The accuracy and precision of equipment for cardiopulmonary exercise testing.

Casper Beijst

November 28, 2011
Abstract

Recently, eighteen professional cyclists of the Rabobank cycling team performed a cardiopulmonary exercise test. Results were inconsistent with exercise physiology suggestive for problems with the accuracy and precision of cardiopulmonary exercise testing (CPET) equipment.

Physicians rely on accurate and precise CPET equipment for a relevant diagnosis and assessment. Therefore, this research focuses on the accuracy and precision of CPET equipment when used to monitor the exercise capacity of athletes and investigates whether the same type of errors play a role when used for patient care in clinical situations. Finding an objective way to study the accuracy and precision of CPET systems is an important part of this study.

Two methods have been used to compare breath-by-breath and mixing chamber principles. First of all, we developed a (theoretical) error analysis based on general error propagation theory. Secondly, calibration measurements using a metabolic simulator were performed.

Error analysis shows that the error in oxygen uptake ($\dot{V}_{O_2}$) and carbon dioxide production ($\dot{V}_{CO_2}$) is smaller for mixing chamber systems than for breath-by-breath systems. The relative error is largely constant over the range of $\dot{V}_{O_2}$, showing relative errors in the range of 9-12% and 5-7% for breath-by-breath and mixing chamber systems respectively. Relative errors in $\dot{V}_{CO_2}$ are found in the range of 5-7% for breath-by-breath systems and approximately 4% for mixing chamber systems. In general, the error of the flow sensor $\delta \dot{V}$ and the delay time error $\delta t_{\text{delay}}$ are significant sources of error.

Measurements using a metabolic simulator show that breath-by-breath systems are less stable for different values of minute ventilation than mixing chamber systems. Moreover, metabolic simulator measurements show that there are significant differences between the accuracy and precision of different commercially available CPET systems. Relative errors in $\dot{V}_{O_2}$ and $\dot{V}_{CO_2}$ as high as 15% are found, whereas relative errors in $\dot{V}_{E}$ are in the range of 0-10%. Corrections by the software and/or hardware of CPET systems can have significant effect on the results of the CPET systems.
# Contents

1 Introduction 1

2 Clinical background 3
   2.1 Terminology 3
   2.2 Exercise physiology 4
   2.3 In practice 6
   2.4 Accuracy 7

3 CPET methods 11
   3.1 Gas theory 11
      3.1.1 Volume correction 11
      3.1.2 Vapor pressure 12
      3.1.3 Gas fractions and humidity 12
   3.2 Methods 13
      3.2.1 Douglas bag method 13
      3.2.2 Breath-by-breath 15
      3.2.3 Mixing chamber 17

4 Problem exploration 19
   4.1 Hypotheses 19
   4.2 Flow calibration 22
   4.3 Frequency response 23
   4.4 Continuation 25

5 Publication 27

6 Additional results 41
   6.1 Metabolic simulator 41

7 General discussion 45
   7.1 Discussion 45
   7.2 Conclusion 46
Chapter 1

Introduction

Several studies have shown that exercise capacity is a strong predictor for overall mortality and that caused by cardiovascular disease, and that exercise capacity can be altered by training [1, 2]. Therefore, exercise tests of all sorts have been developed to determine a patient’s exercise tolerance. Cardiopulmonary Exercise Testing (CPET) is one of these exercise tests and involves the analysis of inspiratory and expiratory breath gasses during exercise to determine oxygen uptake ($\dot{V}_{O_2}$) and carbon dioxide production ($\dot{V}_{CO_2}$). These are valuable parameters in the assessment of pulmonary gas exchange, cardiovascular performance and skeletal muscle metabolism [3].

At the Máxima Medisch Centrum (MMC) CPET is commonly used in clinical practice thanks to an effective collaboration between sports medicine, cardiology and pulmonology specialties. Many years of experience have led to far-reaching implementation of CPET in clinical practice but have also resulted in a critical attitude toward the accuracy and precision of CPET equipment.

Recent tests using the current CPET equipment on elite athletes have shown data inconsistent with physiology, suggestive for accuracy and precision problems of CPET systems. The aim of this study is to investigate the accuracy and precision of CPET equipment, in particular, when used on athletes. We studied what are the major sources of error and whether the same types of error play a role when equipment is used on patients. Developing a quantitative method to study the accuracy and precision of CPET equipment is an important part of this study.

This report is organized as follows. Chapter 2 discusses exercise testing in clinical practice and accuracy/precision problems with CPET equipment. In Chapter 3 an overview of different CPET methods is given and it is explained how the most important CPET parameters are calculated. Chapter 4 then continues with an overview of potential sources of error. Our publication comparing the accuracy and precision of different CPET measurement principles is presented in chapter 5 and chapter 6 discusses additional results. We conclude this report with conclusions and a discussion of our results in chapter 7 and recommendations in chapter 8.
Chapter 2

Clinical background

Exercise tolerance is an important indicator for physical condition and fitness. Therefore, determining exercise tolerance has become an important tool in clinical medicine to assess a patient's physical condition. One of the major advantages of exercise testing is that measurements can be done non-invasively. The outcome of an exercise test gives valuable information about pulmonary gas exchange, cardiovascular performance and skeletal muscle metabolism.

Section 2.1 discusses terminology used in the field of CPET and section 2.2 explains the fundamentals of exercise physiology to understand the physiological response to exercise. This chapter also discusses the use of CPET in clinical practice and the problems with the accuracy and precision of the current CPET equipment leading to the aim of this study.

2.1 Terminology

Because there are different types of exercise testing, nomenclature can sometimes lead to confusion. In this section, we discuss the most important types of exercise testing and the names that will be used in this report. One of the commonly used types of exercise testing is stress testing which involves recording a patient’s electrocardiogram (ECG) during exercise. The patient is subjected to a work rate using a treadmill or bicycle ergometer. The goal of this test is to study exercise tolerance and possible abnormalities in the ECG.

Another type of exercise testing is cardiopulmonary exercise testing (CPET or CPX) or ergospirometry. This involves the analysis of breath gases during exercise to calculate the oxygen uptake $\dot{V}_O_2$ and carbon dioxide production $\dot{V}_C_O_2$ among other parameters. Usually, patients are subjected to a work rate using a treadmill or bicycle ergometer, however portable CPET systems can be used during (almost) any type of exercise. Usually, the patient’s ECG is recorded to provide additional information. Many names are used for systems measuring metabolic gas exchange: ergospirometry system, cardiopulmonary exercise testing (CPET or CPX) system, metabolic cart (MC) and metabolic
measurements cart (MMC). Ergospirometry or CPET should not be confused with spirometry which is used in pulmonary function testing [3, 4, 5]. Figure 2.1 is a picture of an CPET setup in use.

2.2 Exercise physiology

This section describes the fundamentals of physiology, to understand the physiological response of metabolic gas exchange to exercise and the problems discussed in section 2.4. First, we discuss the effect of exercise on respiration, and secondly, an explanation is given for $RQ$ values below 1, according to the overall (net) oxidation reactions of nutrients. Finally, the fundamentals of metabolism are discussed to explain the rise of $RQ$ values above 1.

During physical exercise muscles require an appropriate increase of oxygen delivery and consequently, a larger amount of carbon dioxide is produced on a cellular level. Oxygen ($O_2$) is inspired by respiration, taken up by the lungs and transported to the cells by the cardiovascular system and, vice versa, the produced carbon dioxide ($CO_2$) is transported away from the cells by the cardiovascular system and excreted by the lungs. Therefore, the exercise tolerance of a subject is dependent on three factors; pulmonary gas exchange, cardiovascular performance and skeletal muscle metabolism. This is shown in figure 2.2 illustrating the coupling of cellular and pulmonary respiration with on the left side the muscles and mitochondria where metabolism takes place, in the middle the cardiovascular system transporting $O_2$ and $CO_2$ to and from the mitochondria and on the right side the respiratory system [3].

A subject’s gas exchange can be monitored externally by measuring oxygen uptake ($V_{O_2}$) and carbon dioxide production ($V_{CO_2}$). Both are usually expressed in L/min. An important parameter characterizing metabolic exchange is the gas
2.2. EXERCISE PHYSIOLOGY

Figure 2.2: Schematic drawing of the gas transport mechanisms [3].

exchange ratio \((R)\) or respiratory exchange ratio \((RER)\)

\[
RER = \frac{\dot{V}_{CO_2}}{\dot{V}_{O_2}}.
\]  

(2.1)

The respiratory exchange ratio \((RER)\) is measured externally and only in steady-state reflects gas exchange on a cellular level. The respiratory quotient \((RQ)\) describes the ratio between cellular CO\(_2\) production \(\dot{Q}_{CO_2}\) and cellular oxygen uptake \(\dot{Q}_{O_2}\):

\[
RQ = \frac{\dot{Q}_{CO_2}}{\dot{Q}_{O_2}}.
\]  

(2.2)

\(RQ\) is directly related to the oxidation processes on a cellular level. During oxidation of carbohydrates, the amount of \(O_2\) consumed is equal to the amount of CO\(_2\) produced and corresponds to \(RQ = 1\) as shown by the overall (net) reaction of carbohydrate oxidation [6]

\[
C_6H_{12}O_6 + 6 \, O_2 \to 6 \, CO_2 + 6 \, H_2O
\]  

(2.3)

\(RQ = 6 \, CO_2/6 \, O_2 = 1\).

