{-1}$ ]	14 (14, 16)	14 (14, 16)	14 (14, 16)
Positive end expiratory pressure [cmH <sub>2</sub> O]	5 (5, 6)	6 (5, 6)	5 (5, 6)
Fluid challenge			
SV before fluid challenge [mL]	71 (60, 82)	58 (54, 66)	68 (59, 81)
SV response to fluid challenge [%]	2.1 (0.6, 5.4)	17.3 (13.7, 20.1)	3.8 (1.0, 8.9)

*C. Paper 3*

MAP before fluid challenge [mmHg]	70 (64, 74)	68 (62, 73)	69 (63, 74)
MAP response to fluid challenge [%]	8 (4, 15)	7 (1, 18)	8 (3, 17)

## Figures

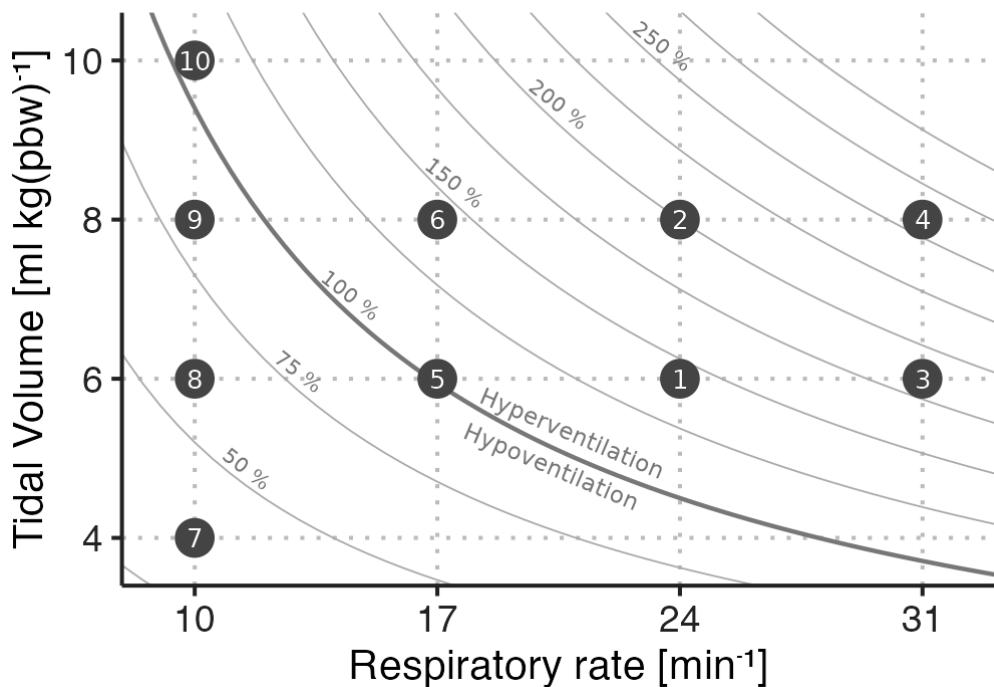


Figure 1: Example of the series of ventilator settings applied. Each numbered dot represents a combination of tidal volume (normalised to predicted body weight (pbw)) and respiratory rate. The numbers represent an example of the order of application of the settings. For each respiratory rate, tidal volumes were applied from low to high. Respiratory rates 17, 24 and 31  $\text{min}^{-1}$  were applied in random order. Respiratory rate 10  $\text{min}^{-1}$  was always applied last to avoid having any recruitment effect from the highest tidal volume ( $10 \text{ ml kg(pbw)}^{-1}$ ) influence the remainder of the recording. Curved lines represent settings with equal alveolar ventilation, assuming a dead space volume of  $1 \text{ ml kg(pbw)}^{-1}$ . The curved lines' labels denote the alveolar ventilation relative to ventilation with a respiratory rate of 14 and a tidal volume of  $7 \text{ ml kg(pbw)}^{-1}$ .

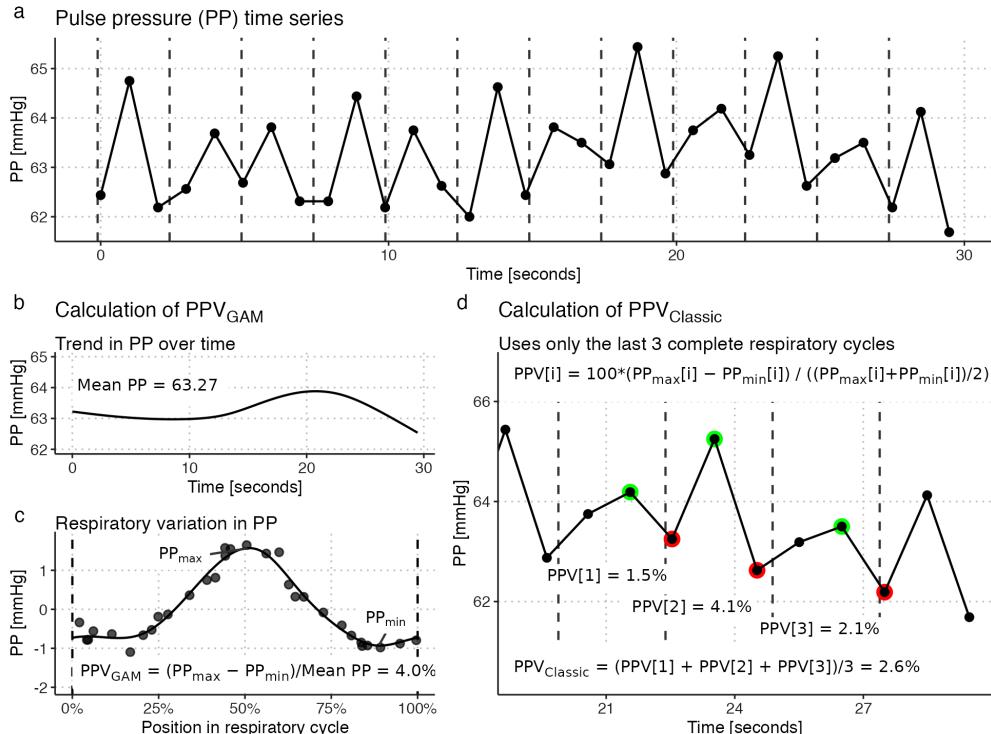


Figure 2: Illustration of the two methods used to calculate pulse pressure variation (PPV). The patient was ventilated with a VT of  $8 \text{ ml kg}^{-1}$  and a respiratory rate of  $24 \text{ min}^{-1}$ . Panel **a**: a 30 second time series of pulse pressure (PP) measurements is available for PPV calculation at each ventilator setting. Panels **b** and **c** illustrate calculation of PPV with a generalized additive model (GAM): the 30 second time series is decomposed into a trend in PP over time (**b**) and the cyclic variation in PP with each respiratory cycle (**c**).  $PPV_{GAM}$  is the variation in PP that is explained by the respiratory component (**c**). Panel **d** illustrates the classic calculation of PPV.

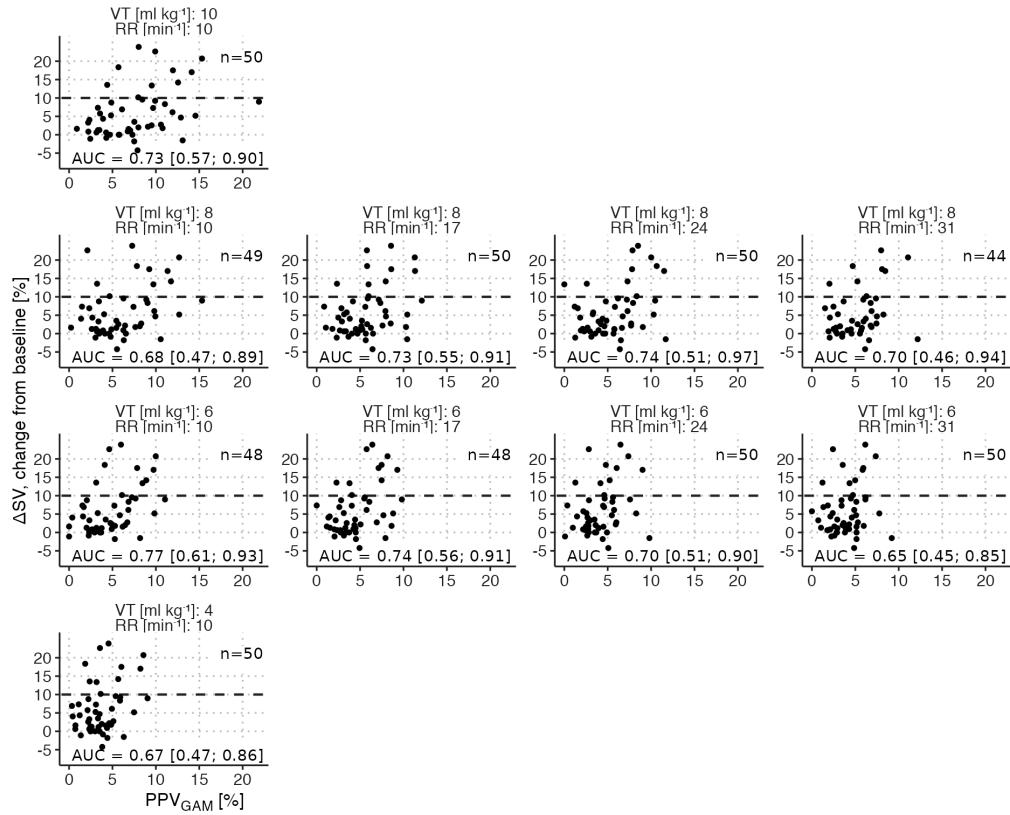


Figure 3: Scatter plots of the relation between PPV derived from a generalized additive model ( $PPV_{GAM}$ ) and the stroke volume response ( $\Delta SV$ ) to a 250 ml fluid challenge. Panels are arranged with tidal volumes (VT) in rows and respiratory rates (RR) in columns. One fluid challenge was evaluated for each patient ( $n = 50$ ), while  $PPV_{GAM}$  was calculated for each of the 10 ventilator settings.