Accordingly, the oxidation of lipids corresponds to \(RQ < 1\). The reaction below shows the oxidation of triacylglycerol, where the amount of \(O_2\) consumed is larger than the amount of CO\(_2\) produced:

\[
2 \, C_{57}H_{110}O_6 + 163 \, O_2 \to 114 \, CO_2 + 110 \, H_2O
\]  

(2.4)

\(RQ = 114 \, CO_2/163 \, O_2 = 0.7\).
The oxidation of proteins in the body is more complicated, as shown in the oxidation reaction below. The oxidation of proteins results in $RQ < 1$:

$$C_{72}H_{112}N_{18}O_{22}S + 77 \text{O}_2 \rightarrow 63 \text{CO}_2 + 38 \text{H}_2\text{O} + \text{SO}_3 + 9 \text{CO(NH}_2\text{)}_2 \quad (2.5)$$

$$RQ = 63 \text{CO}_2/77 \text{O}_2 = 0.818.$$ 

Although the net oxidation reactions provide comprehensive explanation of $RQ$ values below 1, is does not for values higher than 1. However, fundamentals of metabolism can explain the rise of $RQ$ above 1.

Adenosine triphosphate (ATP) provides the source of energy for muscle contraction and is produced largely by oxidative phosphorylation which takes place in the cytoplasm and mitochondria. Oxidative phosphorylation starts with glycolysis metabolizing glycogen and glucose to form pyruvate. In this process, two molecules of ATP are created. Subsequently, pyruvate is oxidized to form acetyl coenzyme A (acetyl CoA), which enters the Krebs cycle in presence of oxygen. The Krebs cycle (or citric acid cycle) is a cyclical series of reactions producing $\text{CO}_2$ and ATP (among other compounds). In the absence of oxygen, pyruvate is oxidized to produce lactic acid. The rise in plasma hydrogen concentration as a consequence of lactic acid production, causes an increased amount of expired $\text{CO}_2$ [3, 4, 7, 8]

$$\text{H}^+ + \text{HCO}_3^- \rightarrow \text{H}_2\text{CO}_3 \rightarrow \text{CO}_2 + \text{H}_2\text{O}. \quad (2.6)$$

### 2.3 CPET in practice

This section shows the results of a typical exercise test and discusses the development of key parameters during the test. Measurements are performed using a ZAN 680 CPET system with Lode cycle ergometer. Figure 2.3a shows the workrate or load as a function of time. Patients are commonly subjected to a work rate using a cycle ergometer or treadmill but also hand ergometers, rowing ergometers etcetera can be used. Different protocols can be used during exercise testing which determine the work rate as a function of time [9, 10]. The protocol defines the length of the warming up phase, the increase of work rate during the exercise phase and the work rate and length of the cooling down phase. The test starts with a warming up phase to reach a steady state in respiratory gas exchange. In this case, the exercise phase is a 10 minute ramp protocol meaning that the load is increased linearly reaching the predicted maximum after 10 minutes. The workrate is increased until the patient reaches exhaustion and further exercise is limited by fatigue, pain or shortness of breath. Subsequently, the cooling down phase is started subjecting the subject to a very low workrate, approximately 10% of the maximum work load.

Figure 2.3c shows the oxygen uptake and carbon dioxide production as a function of time. During the warming up phase, steady state is reached after several minutes. The $\dot{V}_O_2$ and $\dot{V}_CO_2$ increase during the exercise phase as the work rate increases. In figure 2.3b the minute ventilation or exhaled volume
CHAPTER 2. CLINICAL BACKGROUND

2.4. ACCURACY

Figure 2.3: Results of an exercise test using a ZAN 680 CPET system with (a) the work rate as a function of time, (b) \( \dot{V}_E \) as a function of time, (c) \( \dot{V}_{O_2} \) and \( \dot{V}_{CO_2} \) as a function of time and (d) RER as a function of time.

per minute (\( \dot{V}_E \)) is plotted as a function of time. During the exercise phase, \( \dot{V}_E \) increases as workrate is increased until exhaustion is reached and starts decreasing when the cooling down phase is started.

Figure 2.3d shows the gas exchange ratio as a function of time, showing an increase of RER during the exercise phase. Although a combination of nutrients is used throughout the test, in the beginning metabolism is largely fueled by fats and amino acids. Consequently, the RER value is typically lower than 1. As the work rate increases, there is an increased demand for blood glucose and muscle glycogen. The subsequent rise in the RER value is driven by an increased demand for energy in the form of carbohydrates, but also by lactate accumulation. Due to lactate accumulation, the concentration of carbonic acid in the blood increases, which is released in the form of \( CO_2 \).

2.4 CPET accuracy

Physicians rely on correct and reliable CPET measurements for a relevant diagnosis, stressing the importance of accurate and precise measurements of metabolic gas exchange (Appendix A: Accuracy and precision). Recently, CPET
measurements have been performed on professional cyclists. Results were inconsistent with exercise physiology suggesting problems with the accuracy of CPET equipment.

Eighteen professional cyclists of the Rabobank cycling team performed an exercise test on a ZAN 680 CPET system with Lode Excalibur sport cycle ergometer. Typical results are shown in figure 2.4. Tests are done using a relatively long exercise protocol with lengths varying between approximately 20 and 30 minutes until exhaustion is reached. The work rate is not increased linearly but in steps depending on the subject’s weight (figure 2.4a). Figure 2.4c shows the $\dot{V}_O_2$ and $\dot{V}_CO_2$ as a function of time, both increasing as the work rate increases. The respiratory exchange ratio $RER$ is shown as a function of time in figure 2.4d. As shown in figure 2.3, we expect the $RER$ to be lower than 1 at the beginning of exercise and to be higher than 1 at the end of exercise due to anaerobic metabolism. Surprisingly, the measured $RER$ in figure 2.4d is not higher than 1 during any part of the test. We calculated the mean $RER$ during the highest load step and found an $RER_{max} \pm SD$ of 0.91±0.05 suggestive for accuracy and precision problems with the CPET equipment used.

To provide a reliable assessment of athletic exercise capacity, CPET systems
should measure minute ventilation $\dot{V}_E$ in the range of 0-250 L·min$^{-1}$ therefore being able to measure flows in the range of 0-15 L·s$^{-1}$. For athletes, $\dot{V}_{O_2}$ and $\dot{V}_{CO_2}$ as high as 7 L·min$^{-1}$ have been measured. Nevertheless, peak values are lower for the average untrained person, and even more so for patients depending on the severity of the disease. An overview of pathophysiological responses in common disorders has been described by Wasserman et al. [3].

Although the results described above suggest that this CPET system measures athletic exercise capacity inaccurately, the implications for patient care in clinical practice remain unclear. Other studies have shown that the biological variability is in the order of 5% [11, 12, 13, 14], and errors in CPET measurements much larger than the biological variability may hamper clinical application. Furthermore, CPET systems should be able to quantify training effects in the range of 10-15% [15, 16]. The Australian Sports Commission has written guidelines regarding the quality assurance of CPET equipment, suggesting that the precision at maximum work rate should be better than 3% for $\dot{V}_{O_2,max}$ and 5% for $\dot{V}_{E,max}$ [17].

The accuracy of CPET equipment is important but hard to determine. The general problem with validation of CPET systems is that there is no real gold standard for testing of CPET equipment. Often the Douglas bag method is used to validate automated CPET systems, and validation studies using the Douglas bag have found disagreement between the Douglas bag and CPET systems in $\dot{V}_{O_2}$ of up to 15%, and for portable systems up to 22% [18]. However, care should be taken comparing data from completely different systems since validity and reliability issues of the Douglas bag method have been observed many years ago [19]. Moreover, several studies have shown that the largest part of the total variability by calibration with human subjects is associated with the biological variability and only a small part with the variability caused by the accuracy and precision of the measurement [11, 12, 13]. Alternatively, a metabolic simulator has been described by Huszczuk et al. [20] and Gore et al. [21] as a subject independent method to validate CPET equipment.

Validation of CPET equipment can give valuable information about the accuracy and precision, however the origin and relative importance of different sources of error remain unknown. Error analysis can be used to study the absolute and relative importance of different sources of error quantitatively. Error analysis theory has been applied to CPET equipment [22, 23, 24] but a quantitative method to describe the absolute and relative importance of different sources of error has not been described before.

The aim of this study is to investigate the accuracy and precision of CPET equipment, in particular, when used on athletes. We studied what are the major sources of error and whether the same types of error play a role when equipment is used on patients. Developing a quantitative method to study the accuracy and precision of CPET equipment is an important part of this study.
Chapter 3

CPET methods

This chapter describes the most commonly used principles in CPET; the douglas bag method, the breath-by-breath method and the mixing chamber method. Section 3.1 discusses the influence of environmental conditions on volumes of gas, the correction of volumes for different conditions, the calculation of the vapour pressure and the influence of humidity on gas fractions. This theory is used in CPET measurement principles, but also in the error analysis method and metabolic simulator measurements described in chapter 5.

3.1 Gas theory

3.1.1 Volume correction

As described by the ideal gas law, the volume \( V \) of an ideal gas is dependent on the amount of molecules \( n \) in mol, the temperature \( T \) of the gas and the pressure \( p \)

\[
V = \frac{nRT}{p}. \tag{3.1}
\]

This means that the volume of a gas will change whenever the pressure or temperature is changed. When a volume of warm moist air is cooled down, the volume decreases and/or the pressure drops (equation 3.1). Additionally, water vapor in the air will condense as a consequence of the temperature drop, causing and additional pressure drop and/or volume decrease. This is described by the following relation

\[
V_2 = V_1 \cdot \frac{T_2}{T_1} \cdot \frac{p_1 - p_{H_2O,1}}{p_2 - p_{H_2O,2}} \tag{3.2}
\]

where \( V_1 \) and \( V_2 \) are volumes, \( T_1 \) and \( T_2 \) temperatures, \( p_1 \) and \( p_2 \) pressures and \( p_{H_2O,1} \) and \( p_{H_2O,2} \) the partial pressures of gaseous water in state 1 and 2, respectively.