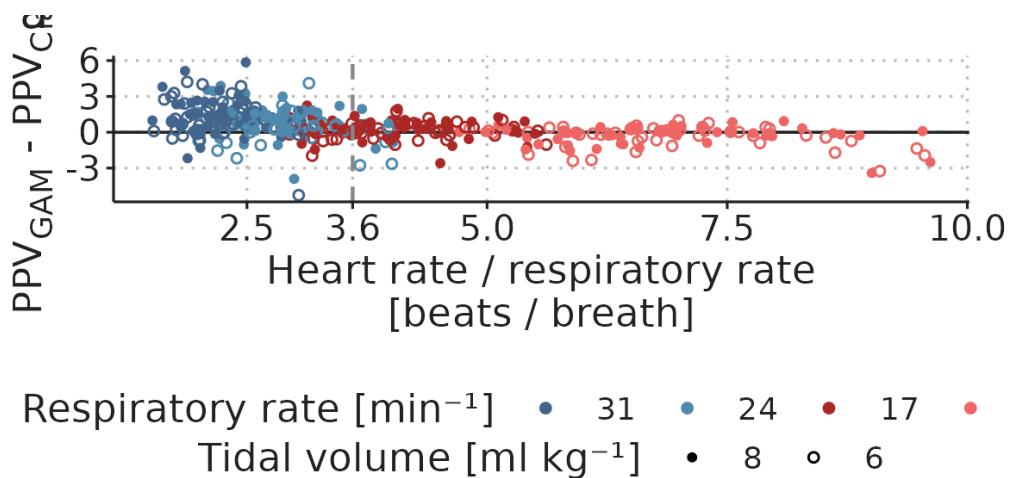


Figure 4: The relationship between heart rate / respiratory rate and PPV when PPV is calculated using the Classic approach (a) and using a generalized additive model (GAM) (b). Panel c shows the difference between PPV<sub>GAM</sub> and PPV<sub>Classic</sub> calculated on the same data (403 total observations in 50 patients).

C. Paper 3

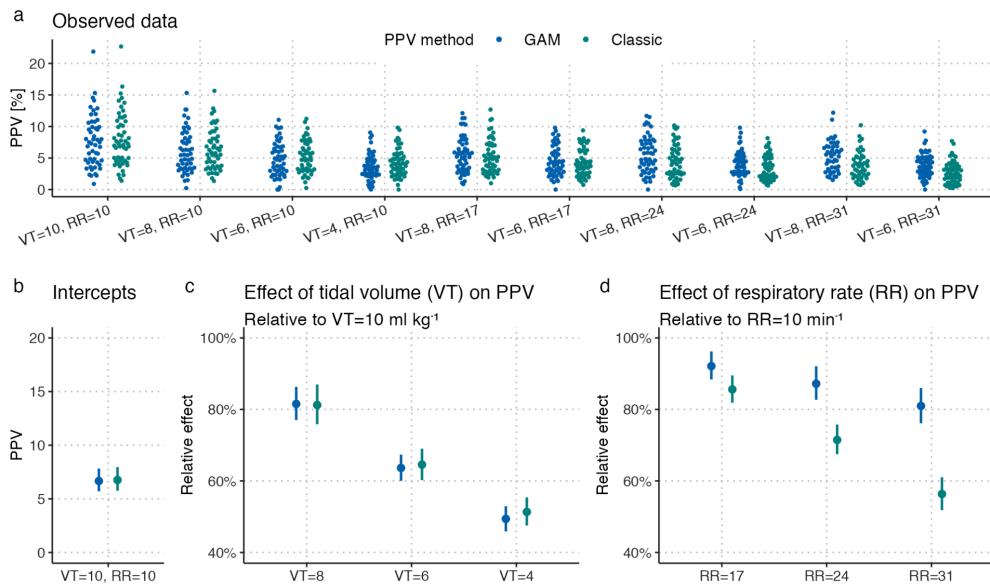


Figure 5: Parameter estimates for a Bayesian mixed-effects model, describing the effect of tidal volume (VT) and respiratory rate (RR) on pulse pressure variation (PPV). Parameters are estimated for both PPV derived using a generalized additive model (GAM) and using a classic approach (Classic). Panel **a** presents the observed PPV values (outcomes) using each method ( $n=507$  for both GAM and Classic). Panel **b**, **c** and **d** present parameter estimates. Vertical bars are 95% compatibility intervals.