Measures of volume can only be compared when taken under identical environmental conditions. In the field of (ergo)spirometry some environmental
3.1. **GAS THEORY**  

## 3.1.2 Vapor pressure

The partial pressure of gaseous water \( p_{H_2O} \) depends on the temperature \( T \). Values of saturated vapor pressure \( p_{H_2O,\text{sat}} \) are tabulated for the normal range of temperatures, but \( p_{H_2O,\text{sat}} \) can also be calculated in mmHg with \( T \) in °C using one of the following empirical formulae [25, 26]

\[
p_{H_2O,\text{sat}} = 47.07 \times 10^{\frac{2.36(T-37)}{273 + T}} \quad (3.3)
\]

or

\[
p_{H_2O,\text{sat}} = 0.1333(9.993 - 0.3952T + 0.03775T^2). \quad (3.4)
\]

In each of these cases, the partial pressure of water in air is determined for a relative humidity of 100%. The partial pressure of water in non-saturated air is determined using

\[
p_{H_2O,RH} = p_{H_2O,\text{sat}} \frac{R_H}{100} \quad (3.5)
\]

with \( R_H \) the relative humidity in %.

### 3.1.3 Gas fractions and humidity

Whenever gas fraction are measures in a volume of gas, the measured gas fraction depends on the relative humidity of the gas since the gas is diluted by gaseous water. This is especially important for the measurement of gas fractions by CPET systems, since expired air is highly humid. The following relation holds
for gas fractions in humid air and dry air [27, 28]

\[ F_{O_2, RH} = F_{O_2, dry} \frac{p_B - p_{H_2O, RH}}{p_B} \]  

(3.6)

where \( p_B \) is the barometric pressure, \( F_{O_2, RH} \) is the oxygen fraction in air with relative humidity \( R_H \) and \( F_{O_2, dry} \) is the oxygen fraction in dry air. A similar relation is valid for carbon dioxide fractions.

Most CPET systems lead the sampled gas through a Permapure Nafion tube to dry the sample. The sampled gas will not be dried completely, but the vapor pressure of the sample equilibrates with the vapor pressure of ambient air [29]. Since we can measure the vapor pressure of the ambient air, we know the vapor pressure of the gas sampled and dry gas fractions can be calculated.

3.2 CPET methods

Traditionally, oxygen uptake (\( \dot{V}_{O_2} \)) and the volume of carbon dioxide produced (\( \dot{V}_{CO_2} \)) were measured using the Douglas bag method, invented in 1911 [30], long before other methods were invented. This involved the use of large bags to collect the expiratory air (cf. section 3.2.1) and the results could only be calculated afterwards. As smaller, faster and cheaper gas sensors were developed, it became easier to analyse breath gasses as they are consumed or produced and two commonly used methods have been developed since. The breath-by-breath method measures inspiratory and expiratory flow together with gas fractions directly outside the facemask [31]. The mixing chamber method measures expiratory flow and collects expired air in a mixing chamber where gas fractions are measured [32, 33]. Both breath-by-breath systems and mixing chamber systems can present the data in real time.

3.2.1 Douglas bag method

The Douglas bag method involves the collection of expired air in a large bag to determine metabolic gas exchange [30]. The expired air of the subject is collected during physical exercise, and analyzed afterwards. The total volume of expired air \( V_{coll} \), the average expired oxygen fraction \( F_{EO_2} \) and the average expired carbon dioxide fraction \( F_{ECO_2} \) are measured and used to calculate the average oxygen uptake \( \dot{V}_{O_2} \) and carbon dioxide production \( \dot{V}_{CO_2} \). Because the expired air from different breaths is mixed in the collection bag, the Douglas bag method measures the average metabolic gas exchange over the collection time. Figure 3.1 shows a schematic drawing of a Douglas bag system [34].

\( \dot{V}_{O_2} = (\dot{V}_I \cdot F_{IO_2}) - (\dot{V}_E \cdot F_{EO_2}) \)  

(3.7)

where \( \dot{V}_I \) is the inspired volume per minute, \( F_{IO_2} \) is the average inspired oxygen fraction, \( \dot{V}_E \) is the expired volume per minute or minute ventilation and \( F_{EO_2} \)
is the average expired oxygen fraction. $F_{IO_2}$ is assumed to be equal to the atmospheric oxygen fraction and $F_{EO_2}$ is determined by measuring gas fractions in expired air from the collection bag. Minute ventilation $\dot{V}_E$ in L·min$^{-1}$ is calculated using the collected gas volume $V_{coll}$ and the collection time $t_{coll}$ in s

$$\dot{V}_E = V_{coll} \cdot \frac{60}{t_{coll}}. \quad (3.8)$$

The inspired volume $\dot{V}_I$ cannot be measured directly since only expired air is collected in the bag. We cannot assume the inspired volume is equal to the expired volume due to metabolic gas exchange. However, we do know that the amount of molecules nitrogen $N_2$ (and other inert gasses) does not change as air is respired. The nitrogen balance can be written in the following way

$$\dot{V}_I \cdot F_{IN_2} = \dot{V}_E \cdot F_{EN_2} \quad (3.9)$$

where $F_{IN_2}$ is the average fraction of $N_2$ (and other inert gasses) in inspired air and $F_{EN_2}$ is the average fraction of $N_2$ (and other inert gasses) in expired air. Since $F_{IN_2}$ and $F_{EN_2}$ are not measured directly, they are calculated using

$$F_{IN_2} = 1 - F_{IO_2} - F_{ICO_2} \quad (3.10)$$

and

$$F_{EN_2} = 1 - F_{EO_2} - F_{ECO_2}, \quad (3.11)$$
where $F_{ICO_2}$ and $F_{ECO_2}$ are the inspired and expired average CO\textsubscript{2} fractions. Inspired O\textsubscript{2} and CO\textsubscript{2} fractions are assumed to be equal to atmospheric O\textsubscript{2} and CO\textsubscript{2} fractions $F_{O_2,atm}$ and $F_{CO_2,atm}$, whereas expired O\textsubscript{2} and CO\textsubscript{2} fractions are measured by sampling gas from the Douglas bag. Equations 3.9, 3.10 and 3.11 are combined and rewritten to obtain an expression for the inspired volume $\dot{V}_I$

$$\dot{V}_I = \frac{1 - F_{EO_2} - F_{ECO_2}}{1 - F_{IO_2} - F_{ICO_2}} \cdot \dot{V}_E (3.12)$$

and thus $\dot{V}_{O_2}$ can be calculated from equation 3.7.

**Carbon dioxide production** $\dot{V}_{CO_2}$ is calculated in a similar manner. The expired amount of CO\textsubscript{2} can be subtracted from the inspired amount

$$\dot{V}_{CO_2} = (\dot{V}_E \cdot F_{ECO_2}) - (\dot{V}_I \cdot F_{ICO_2}) (3.13)$$

however, the inspired CO\textsubscript{2} fraction is much smaller than the expired CO\textsubscript{2} fraction ($F_{ICO_2} \ll F_{ECO_2}$). That is why the expression for $\dot{V}_{CO_2}$ is often simplified to

$$\dot{V}_{CO_2} = \dot{V}_E \cdot F_{ECO_2}. (3.14)$$

Note that measures of the collected gas volume $V_{coll}$ depend on the environmental conditions. Conventionally, $\dot{V}_{O_2}$ and $\dot{V}_{CO_2}$ are expressed under STPD conditions and equation 3.2 should be used to correct measures of the collected gas volume $V_{coll}$. Furthermore, dry gas fractions should be used to calculate $\dot{V}_{O_2}$ and $\dot{V}_{CO_2}$ under STPD conditions (equation 3.7) [27].

### 3.2.2 Breath-by-breath

Due to advances in CPET, exercise testing has become less laborious and has gained temporal resolution [37]. Unlike the Douglas bag method, the breath-by-breath method is able to calculate metabolic gas exchange during an exercise test and calculates $\dot{V}_{O_2}$, $\dot{V}_{CO_2}$, $RER$ and many other parameters real time for each breath. Breath-by-breath data is calculated by the software using raw
data from a flow sensor, gas sensors and sensors measuring ambient conditions. The flow and gas sensors typically produce raw data with a sampling frequency $f_s=25-100$ Hz. Figure 3.3 shows an example of raw data produced by the flow and gas sensors during an exercise test.

**Oxygen uptake** $\dot{V}_O_2$ is measured by subtracting the exhaled amount of oxygen from the inhaled amount of oxygen

$$\dot{V}_O_2 = (\dot{V}_I \cdot F_{IO_2}) - (\dot{V}_E \cdot F_{EO_2}). \quad (3.15)$$

The volume of expired air $V_{ex}$ is used to calculate minute ventilation $\dot{V}_E$ by integrating the flow $\dot{V}$ during expiration \cite{31, 3}

$$V_{ex} = \int_{t_{be}}^{t_{ee}} \dot{V}(t) \, dt \quad (3.16)$$

with $t_{be}$ the time at the beginning of expiration and $t_{ee}$ the time at the end of expiration. However, the raw data produced by the sensors is digital so integration is approximated by a summation to calculate $V_{ex}$

$$V_{ex} = \sum_{t_{be}}^{t_{ee}} \dot{V}(t) \cdot t_s \quad (3.17)$$

where $t_s$ is the sampling time $(1/f_s)$. Subsequently, the minute ventilation $\dot{V}_E$ is calculated

$$\dot{V}_E = \frac{60 \cdot V_{ex}}{t_{in} + t_{ex}} \quad (3.18)$$

with $t_{ex}$ the time of expiration and $t_{in}$ the time of inspiration.
CHAPTER 3. CPET METHODS

3.2. METHODS

Most breath-by-breath systems do not measure inspired volume \( \dot{V}_I \) by integrating the flow during inspiration but calculate inspired volume using nitrogen balance (equation 3.9). Wilmore et al. have shown that a true retention or production of \( \text{N}_2 \) has little effect on \( \dot{V}_O_2 \) calculation [38], justifying the use of nitrogen balance. Error analysis has shown that calculating instead of measuring inspired volume improves accuracy of respiratory gas exchange measurements for patients that are not artificially ventilated with oxygen enriched-air increasing \( F_{I O_2} \) [23].

Average expired gas fractions are calculated using

\[
F_{EO_2} = \frac{\sum_{t_{ee}}^{t_{eb}} \dot{V}(t) \cdot F_{O_2}(t) \cdot t_s}{\sum_{t_{ee}}^{t_{eb}} \dot{V}(t) \cdot t_s} = \frac{\sum_{t_{ee}}^{t_{eb}} \dot{V}(t) \cdot F_{O_2}(t) \cdot t_s}{V_{ex}}
\]

and

\[
F_{ECO_2} = \frac{\sum_{t_{ee}}^{t_{eb}} \dot{V}(t) \cdot F_{CO_2}(t) \cdot t_s}{\sum_{t_{ee}}^{t_{eb}} \dot{V}(t) \cdot t_s} = \frac{\sum_{t_{ee}}^{t_{eb}} \dot{V}(t) \cdot F_{O_2}(t) \cdot t_s}{V_{ex}}
\]

**Carbon dioxide production** \( \dot{V}_{CO_2} \) is calculated using

\[
\dot{V}_{CO_2} = \dot{V}_E \cdot F_{ECO_2}.
\]

Note that the flow sensor is usually mounted on the facemask and therefore measures warm and humid air. Volumes measured by the flow sensor should be corrected to STPD conditions using equation 3.2. Since \( \dot{V}_{O_2} \) and \( \dot{V}_{CO_2} \) are expressed at STPD, dry gas fractions should be used to calculate \( \dot{V}_{O_2} \) and \( \dot{V}_{CO_2} \) [27].

### 3.2.3 Mixing chamber

Mixing chamber systems collect expired air in a mixing chamber using a two-way-valve and flexible tubing [32, 33], as shown in figure 3.4. There is no standard size for a mixing chamber, but usually the expired air from several breaths (2-5 breaths) is collected in the mixing chamber. Cosmed systems use mixing chambers with a volume of 7L. A continuous sample is drawn from the mixing chamber and led to the oxygen and carbon dioxide analyzer therefore analyzing only mixed expired air. A flow sensor is mounted either between the facemask and the two-way-valve, between the flexible tubing and the mixing chamber or at the end of the mixing chamber. Figure 3.5 shows an example of raw data produced by the flow and gas sensors during an exercise test.

Expired volumes \( V_{ex} \) and minute ventilation \( \dot{V}_E \) are calculated in the same way as for breath-by-breath systems (equations 3.17 and 3.18). Nitrogen balance is used to calculate inspired volume per minute \( \dot{V}_I \). The mixing chamber enables direct measurement of \( F_{EO_2} \) and \( F_{ECO_2} \) by sampling gas from the mixing chamber, unlike breath-by-breath systems. Since the expired air of several breaths is collected in the mixing chamber, the temporal resolution of mixing chamber systems is lower than of breath-by-breath systems [37].
3.2. METHODS

CHAPTER 3. CPET METHODS

Figure 3.4: Schematic drawing of a mixing chamber system.

Figure 3.5: Examples of raw data from flow sensor (a) and gas sensors (b), measured with a sampling frequency $f_s=25$ Hz, measured by a Cosmed Quark CPET mixing chamber system.
Chapter 4

Problem exploration

As described in section 2.4, problems were encountered with the use of CPET equipment in clinical practice. Unfortunately, the quantification of measurement errors during exercise tests on human subjects is difficult due to biological variability [11, 12, 13] and the lack of standards. Moreover, most manufacturers do not mention the accuracy of commercially available equipment in ˙\(V_{O_2}\) and ˙\(V_{CO_2}\). This leaves two important questions to be answered: what are the potential sources of error and what is their relative importance? And secondly, can we quantify this error?

4.1 Hypotheses

CPET equipment measures the flow ˙\(V(t)\), oxygen concentration ˙\(F_{O_2}(t)\) and carbon dioxide concentration ˙\(F_{CO_2}(t)\) of expired and inspired air time dependently, with a typical sampling frequency (25-100 Hz) much higher than the breathing frequency. CPET equipment often also measures barometric pressure ˙\(p_B\), ambient temperature ˙\(T_A\) and relative humidity ˙\(R_H\). Subsequently, raw data from the sensors is used by the software to calculate the results of CPET measurements. Obviously, errors in CPET results can be caused by error in sensors for flow (a), gas fractions (b) or errors in the software (c). Other potential sources of error (d) are errors measuring ambient conditions or operation errors, such as failing to attaching the facemask tight enough allowing leakage. Figure 4.1 gives an overview of potential sources of error, that will be discussed in this section. The numbering of potential sources of error used in this section corresponds with the numbering used in figure 4.1.

In the field of (ergo)spirometry, various types of flow meters are used, such as turbine flow meters, variable orifice flow meters and pneumotachometers of the Fleisch, Lilly or Silverman-type [18]. In the equipment we studied, both variable orifice flow meters and turbine flow meters are used. Variable orifice flow meters [39] are difficult to calibrate and non-linearity (a1) and non-reproducibility (a2) are potential issues as well as errors caused by the frequency response (a3).
These issues are discussed in sections 4.2 and 4.3. For the accuracy of the turbine flow sensors, two phenomena are important. Leg-before-start and spin-after-stop effects are caused by angular momentum of the vane (a4) and most manufacturers correct the raw data from flow sensors for these effects [41]. The errors caused by the flow sensor are included in our error analysis method described in chapter 5.

Issues with the measurement of fast changing gas fractions are largely caused by the response time of gas analyzers (b1). Therefore, measurements of gas fractions are often corrected for the response time of the analyzer by advanced algorithms [42, 43, 44, 45, 46, 47, 48, 49]. Moreover, correct gas measurements also take into account the influence of humidity (b2) [27, 28]. Validity of the gas measurement depends on the correction of data from the gas sensors. An error often overlooked is the delay time error (b3). Delay time is defined as the
travel time of the sampling gas through the sampling tube, sampled from the flow sensor and led to the gas analyzers [50, 51]. Most systems measure the delay time of the sampling tube prior to measurement during the calibration of the gas sensors, and some systems measure the delay time during the test. Errors in the delay time are caused either by uncertainty in the measurement of the delay time or by the change of delay time during an exercise test. Errors in the measurement of gas fractions (b1-3) influence the accuracy and precision of CPET measurements and are studied in more detail in chapter 5.

The software uses the raw data from the sensors to calculate the results of the exercise test and applies the corrections described above (c1-2). Separate calibration of sensors can give valuable information about the validity of measurement, however the validity also depends on the correction applied by the software. For example, the accuracy and precision of the flow measurement also depends on the validity of the correction for the angular momentum of the vane applied by the software. Nevertheless, errors caused by software are hard to investigate. Most techniques are not patented and manufacturers are determined to keep specific details about the software secret. Therefore, it is difficult to determine whether other corrections are applied (c3) to meet the expectation of the customer.

The software also corrects for the delay time of the gas analysis (c4). The error analysis method described in chapter 5 investigates the relative and absolute importance of errors caused by the delay time error.

According to the former manufacturer (personal communication Mr. Alfred Albert, Medical Equipment Europe GmbH, Germany) of the ZAN680 CPET system, asynchrony of datastreams was sometimes encountered (c5), in such a way that a misalignment of signals from the gas sensors and flow sensors occurred. This problem was believed to be caused by miscommunication between the software and hardware components. We chose not to investigate this problem further for three reasons. First of all, this was believed to be a small problem, and secondly, this was a problem specific for CPET systems from one manufacturer only. Thirdly, a better understanding of this problem would not lead to a better understanding of the accuracy and precision of CPET systems in general.

The temperature of the exhaled air $T_B$ is important for the measurement of exhaled volume. Some CPET systems assume the temperature of exhaled air to be 34°C, others assume it to be 37°C. Nevertheless, the temperature of exhaled air depends on many factors such as temperature of ambient air, metabolic rate, exercise phase, ventilation etcetera. Significant variations in the temperature of exhaled air have been reported many years ago [52, 53, 54, 55]. Errors are introduced when the temperature of the exhaled air is assumed to be constant (d1), as shown in chapter 5.

The outcome of an exercise test depends on the ambient conditions, and errors in the ambient conditions cause errors in CPET results (d2). Barometric
pressure $p_B$ is measured to convert volumes measured in BTPS\textsuperscript{1} to STPD\textsuperscript{2} and the error in $p_B$ depends on the accuracy of the pressure sensor. Inspiratory oxygen and carbon dioxide fractions ($F_{IO_2}$ and $F_{ICO_2}$) are often assumed to be equal to atmospheric fractions. However, laboratory conditions can vary due to oxygen consumption and carbon dioxide accumulation \cite{56}. Error analysis (chapter 5) is used to study the influence of these errors.

Other sources of error are leakage of the two-way-valve (d3), leakage of the face mask (d4) and saliva blocking the sampling tube or adhering to the flow sensor (d5). Errors like these can often be prevented by cleaning the equipment prior to measurements and by strictly following the (international) guidelines for exercise testing \cite{57, 58}.

### 4.2 Flowmeter calibration

This section describes stationary calibration measurements of a variable orifice flow meter as used in the ZAN 680 CPET system. A variable orifice flow meter measures the pressure drop $\Delta p$ over a flexible plastic flap as a measure of flow \cite{39}. The mechanical properties and geometry of the plastic flap largely determine the characteristics of the flow meter. A schematic drawing of the variable orifice flow meter is shown in figure 4.2.