*C. Paper 3*

### **C.1 Paper 3 - Supplementary figures and tables**

The following pages contain Paper 3's Supplementary 3. Supplementary 1 and 2, data and analysis code are available at [DOI:10.5281/zenodo.6984310](https://doi.org/10.5281/zenodo.6984310).

## Supplementary figures and tables

**Figure S1**

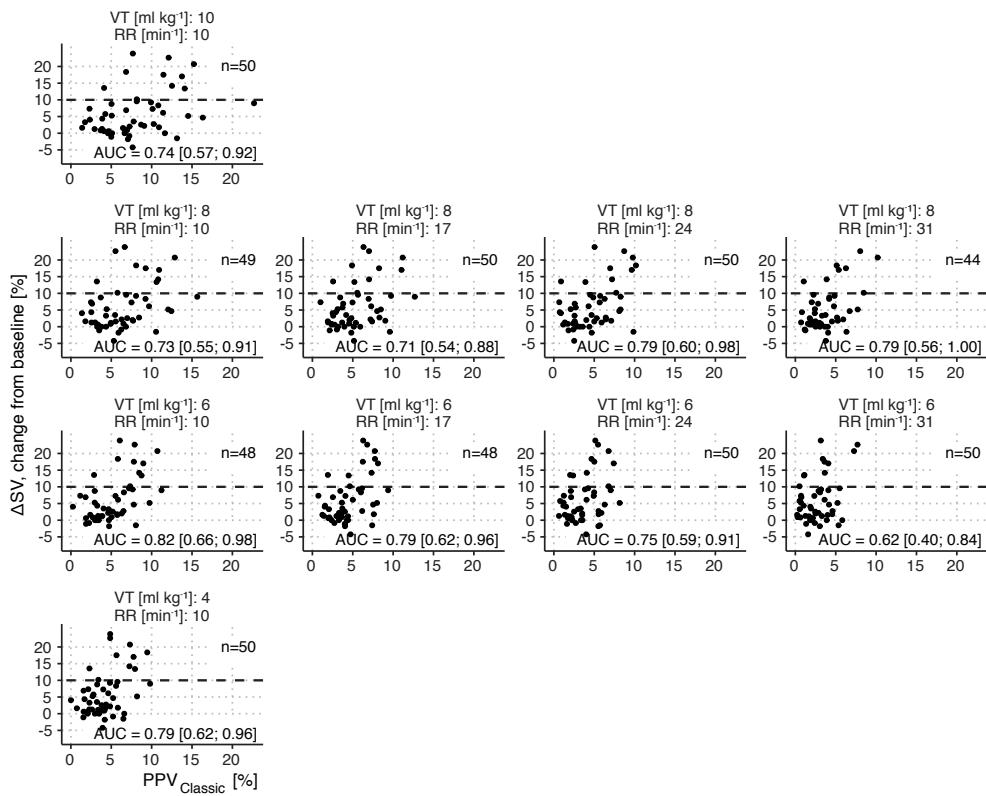


Figure S1: Scatter plots of the relation between PPV calculated with the classic algorithm ( $PPV_{Classic}$ ) and the stroke volume response ( $\Delta SV$ ) to a 250 ml fluid challenge. Panels are arranged with tidal volumes (VT) in rows and respiratory rates (RR) in columns . One fluid challenge was evaluated for each subject ( $n = 50$ ), while  $PPV_{Classic}$  was calculated for each of the 10 ventilator settings.

**Figure S2**

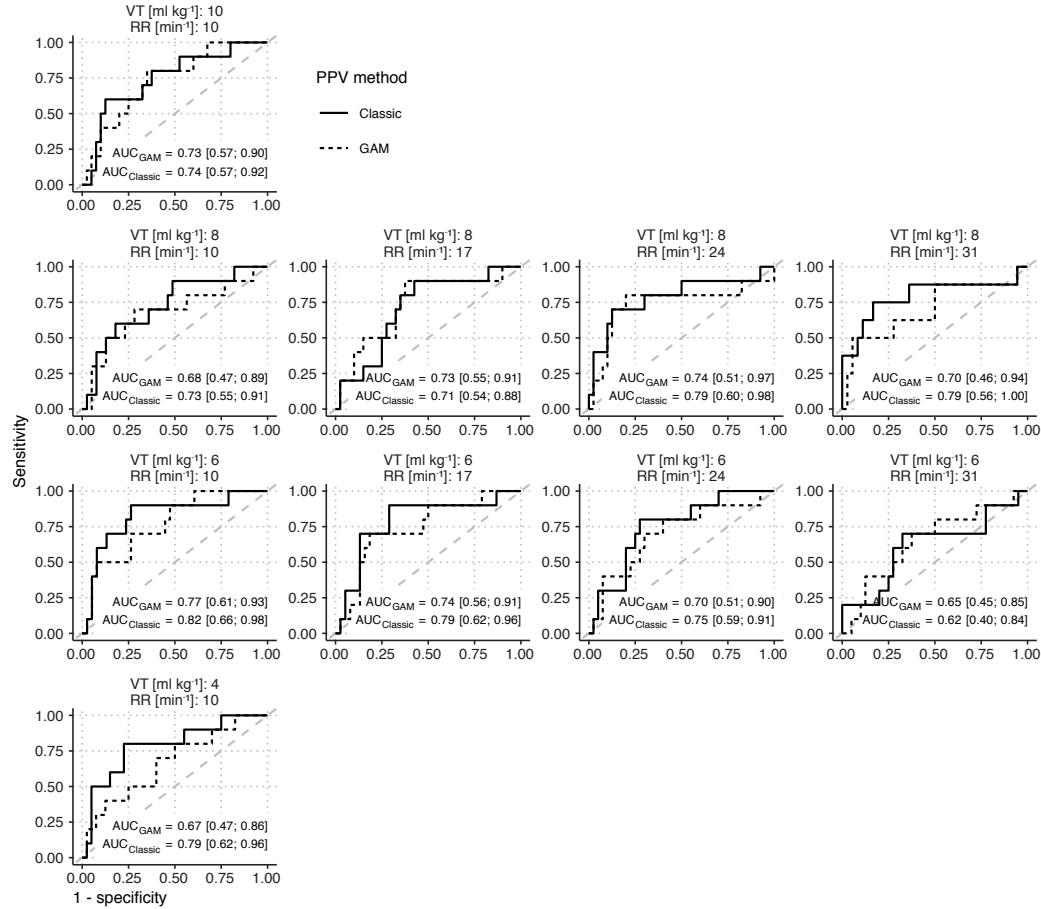


Figure S2: Receiver operating characteristic (ROC) curves for capacity of pulse pressure variation (PPV) to classify fluid responsiveness (stroke volume change > 10%). Panels are arranged with tidal volumes (VT) in rows and respiratory rates (RR) in columns. One fluid challenge was evaluated for each subject ( $n = 50$ ), while  $\text{PPV}_{\text{GAM}}$  and  $\text{PPV}_{\text{Classic}}$  was calculated for each of the 10 ventilator settings.

C. Paper 3

**Figure S3**

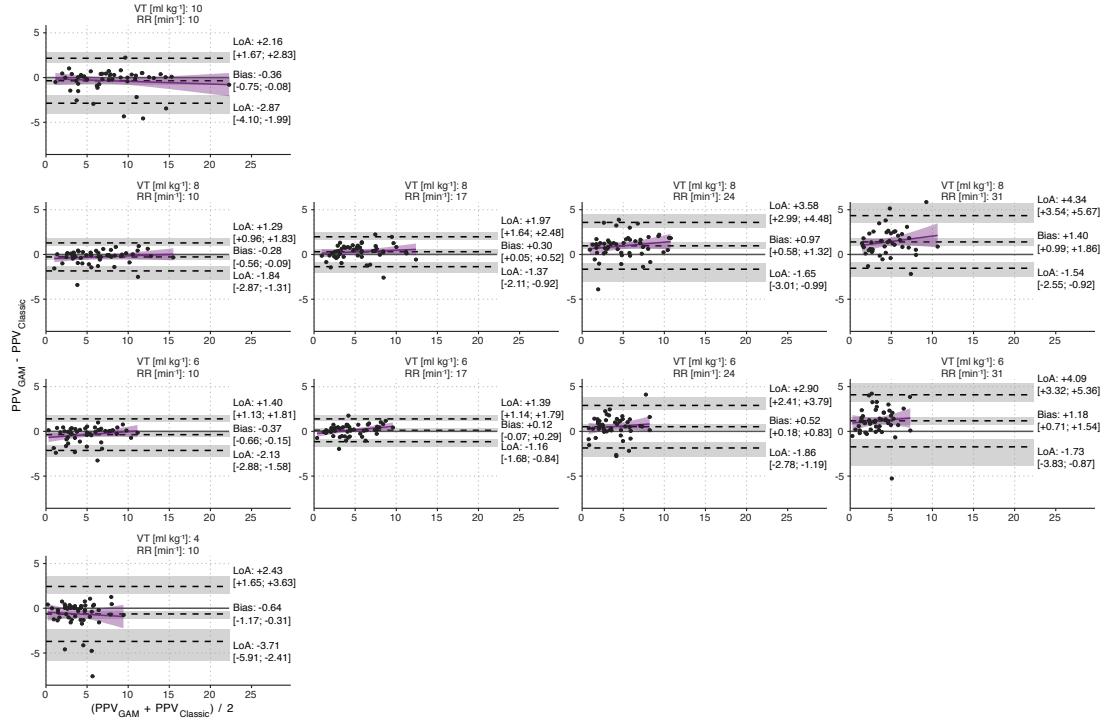


Figure S3: Bland-Altman plots showing the relation between PPV<sub>GAM</sub> and PPV<sub>Classic</sub>. The outer dashed lines represent 95% limits of agreement (LoA). Grey areas are 95% confidence intervals for bias and LoA. The purple lines and areas are linear regression fits with 95% confidence intervals. Panels are arranged with tidal volumes (VT) in rows and respiratory rates (RR) in columns.