Figure 4.3 shows a schematic drawing of the calibration setup used for calibration measurements. A vacuum cleaner (AEG ergoessence 4599, 2000W) is used to create a stationary flow through the flow sensor. A calibrated pressure sensor (Datametrics Dresser, 1400 electronic manometer) is used to measure the corresponding pressure drop $\Delta p$ over the plastic flap and calibrated flow meters are placed in series with the variable orifice flow meter to measure the flow. Three calibrated parallel flow meters are used to measure the flow; two rotameters and one mass flow meter (FMA 1700/1800, Omega engineering 0-500L/min). The flow through the system is varied by partially opening or closing

\textsuperscript{1}Body Temperature and Pressure, Saturated
\textsuperscript{2}Standard Temperature and Pressure, Dry
4.3. FREQUENCY RESPONSE

Results of the calibration measurements are shown in figure 4.4 with the pressure drop $\Delta p$ over the flexible plastic flap as a function of flow. Calibration measurements show a non-linear relation between flow and pressure difference. Because the relation between the pressure difference and flow is non-linear, CPET systems use lookup tables to relate the measured pressure difference to flow. We do believe variable orifice flow meters are prone to error due to the change of mechanical properties of the plastic flap over time. Moreover, the mechanical properties of the plastic flap vary for each flow meter ideally requiring a separate calibration for each flow meter.

Unfortunately, the setup does not allow calibration of the sensor in the entire range of flows up to 15 L·s$^{-1}$ potentially produced by elite athletes. Furthermore, this setup does not allow investigation of time dependent effects. These measurements show that it is possible to use variable orifice flow meters in CPET equipment, but correct calibration is crucial.

4.3 Frequency response

The following measurements are performed using a Vacumed metabolic simulator (Vacumed, USA), which is essentially a piston pump creating sinusoidal
flows with fixed tidal volumes $V_T$. The metabolic simulator is a device that can be connected to a CPET system to perform calibration measurements and is able to simulate respiratory function. Titration of calibration gas (21% carbon dioxide, 79% nitrogen) dilutes the room air inside piston pump simulating oxygen consumption $\dot{V}_{O_2}$ and addition of extra carbon dioxide simulates carbon dioxide production $\dot{V}_{CO_2}$ [20, 21]. In the next chapter, the metabolic simulator is discussed in more detail.

For this experiment, we do not use any calibration gas and only use the metabolic simulator as a motorized calibration syringe, creating sinusoidal flows with fixed tidal volumes $V_T$. The experiment studies the effect of changing breathing frequencies of the metabolic simulator on the inspiratory and expiratory volumes measured by the CPET system, in this case a ZAN 680 (nSpire Health, Germany). We measured the inspired volume $V_{in}$ and expired volume $V_{ex}$ as a function of the breathing frequency $f_B$ and repeated the experiment for different tidal volumes of the metabolic simulator ($V_T=1, 2, 3$ and 4L), as shown in figure 4.5.

In general, the measured inspired and expired volumes $V_{in}$ and $V_{ex}$ first increase, and then decrease, as the breathing frequency is increased. Errors in the inspired and expired volumes $V_{in}$ and $V_{ex}$ of up to 8% are found. During the experiment, the metabolic simulator creates a sinusoidal flow pattern. Since the frequency of the piston pump is changed as the tidal volume is kept constant, both the frequency as the amplitude of the sinusoidal flow pattern change. Therefore, changes in the measured inspired and expired volumes $V_{in}$ and $V_{ex}$ are not only caused by the changing frequency, but also by the changing amplitude. Errors in the inspired and expired volumes $V_{in}$ and $V_{ex}$ can be caused by miscalibration of the flow sensor and by resonance of the variable orifice. These
4.4 Continuation of the study

The calibration of individual sensors gives valuable information about the error of measurements, but only partially helps in understanding the accuracy and precision of CPET systems as a whole. To understand the origin of accuracy problems of CPET systems, we need to know the relative importance of different sources of error, preferably in a quantitative manner.

The remainder of this thesis describes two methods that are used to study the error of CPET equipment quantitatively, namely, an error analysis method based on general error propagation theory and measurements using a metabolic simulator. Chapter 5 is a publication about the comparison of breath-by-breath and mixing chamber systems using these two methods. In chapter 6, the results of metabolic simulator measurements on several commercially available CPET systems are shown.
Chapter 5

Publication

Accuracy and precision of CPET equipment: a comparison of breath-by-breath and mixing chamber systems

C. Beijst, G. Schep, E. van Breda, P. Wijn and C. van Pul

Abstract

Cardiopulmonary exercise testing (CPET) has become an important diagnostic tool for patients with cardiorespiratory disease and can monitor athletic performance measuring maximal oxygen uptake $\dot{V}_{O_2,\text{max}}$. The accuracy and precision of CPET equipment are important but hard to determine. The aim of this study is to compare the accuracy and precision of a breath-by-breath and a mixing chamber system. Two methods have been used to compare breath-by-breath and mixing chamber principles. First of all, we developed a (theoretical) error analysis based on general error propagation theory. Secondly, calibration measurements using a metabolic simulator were performed. This study shows that mixing chamber systems have better accuracy and precision than breath-by-breath systems, based on theoretical error analysis based on general error propagation theory. Measurements using a metabolic simulator show that breath-by-breath systems are less stable for different values of minute ventilation than mixing chamber systems. Generally, the flow error $\delta \dot{V}$, the delay time error $\delta t_{\text{delay}}$ and the error in temperature of expired air $\delta T_B$ are significant sources of error for the computation of $\dot{V}_{O_2}$ and $\dot{V}_{CO_2}$. Error analysis also shows that the error in $\dot{V}_{O_2}$ is larger than in $\dot{V}_{CO_2}$. 

CHAPTER 5. PUBLICATION

Introduction

For many years exercise testing has been used in clinical practice to determine a patient’s exercise tolerance. Cardiopulmonary exercise testing (CPET) is the preferred method for exercise testing and involves the analysis of inspiratory and expiratory breath gases during exercise to determine oxygen uptake ($\dot{V}_O_2$) and carbon dioxide production ($\dot{V}_C_O_2$) [57, 58]. CPET has become an important diagnostic tool for patients with cardiorespiratory disease and can monitor athlete performance measuring maximal oxygen uptake $\dot{V}_O_2_{max}$ [3]. Commonly, automated systems are used to perform breath-by-breath analysis by sampling directly at the facemask in a way first described by Beaver et al. [31]. As an alternative, mixing chamber systems can be used to measure gas fractions from several breaths collected in a mixing chamber [32, 33]. Differences between breath-by-breath and mixing chamber systems have been studied before [59] and a lower precision in $\dot{V}_O_2$ and $\dot{V}_C_O_2$ in the breath-by-breath mode compared to mixing chamber mode has been found. On the other hand, mixing chamber systems measure metabolic gas exchange with a lower temporal resolution [37].

The accuracy and precision of CPET equipment are important but hard to determine. Validation studies are commonly used to determine accuracy and precision, however, the general problem with validation of CPET systems is that there is no real gold standard for testing of CPET equipment. Often the Douglas bag method is used to validate automated CPET systems. Nevertheless, care should be taken when comparing data from completely different systems since validity and reliability issues of the Douglas bag method have been observed many years ago [19]. Alternatively, a metabolic simulator has been described by Huszczuk et al. [20] and Gore et al. [21] as a subject independent method to determine the accuracy and precision of CPET equipment.

Validation of CPET equipment can give valuable information about the accuracy and precision, however the origin and relative importance of different sources of error remain unknown. This paper introduces an error analysis method to study the absolute and relative importance of different sources of error quantitatively. Error analysis theory has been applied to CPET equipment [22, 23, 24] but a quantitative method to describe the absolute and relative importance of different sources of error has not been described before.

The aim of this study is to compare the accuracy and precision of a breath-by-breath and mixing chamber system. Two methods have been used to compare breath-by-breath and mixing chamber principles. First of all, we developed a (theoretical) error analysis based on general error propagation theory. Secondly, calibration measurements using a metabolic simulator are performed.
Methods

Equipment

This paper describes a comparative study of a breath-by-breath system and a mixing chamber system. CPET data is gathered using the Cosmed Quark CPET system (Cosmed, Italy) in breath-by-breath and mixing chamber mode. The breath-by-breath method measures flow, oxygen concentration and carbon dioxide concentration time dependently with a typical frequency (25Hz) much higher than the maximum breathing frequency (∼1Hz). The mixing chamber and breath-by-breath system measure minute ventilation and expired volumes identically. However, mixing chamber systems measure $F_{EO2}$ and $F_{ECO2}$ directly in the mixing chamber reflecting average expired gas fractions. Breath-by-breath systems integrate the product of flow and gas concentration to obtain expired gas volumes

$$F_{EO2} = \frac{\sum_{t_{be}}^{t_{ee}} \dot{V}(t) \cdot F_{O2}(t) \cdot t_s}{\sum_{t_{be}}^{t_{ee}} \dot{V}(t) \cdot t_s} \quad (5.1)$$

where $t_{be}$ is the time at the beginning of expiration, $t_{ee}$ is the time at the end of expiration and $t_s$ is the sampling time. $F_{ECO2}$ is obtained is a similar way.