**Figure S4**

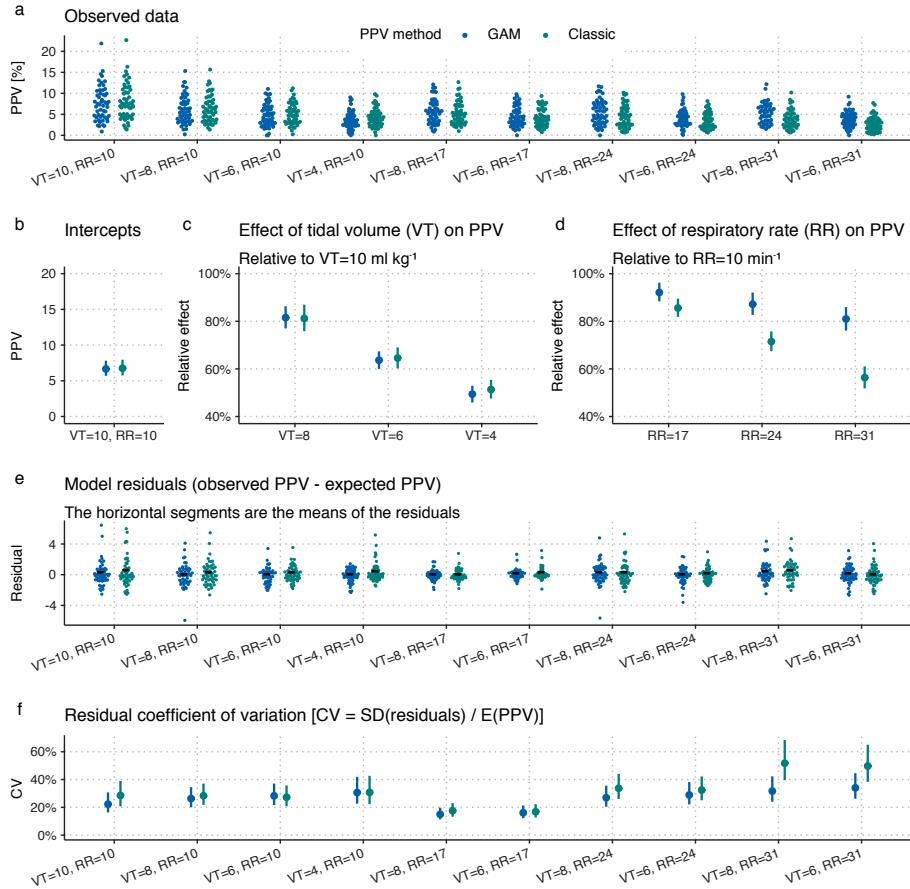


Figure S4: (Extended version of the paper's Fig. 5): Parameter estimates for a Bayesian mixed-effects model, describing the effect of tidal volume (VT) and respiratory rate (RR) on pulse pressure variation (PPV). Parameters are estimated for both PPV derived using a generalized additive model (GAM) and using a classic approach (Classic). Panel **a** presents the observed PPV values (outcomes) using each method ( $n=507$  for both GAM and Classic). Vertical bars are 95% compatibility intervals. Panel **b**, **c** and **d** present parameter estimates. Vertical bars are 95% compatibility intervals. Panel **e** shows model residuals, and panel **f** shows the residual variation relative to the estimated value of PPV.

C. Paper 3

**Table S1**

Receiver operating characteristic analysis of fluid-responsiveness prediction using pulse pressure variation (PPV). A positive fluid response was a >10% increase in stroke volume from a 250 ml fluid bolus. PPV was calculated using a classic method and derived from a generalized additive model (GAM). Results are presented as *estimate* [95% confidence interval].