Generally, $\dot{V}_{O2}$ is calculated using

$$\dot{V}_{O2} = \left(\dot{V}_I \cdot F_{IO2} - \dot{V}_E \cdot F_{EO2}\right) \quad (5.2)$$

where $F_{IO2}$ is the oxygen fraction of inspired air and $\dot{V}_I$ is the inspired volume, calculated using the Haldane transformation [38]

$$\dot{V}_I = \frac{1 - F_{EO2} - F_{ECO2}}{1 - F_{IO2} - F_{ICO2}} \cdot \dot{V}_E \quad (5.3)$$

and $\dot{V}_{CO2}$ is calculated using

$$\dot{V}_{CO2} = \dot{V}_E \cdot F_{ECO2}. \quad (5.4)$$

Error analysis

Error analysis is used to estimate the error in $\dot{V}_{O2}$, $\dot{V}_{CO2}$ and $\dot{V}_E$. General error propagation theory is applied to CPET raw data using the Matlab software package (Mathworks, USA).

Error propagation

Applying general error propagation theory [60] (Appendix B: Error propagation), the error $\delta y$ in $y$ can be estimated, assuming variables $x_1, ..., x_N$ are measured with uncertainties $\delta x_1, ..., \delta x_N$ as follows:

$$\delta y \approx \sum_{n=1}^{N} \left| \frac{\partial f(x_1, ..., x_N)}{\partial x_n} \right| \delta x_n. \quad (5.5)$$
Table 5.1: Overview of errors used in the error analysis.

<table>
<thead>
<tr>
<th>Error</th>
<th>Value</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\frac{\Delta V}{V}$</td>
<td>2%</td>
<td>[61]</td>
</tr>
<tr>
<td>$\delta F_{O_2}$</td>
<td>0.03%</td>
<td>[61]</td>
</tr>
<tr>
<td>$\delta F_{CO_2}$</td>
<td>0.03%</td>
<td>[61]</td>
</tr>
<tr>
<td>$\delta T_B$</td>
<td>2°C</td>
<td>[52, 53, 54, 55]</td>
</tr>
<tr>
<td>$\delta p_B$</td>
<td>1 mmHg</td>
<td>[61]</td>
</tr>
<tr>
<td>$\delta F_{I_{O_2}}$</td>
<td>0.04%</td>
<td>[56, 62]</td>
</tr>
<tr>
<td>$\delta F_{I_{CO_2}}$</td>
<td>0.03%</td>
<td>[56, 62]</td>
</tr>
<tr>
<td>$\delta t_{delay}$</td>
<td>40 ms</td>
<td>[63, 37]</td>
</tr>
</tbody>
</table>

Similarly, the error $\delta \dot{V}_{O_2}$ in $\dot{V}_{O_2}$ is calculated, assuming variables $\dot{V}$, $F_{O_2}$, $F_{CO_2}$, $T_B$, $p_B$, $F_{I_{O_2}}$, $F_{I_{CO_2}}$ and $t_{delay}$ are measured with uncertainties $\delta \dot{V}$, $\delta F_{O_2}$, $\delta F_{CO_2}$, $\delta T_B$, $\delta p_B$, $\delta F_{I_{O_2}}$, $\delta F_{I_{CO_2}}$ and $\delta t_{delay}$ as follows

$$
\delta \dot{V}_{O_2} \approx \sum_{t=t_{on}}^{t_{off}} \frac{\partial \dot{V}_{O_2}}{\partial \dot{V}(t)} \delta \dot{V}(t) + \frac{\partial \dot{V}_{O_2}}{\partial F_{O_2}} \delta F_{O_2} \\
+ \frac{\partial \dot{V}_{O_2}}{\partial F_{CO_2}} \delta F_{CO_2} + \frac{\partial \dot{V}_{O_2}}{\partial T_B} \delta T_B \\
+ \frac{\partial \dot{V}_{O_2}}{\partial p_B} \delta p_B + \frac{\partial \dot{V}_{O_2}}{\partial F_{I_{O_2}}} \delta F_{I_{O_2}} \\
+ \frac{\partial \dot{V}_{O_2}}{\partial F_{I_{CO_2}}} \delta F_{I_{CO_2}} + \delta \dot{V}_{O_2, delay}.
$$

(5.6)

where $T_B$ is the temperature of the expired air, $p_B$ the barometric pressure and $t_{delay}$ the delay time. The errors $\delta \dot{V}_{CO_2}$ and $\delta \dot{V}_E$ are calculated in the same manner, though less variables are used to calculate $\delta \dot{V}_{CO_2}$ and $\delta \dot{V}_E$.

The equations above show that the error caused by the delay time error $\delta t_{delay}$ is not determined by calculating the partial derivative according to general error propagation theory. Instead, CPET results are re-calculated for the correct delay time, the delay time minus 40 ms and the delay time plus 40 ms [37]. For each breath, the influence of changing delay times is studied by calculating $\dot{V}_{O_2, t_{delay}} - \dot{V}_{O_2, (t_{delay}+40\text{ ms})}$ and $\dot{V}_{O_2, t_{delay}} - \dot{V}_{O_2, (t_{delay}-40\text{ ms})}$. The error in $\dot{V}_{O_2}$ caused by the delay time error $\delta \dot{V}_{O_2, delay}$ is defined as the maximum value of $|\dot{V}_{O_2, t_{delay}} - \dot{V}_{O_2, (t_{delay}+40\text{ ms})}|$ and $|\dot{V}_{O_2, t_{delay}} - \dot{V}_{O_2, (t_{delay}-40\text{ ms})}|$. The error $\delta \dot{V}_{CO_2, delay}$ is calculated in the same way.
Measurement errors

Flow is measured time dependently by a turbine flow sensor in both Cosmed systems. The uncertainty in the flow measurement used for error analysis was chosen according to the accuracy of the sensor as specified by the manufacturer [61]. An overview of errors used for error analysis is given in table 5.1.

Gas concentrations of respired air are measured time dependently by gas analyzers. Both Cosmed systems tested in this paper use paramagnetic oxygen sensors and measure carbon dioxide fractions using infrared sensors. The uncertainty of the gas analyzers used for the error analysis described in this paper assumes an accuracy of the sensor as specified by the manufacturer [61].

A quantity generally not measured is the temperature of the expired air $T_B$. Cosmed systems assume the temperature of expired air to be 34°C. Nevertheless, the temperature of expired air depends on many factors such as temperature of ambient air, metabolic rate, exercise phase, ventilation etcetera. Significant variations in the temperature of expired air have been reported many years ago [52, 53, 54, 55], therefore an uncertainty in $T_B$ of 2°C is used.

Barometric pressure $p_B$ is measured to convert volumes measured in BTPS\textsuperscript{1} to STPD\textsuperscript{2}. The uncertainty of $p_B$ is dependent on the accuracy of the pressure sensor, which is specified by the manufacturer [61].

Inspiratory oxygen and carbon dioxide fractions ($F_{IO_2}$ and $F_{ICO_2}$) are often assumed to be equal to atmospheric fractions. However laboratory conditions can vary due to oxygen consumption and carbon dioxide accumulation, depending on the amount of ventilation as shown for 38 exercise tests in a laboratory with one exercising subject and two experimenters with $F_{IO_2}$ and $F_{ICO_2}$ varying 0.04% and 0.03% respectively [56].

An error often overlooked is the delay error caused by suction of sample gas through sampling tube. Delay time is defined by the travel time of the sampling gas through the sampling tube, sampled from the flow sensor and led to the gas analyzers [50, 51]. Errors of delay time influence the accuracy of CPET measurements and are caused either by uncertainty in the measurement of the delay time or the change of delay time during testing. Therefore, the error analysis method described in this paper included errors in delay time and assumed a maximum error in delay time of 40 ms [63, 37].

Error sensitivity

Error sensitivity analysis is done to investigate the dependence of the errors $\delta \dot{V}_{O_2}$ and $\delta \dot{V}_{CO_2}$ on the input errors, since $\delta \dot{V}_{O_2}$ and $\delta \dot{V}_{CO_2}$ strongly depend on the input errors. Input errors are multiplied individually by factors 0.2, 0.5, 1, 2, 5 and the percentage of increase or decrease in the mean error $\delta \dot{V}_{O_2}$ and $\delta \dot{V}_{CO_2}$ is calculated.

\textsuperscript{1}Body Temperature and Pressure, Saturated
\textsuperscript{2}Standard Temperature and Pressure, Dry
Since there is no real gold standard for CPET, calibration with human subjects comparing the outcome of different metabolic measurement systems has become very common. However several studies have shown that the largest part of the total variability is associated with the biological variability and only a small part with the variability caused by the accuracy and precision of the measurement [11, 12, 13]. A method avoiding calibration with human subjects has been proposed by Huszczuk et al. [20] and Gore et al. [21], involving calibration with a metabolic simulator. Calibration measurements in our study are performed using a commercially available metabolic simulator (Vacumed, USA).

The metabolic simulator creates gas flows of known composition and volume, simulating carbon dioxide production and oxygen consumption. A motorized calibration syringe can be set to deliver several tidal volumes $V_T$ and breathing frequencies $f_B$, creating the desired minute ventilation $V_E$. Titration of calibration gas (21% carbon dioxide, 79% nitrogen) dilutes the room air inside piston pump simulating oxygen consumption $\dot{V}_{O_2}$ and addition of extra carbon dioxide simulates carbon dioxide production $\dot{V}_{CO_2}$. Therefore $\dot{V}_{O_2}$ and $\dot{V}_{CO_2}$ are proportional to the gas flow of calibration gas from the tank $\dot{V}_{tank}$ [64]. To compare the $\dot{V}_{O_2}$ and $\dot{V}_{CO_2}$ of the metabolic simulator to the measured CPET value, CPET values are corrected. The correction is applied because CPET systems assume expired air is close to 37°C and saturated, but actually, air expired by the metabolic simulator is at ambient temperature and humidity.