Tidal volume [ml kg <sup>-1</sup> ]	Respiratory rate [min <sup>-1</sup> ]	GAM PPV				Classic PPV			
		Optimal threshold <sup>†</sup>	Sensitivity	Specificity	AUC	Optimal threshold <sup>†</sup>	Sensitivity	Specificity	AUC
10	10	7.9% [0.50;1.00]	0.80 [0.50;0.80]	0.65 [0.50;0.80]	0.73 [0.57;0.90]	11.4% [0.30;0.90]	0.60 [0.75;0.97]	0.88 [0.57;0.92]	0.74
8	10	6.9% [0.40;0.90]	0.70 [0.56;0.85]	0.72 [0.47;0.89]	0.68 [0.47;0.89]	8.1% [0.30;0.90]	0.60 [0.69;0.92]	0.82 [0.55;0.91]	0.73
6	10	5.9% [0.40;1.00]	0.70 [0.61;0.87]	0.74 [0.61;0.93]	0.77 [0.61;0.93]	5.8% [0.70;1.00]	0.90 [0.61;0.87]	0.74 [0.66;0.98]	0.82
4	10	3.1% [0.50;1.00]	0.80 [0.35;0.65]	0.50 [0.35;0.65]	0.67 [0.47;0.86]	4.9% [0.50;1.00]	0.80 [0.65;0.90]	0.78 [0.62;0.96]	0.79
8	17	5.7% [0.70;1.00]	0.90 [0.47;0.78]	0.62 [0.47;0.78]	0.73 [0.55;0.91]	4.9% [0.70;1.00]	0.90 [0.42;0.72]	0.57 [0.54;0.88]	0.71
6	17	5.7% [0.40;1.00]	0.70 [0.68;0.92]	0.82 [0.56;0.91]	0.74 [0.56;0.91]	4.4% [0.70;1.00]	0.90 [0.58;0.84]	0.71 [0.62;0.96]	0.79
8	24	7.2% [0.50;1.00]	0.80 [0.68;0.93]	0.80 [0.51;0.97]	0.74 [0.51;0.97]	6.7% [0.40;0.90]	0.70 [0.78;0.97]	0.88 [0.60;0.98]	0.79
6	24	4.6% [0.40;1.00]	0.70 [0.55;0.82]	0.70 [0.51;0.90]	0.70 [0.51;0.90]	4.1% [0.50;1.00]	0.80 [0.57;0.85]	0.72 [0.59;0.91]	0.75
8	31	7.7% [0.12;0.88]	0.50 [0.86;1.00]	0.94 [0.46;0.94]	0.70 [0.46;0.94]	5.0% [0.50;1.00]	0.75 [0.69;0.94]	0.83 [0.56;1.00]	0.79
6	31	4.3% [0.40;1.00]	0.70 [0.47;0.78]	0.62 [0.45;0.85]	0.65 [0.45;0.85]	3.1% [0.40;1.00]	0.70 [0.53;0.82]	0.68 [0.40;0.84]	0.62

<sup>†</sup> Threshold with maximum Youden index.

# D

Declarations of co-authorship

D. Declarations of co-authorship



**Declaration of co-authorship concerning article for PhD dissertations**

Full name of the PhD student: Johannes Aagaard Enevoldsen

This declaration concerns the following article/manuscript:

Title:	Existing fluid responsiveness studies using the mini-fluid challenge may be misleading: Methodological considerations and simulations
Authors:	Johannes Enevoldsen, Thomas W. L. Scheeren, Jonas M. Berg, Simon T. Vistisen

The article/manuscript is: Published  Accepted  Submitted  In preparation

If published, state full reference: 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. 2022; 66: 17– 24. <https://doi.org/10.1111/aas.13965>

If accepted or submitted, state journal: Acta Anaesthesiologica Scandinavica

Has the article/manuscript previously been used in other PhD or doctoral dissertations?

No  Yes  If yes, give details:

**Your contribution**

Please rate (A-F) your contribution to the elements of this article/manuscript, **and** elaborate on your rating in the free text section below.

- A. Has essentially done all the work (>90%)
- B. Has done most of the work (67-90 %)
- C. Has contributed considerably (34-66 %)
- D. Has contributed (10-33 %)
- E. No or little contribution (<10%)
- F. N/A

Category of contribution	Extent (A-F)
The conception or design of the work:	C
<i>Free text description of PhD student's contribution (mandatory)</i>	
The problem was originally identified by Simon Vistisen. Johannes helped specify it into two sub-problems and proposed a simulation-based presentation	
The acquisition, analysis, or interpretation of data:	B
<i>Free text description of PhD student's contribution (mandatory)</i>	
Johannes created the simulations of the problem, and, with Simon Vistisen, formulated the reanalysis of data from an existing MFC study.	
Drafting the manuscript:	B
<i>Free text description of PhD student's contribution (mandatory)</i>	
Johannes formulated the first draft with significant contributions from Simon Vistisen. Johannes prepared the manuscript for submission after revisions from co-authors.	
Submission process including revisions:	A

*D. Declarations of co-authorship*



*Free text description of PhD student's contribution (mandatory)*

Johannes submitted the manuscript and handled revisions after discussion with co-authors.

**Signatures of first- and last author, and main supervisor**

Date	Name	Signature
26-10-2022	Johannes Enevoldsen	
26-10-2022	Simon Vistisen	

Date: 26-10-2022

  
\_\_\_\_\_  
Signature of the PhD student

*D. Declarations of co-authorship*



**Declaration of co-authorship concerning article for PhD dissertations**

Full name of the PhD student: Johannes Aagaard Enevoldsen

This declaration concerns the following article/manuscript:

Title:	Using generalized additive models to decompose time series and waveforms, and dissect heart-lung interaction physiology
Authors:	Johannes Enevoldsen, Gavin L. Simpson, Simon T. Vistisen

The article/manuscript is: Published  Accepted  Submitted  In preparation

If published, state full reference: Enevoldsen, J., Simpson, G.L. & Vistisen, S.T. Using generalized additive models to decompose time series and waveforms, and dissect heart-lung interaction physiology. *J Clin Monit Comput* (2022). <https://doi.org/10.1007/s10877-022-00873-7>

If accepted or submitted, state journal: Journal of Clinical Monitoring and Computing

Has the article/manuscript previously been used in other PhD or doctoral dissertations?

No  Yes  If yes, give details:

**Your contribution**

Please rate (A-F) your contribution to the elements of this article/manuscript, **and** elaborate on your rating in the free text section below.

- A. Has essentially done all the work (>90%)
- B. Has done most of the work (67-90 %)
- C. Has contributed considerably (34-66 %)
- D. Has contributed (10-33 %)
- E. No or little contribution (<10%)
- F. N/A

Category of contribution	Extent (A-F)
The conception or design of the work:	A
<i>Free text description of PhD student's contribution (mandatory)</i>	
Johannes came up with the idea, that GAM's could be used for fitting waveforms, and, together with Simon Vistisen descided that it could be presented in combination with a more general introduction to GAM's	
The acquisition, analysis, or interpretation of data:	B
<i>Free text description of PhD student's contribution (mandatory)</i>	
Johannes acquired sample data as part of the study presented in Paper 3, did the initial data analysis, and discussed interpretations and model improvements with co-authors.	
Drafting the manuscript:	B
<i>Free text description of PhD student's contribution (mandatory)</i>	
Johannes prepared the initial draft, which was then revised and supplemented by co-authors	
Submission process including revisions:	A

*D. Declarations of co-authorship*



*Free text description of PhD student's contribution (mandatory)*

Johannes submitted the manuscript and handled revisions after discussion with co-authors.

**Signatures of first- and last author, and main supervisor**

Date	Name	Signature
26-10-2022	Johannes Enevoldsen	
26-10-2022	Simon Vistisen	

Date: 26-10-2022

Signature of the PhD student

D. Declarations of co-authorship



**Declaration of co-authorship concerning article for PhD dissertations**

Full name of the PhD student: Johannes Aagaard Enevoldsen

This declaration concerns the following article/manuscript:

Title:	The Effects of Respiratory Rate and Tidal Volume on Pulse Pressure Variation in Healthy Lungs—A Generalized Additive Model Approach May Help Overcome Limitations
Authors:	JJohannes Enevoldsen Birgitte Brandsborg Peter Juhl-Olsen Stephen Edward Rees Henriette Vind Thaysen Thomas W.L. Scheeren Simon Tilma Vistisen

The article/manuscript is: Published  Accepted  Submitted  In preperation

If published, state full reference:

If accepted or submitted, state journal: British Journal of Anaesthesia

Has the article/manuscript previously been used in other PhD or doctoral dissertations?

No  Yes  If yes, give details:

**Your contribution**

Please rate (A-F) your contribution to the elements of this article/manuscript, and elaborate on your rating in the free text section below.

- A. Has essentially done all the work (>90%)
- B. Has done most of the work (67-90 %)
- C. Has contributed considerably (34-66 %)
- D. Has contributed (10-33 %)
- E. No or little contribution (<10%)
- F. N/A

Category of contribution	Extent (A-F)
The conception or design of the work:	C
<i>Free text description of PhD student's contribution (mandatory)</i>	
Johannes discussed different parts of the study design with the co-authors with relevant expertise, and formulated the final protocol with Simon Vistisen.	
The acquisition, analysis, or interpretation of data:	B
<i>Free text description of PhD student's contribution (mandatory)</i>	
Johannes included patients, recorded data, formulated initial statistical models and wrote analysis code. Interpretation and revisions of the models were discussed with co-authors.	
Drafting the manuscript:	B

*D. Declarations of co-authorship*



*Free text description of PhD student's contribution (mandatory)*

Johannes prepared the initial draft, which was then revised and supplemented by co-authors

Submission process including revisions:	A
---	---

*Free text description of PhD student's contribution (mandatory)*

Johannes prepared and submitted the manuscript.

**Signatures of first- and last author, and main supervisor**

Date	Name	Signature
26-10-2022	Johannes Enevoldsen	
26-10-2022	Simon Vistisen	

Date: 26-10-2022

  
\_\_\_\_\_  
Signature of the PhD student