The breathing frequency and tidal volume can be configured independently from the gas flow of calibration gas allowing $\dot{V}_E$ to be set independently from $\dot{V}_{O_2}$ and $\dot{V}_{CO_2}$. This enables the investigator to study the stability of $\dot{V}_{O_2}$ and $\dot{V}_{CO_2}$ measured by the CPET system for varying $\dot{V}_E$. Table 5.2 shows the protocol used for calibration using the metabolic simulator. CPET systems are calibrated for 4 different values of $\dot{V}_{O_2}$ and $\dot{V}_{CO_2}$. The stability of all 4 metabolic rates is tested with 4 different values of minute ventilation.

### Table 5.2: Protocol during measurements with metabolic simulator.

<table>
<thead>
<tr>
<th>$\dot{V}_{O_2}$</th>
<th>$V_T$</th>
<th>$BF$</th>
<th>$V_E$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>2</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>15</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>3</td>
<td>30</td>
</tr>
<tr>
<td>3.5</td>
<td>2</td>
<td>3</td>
<td>40</td>
</tr>
</tbody>
</table>

Metabolic simulator
CHAPTER 5. PUBLICATION

Figure 5.1: Bar graph showing the mean theoretical error in $\dot{V}_{O_2}$ and $\dot{V}_{CO_2}$ during an exercise test in the ranges 0-1, 1-2, 2-3 and 3-4 L·min$^{-1}$, calculated for breath-by-breath systems (left) and mixing chamber systems (right).

Results

Error analysis

The results of the theoretical error analysis on $\dot{V}_{O_2}$ and $\dot{V}_{CO_2}$ for breath-by-breath and mixing chamber systems are shown in figure 5.1 in the ranges 0-1, 1-2, 2-3 and 3-4 L·min$^{-1}$. The different colors indicate the different input errors described in table 5.1. Figure 5.2 shows the relative error in $\dot{V}_{O_2}$ and $\dot{V}_{CO_2}$ for breath-by-breath and mixing chamber systems in the ranges 0-1, 1-2, 2-3 and 3-4 L·min$^{-1}$.

It is directly visible that for breath-by-breath systems the flow error $\delta \dot{V}$ and delay time error $\delta t_{\text{delay}}$ are the main contributors, whereas for mixing chamber systems the delay time error $\delta t_{\text{delay}}$ is not present. Furthermore, the assumed temperature of the expired air $T_B$ is important for the accuracy and precision of CPET results. The uncertainty in the oxygen fraction of inspired air $\delta F_{IO_2}$ is important for the error in $\dot{V}_{O_2}$ for both breath-by-breath and mixing chamber systems. The relative error is largely constant over the range of $\dot{V}_{O_2}$, show-
Figure 5.2: Bar graph showing the mean relative theoretical error in $\dot{V}_{O_2}$ and $\dot{V}_{CO_2}$ during an exercise test in the ranges 0-1, 1-2, 2-3 and 3-4 L·min$^{-1}$, calculated for breath-by-breath systems (left) and mixing chamber systems (right).

The theoretical error of $\dot{V}_{CO_2}$ is smaller than of $\dot{V}_{O_2}$ for both mixing chamber and breath-by-breath systems. For mixing chamber systems, the difference between $\delta\dot{V}_{O_2}$ and $\delta\dot{V}_{CO_2}$ is largely caused by $\delta F_{IO_2}$ and $\delta F_{ICO_2}$. Overall, our results show that the theoretical error in $\dot{V}_{O_2}$ and $\dot{V}_{CO_2}$ is much smaller for mixing chamber systems than for breath-by-breath systems. Differences between the theoretical error of breath-by-breath and mixing chamber systems are largely caused by the delay time error $\delta t_{delay}$ in breath-by-breath systems, but also due to differences in the flow error $\delta\dot{V}$. Surprisingly, the error caused by $\delta t_{delay}$ varies with the range of $\delta\dot{V}_{O_2}$ and $\delta\dot{V}_{CO_2}$. Relative errors in $\dot{V}_{CO_2}$ are found in the range of 5-7% for breath-by-breath systems and 4% for mixing chamber systems approximately, meaning that the absolute error increases almost linearly.

ing relative errors in the range of 9-12% and 5-7% for breath-by-breath and mixing chamber systems respectively. Relative errors in $\dot{V}_{CO_2}$ are found in the range of 5-7% for breath-by-breath systems and 4% for mixing chamber systems approximately, meaning that the absolute error increases almost linearly.
CHAPTER 5. PUBLICATION

Figure 5.3: The results of error sensitivity calculations showing the percentage of increase or decrease in mean $\delta \dot{V}_O_2$ and $\delta \dot{V}_C_O_2$ for errors multiplied by a factor of 0.2, 0.5, 1, 2 and 5 for breath-by-breath systems (left) and mixing chamber systems (right).

Error sensitivity

The results of error sensitivity calculations are shown in figure 5.3. It is directly clear in figure 5.3 that variations in flow error $\delta \dot{V}$ and temperature of expired air $T_B$ have significant effects on outcome of $\delta \dot{V}_O_2$ and $\delta \dot{V}_C_O_2$. The error in oxygen fraction $\delta F_{O_2}$ and inspired oxygen fraction $\delta F_{I,O_2}$ are important for the computation of $\delta \dot{V}_O_2$. For $\delta \dot{V}_C_O_2$, the error in carbon dioxide fraction $\delta F_{C_O_2}$ is of importance.

Metabolic simulator

Bland Altman plots are used to compare the results of CPET systems to the configuration of the metabolic simulator (figure 5.4). Figure 5.4a shows the results for $\dot{V}_O_2$. The difference between measurements of the mixing chamber system and the metabolic simulator is almost zero whereas a clear difference is visible for the breath-by-breath system. Moreover, the measured $\dot{V}_O_2$ of the mixing chamber system shows less dependence on $\dot{V}_E$ shown by a smaller spread of data. Theoretical uncertainties from error analysis shown by the lines are larger than the measured difference between metabolic simulator and CPET.
Figure 5.4: Bland Altman plots showing the results of \( V_{O_2} \), \( V_{CO_2} \) and \( V_E \) measurements using a metabolic simulator [65]. The vertical axis shows the error or difference between the measured value by the CPET system (CPET) and the input of the metabolic simulator (MS) and the horizontal axis is the input of the metabolic simulator (MS). Open dots represent data from the breath-by-breath system, filled dots represent data from the mixing chamber system, acquired at different values of minute ventilation \( V_E \). Lines represent the mean uncertainties calculated with the error analysis for breath-by-breath (dashed) and mixing chamber (solid) systems.
system.

The results of $\dot{V}_{CO_2}$ calibration are shown in figure 5.4b. The $\dot{V}_{CO_2}$ measured by the mixing chamber system shows less dependence on $\dot{V}_E$ than the breath-by-breath system. Surprisingly, the largest error $\dot{V}_{CO_2}$ measured by calibration with the metabolic simulator is larger than the theoretical uncertainty calculated.

Figure 5.4c shows the results of $\dot{V}_E$ calibration. Despite the use of identical flow sensors for mixing chamber and breath-by-breath systems, a difference between the error in $\dot{V}_E$ is observed. The $\dot{V}_E$ measured by the mixing chamber was more accurate.

The results of calibration measurements with the metabolic simulator are shown for Cosmed breath-by-breath and mixing chamber systems only. Similar measurements on equipment of other manufacturers have shown that the error observed for Cosmed systems was comparable to or smaller than the error of other commercially available systems (Beijst et al., unpublished data).

Discussion

The results of our theoretical error analysis show that the error of breath-by-breath systems is smaller than of mixing chamber systems. Delay time is the major cause of differences between breath-by-breath and mixing chamber systems. Other differences are caused by the fact that the breath-by-breath and mixing chamber methods compute the average expired oxygen and carbon dioxide fractions $F_{EO_2}$ and $F_{ECO_2}$ in a different way. Breath-by-breath systems calculate the expired gas volumes by integrating the product of flow and gas fractions, whereas mixing chamber systems measure $F_{EO_2}$ and $F_{ECO_2}$ directly in the mixing chamber [31, 32, 33]. Theoretical error analysis also shows that the error in $\dot{V}_{O_2}$ is larger than in $\dot{V}_{CO_2}$, which is largely caused by the fact that the Haldane transformation is needed to calculate inspired volumes for $\dot{V}_{O_2}$. Generally, the flow error $\delta \dot{V}$, the delay time error $\delta t_{delay}$ and the error in temperature of expired air $\delta T_B$ are significant sources of error for the computation of $\dot{V}_{O_2}$ and $\dot{V}_{CO_2}$.

The validity of our theoretical error analysis method depends on the validity of input errors. Raw data from the gas sensors is corrected for the delay and response times of the analyzers [42, 43, 44, 45, 46, 47, 48, 49] and the accuracy of the measured gas fractions depends on the validity of the correction applied to the raw data. The same is true for corrections of leg-before-start and spin-after-stop effects of the flow sensor.

To investigate the influence of changing input errors on $\delta \dot{V}_{O_2}$ and $\delta \dot{V}_{CO_2}$, we performed error sensitivity analysis. The input errors are multiplied individually by factors 0.2, 0.5, 1, 2 and 5, although it is unlikely that input errors are in fact a factor 5 larger or smaller, as long as CPET systems are calibrated prior to measurements and are validated regularly. Relatively, changes in input errors have a larger effect on the error of mixing chamber systems, since the delay time error is not present for mixing chamber systems and therefore, absolute total errors are smaller.
A limitation of the theoretical analysis method is that we did not include
the influence of breath detection algorithms and the use of two-way valves.
Furthermore, this study does not include an analysis of temporal resolution
of breath-by-breath and mixing chamber systems. Although mixing chamber sys-
tems might be more accurate in some situations, only breath-by-breath systems
are capable of measuring metabolic gas exchange with high temporal resolution
[37].

Measurements using a metabolic simulator show that for both $\dot{V}_O_2$ and $\dot{V}_C_O_2$, 
breath-by-breath data are less stable for different values of minute ventilation $V_E$ than mixing chamber data. Larger fluctuations in the measurements of breath-by-breath systems compared to mixing chamber systems have been ob-
served before [59, 66].

Measurements of $\dot{V}_E$ show that a difference between the error in $\dot{V}_E$ for
mixing chamber and breath-by-breath systems is observed, despite the use of
identical flow sensors. Two phenomena are important for the accuracy of the
turbine flow sensor. Leg-before-start and spin-after-stop effects are caused by
angular momentum of the vane and most manufacturers correct the raw data
from flow sensors for these effects [41]. Differences in the $\dot{V}_E$ error of mixing
chamber and breath-by-breath systems can be caused by the use of different
corrections.

Surprisingly, the largest error in $\dot{V}_C_O_2$ measured by calibration with the
metabolic simulator is larger than the theoretical uncertainty calculated. The
same is true for the largest error measured in $\dot{V}_E$. The repeatability of the mea-
surements has not been investigated and measurements of the breath-by-breath
and mixing chamber systems have been performed on two consecutive days,
although ambient conditions have been measured and taken into account. Dis-
agreement between the error analysis and the metabolic simulator can be caused
by limitations of the metabolic simulator. The validity of measurements with
the metabolic simulator relies on correct calibration of the metabolic simulator.
Since $\dot{V}_O_2$ and $\dot{V}_C_O_2$ are proportional to the gas flow from the calibration gas
tank measured by a mass flow meter, correct calibration of the mass flow meter
is crucial. Errors in the mass flow meter influence $\dot{V}_O_2$ and $\dot{V}_C_O_2$ proportion-
ally. Miscalibration of the mass flow meter does not influence $\dot{V}_E$, since $\dot{V}_E$ is
independent of the flow of calibration gas, but dependent on the frequency and
volume of the motorized calibration syringe or piston pump. Ideally, the mass
flow meter is recalibrated regularly but we expect the mass flow meter to be
stable since it is used to measure dry and clean gas only. The last calibration
of the mass flow meter has been performed in May 2009. Clearly, miscalibration
can be a limitation of the metabolic simulator that hampers the use of absolute
ersrors. However, miscalibration cannot explain the issues previously discussed,
such as the stability of $\dot{V}_O_2$ and $\dot{V}_C_O_2$ for different values of $\dot{V}_E$. Another limi-
tation of the metabolic simulator is that it does not produce warm and humid
air which can cause considerable disturbance measuring gas fractions [21].

Theoretical error analysis based on general error propagation theory shows
that mixing chamber systems are more accurate and precise than breath-by-
breath systems. Measurements using a metabolic simulator show that the $\dot{V}_O_2$,
and $\dot{V}_{CO_2}$ measured by breath-by-breath systems are less stable for different values of minute ventilation than mixing chamber systems.

**Acknowledgements**

We thank Cosmed Italy for providing the CPET equipment and Accuramed Belgium for providing the metabolic simulator used in this study. The authors have no financial or business interests in any of the companies mentioned in this paper.
Chapter 6

Additional results and discussion

6.1 Metabolic simulator

To compare the accuracy of several commercially available CPET systems for acquisition purposes, measurements using a Vacumed metabolic simulator (Vacumed, USA) are performed using the same methods described in chapter 5. Figure 6.1 shows the results of calibration measurements using a metabolic simulator on the following CPET systems:

- Quark CPET (Cosmed, Italy)
- Oxycon Mobile (Carefusion Germany GmbH, Germany)
- Ergostik (Geratherm Respiratory GmbH, Germany)
- ZAN 680 (nSpire Health, Germany).

All of the above mentioned systems measure metabolic gas exchange according to the breath-by-breath principle, except for the Cosmed Quark CPET, which can measure in both breath-by-breath and mixing chamber mode. Bland Altman plots [65] are used to compare the input of the metabolic simulator to the result of the CPET systems. Figure 6.1 shows the absolute error in (a) $\dot{V}_O_2$, (b) $\dot{V}_CO_2$ and (c) $\dot{V}_E$ and the relative error in (d) $\dot{V}_O_2$, (e) $\dot{V}_CO_2$ and (f) $\dot{V}_E$.

Overall, there are significant differences between the error of equipment from different manufacturers. A smaller spread of data is found for measurements by the Cosmed Quark CPET mixing chamber system, in other words, the mixing chamber system is most stable for different values of ventilation $\dot{V}_E$. The $\dot{V}_O_2$ measured by the mixing chamber system is measured with a small error over the entire range compared to other systems. On the other hand, the $\dot{V}_CO_2$ of the Cosmed Quark CPET mixing chamber system is measured with a comparable error to the Cosmed Quark CPET breath-by-breath system. The differences
Figure 6.1: Bland Altman plots showing the results of $\dot{V}_{O_2}$, $\dot{V}_{CO_2}$ and $\dot{V}_E$ calibration [65]. The vertical axis of (a), (b) and (c) shows the error or difference between the measured value by the CPET system and the configuration of the metabolic simulator (MS), whereas the vertical axis of (d), (e) and (f) shows the relative error or difference between the measured value by the CPET system and the configuration of the metabolic simulator (MS) divided by the configuration of the metabolic simulator. The horizontal axis is configuration of the metabolic simulator.
between the breath-by-breath and the mixing chamber systems are discussed in more detail in chapter 5.

The nSpire ZAN680 underestimates $\dot{V}_E$ significantly over the entire range and shows underestimation of $\dot{V}_{O_2}$ and $\dot{V}_{CO_2}$ in the high range. It is remarkable that the relative error in $\dot{V}_E$ increases with $\dot{V}_E$. The Geratherm Ergostik system overestimates $\dot{V}_{O_2}$, $\dot{V}_{CO_2}$, and $\dot{V}_E$, and shows an increase in relative error as $\dot{V}_{O_2}$, $\dot{V}_{CO_2}$, and $\dot{V}_E$ are increased. Surprisingly, the Carefusion Oxycon Mobile overestimates $\dot{V}_{O_2}$ and underestimates $\dot{V}_{CO_2}$.

According to formulae, an underestimation in $\dot{V}_E$ causes an underestimation in $\dot{V}_{O_2}$ and $\dot{V}_{CO_2}$, since $\dot{V}_{O_2}$ and $\dot{V}_{CO_2}$ are proportional to $\dot{V}_E$ (section 3.2). Vice versa, an overestimation in $\dot{V}_E$ causes an overestimation in $\dot{V}_{O_2}$ and $\dot{V}_{CO_2}$. Surprisingly, this is not the case for metabolic simulator measurements of the Carefusion Oxycon Mobile and the Cosmed Quark CPET equipment. This could be explained by the fact that raw data signals are corrected for the response time of the gas analyzers, and that corrections for the signal of the oxygen and carbon dioxide sensors are significantly different.

Rosdahl et al. have published a validation study of the Carefusion Oxycon Mobile based on biological calibration with the Douglas bag method, and found an overestimation of $\dot{V}_{CO_2}$ [67]. This is a remarkable result, since our measurements using the metabolic simulator show an underestimation of $\dot{V}_{CO_2}$ and an overestimation of $\dot{V}_{O_2}$. The overestimation of $\dot{V}_{CO_2}$ found by Rosdahl et al. was explained in their discussion by a 'faulty setting' in the software affecting the calculation of $\dot{V}_{CO_2}$. Supposedly, this was corrected by the manufacturer afterwards, and verified by the author.

A validation study of the nSpire ZAN680 has been performed by Kusch et al. [68], using a metabolic simulator as described by Prieur et al. [69]. This study observed an overestimation of expired tidal volume $V_T$, and therefore an overestimation of minute volume $\dot{V}_E$. Errors in $\dot{V}_{O_2}$ and $\dot{V}_{CO_2}$ were reported as the maximum flow exceeded 9 L·s$^{-1}$, due to an overestimation of $V_T$. Our results show an underestimation of $\dot{V}_E$ and an underestimation in $\dot{V}_{O_2}$ and $\dot{V}_{CO_2}$ in the high range. Differences between the results of the two studies can be explained by differences in the equipment used. Our ZAN680 CPET system uses a newer software version and a newer version of the flow sensor, showing that differences in components can influence the accuracy of CPET measurements.

This study uses a metabolic simulator to validate CPET equipment. However, the metabolic simulator does not produce warm and humid air which can cause considerable disturbance when determining gas fractions [21]. Furthermore, the validity of measurements with the metabolic simulator relies on correct calibration of the metabolic simulator. Since $\dot{V}_{O_2}$ and $\dot{V}_{CO_2}$ are proportional to the gas flow from the calibration gas tank measured by a mass flow meter, correct calibration of the mass flow meter is crucial. Errors in the mass flow meter influence $\dot{V}_{O_2}$ and $\dot{V}_{CO_2}$ proportionally. Miscalibration of the mass flow meter does not influence $\dot{V}_E$, since $\dot{V}_E$ is independent of the flow of calibration gas, but dependent on the frequency and volume of the motorized calibration syringe or piston pump. Ideally, the mass flow meter is recalibrated regularly. The mass flow meter of the metabolic simulator has been recalibrated.
between measurements of the Cosmed systems and measurements of the Carefusion and nSpire systems. Measurements of the Geratherm system have been performed using a different metabolic simulator using a rotameter to measure the flow of calibration gas. Clearly, miscalibration can be a limitation of the metabolic simulator that hampers the use of absolute errors. However, miscalibration cannot explain the issues previously discussed in this section, such as the discrepancy between $\dot{V}_{O_2}$, $\dot{V}_{CO_2}$ and $\dot{V}_E$.

Measurement of minute ventilation by the systems using variable orifice flow meters, the Geratherm Ergostik and the nSpire ZAN 680, are less accurate than systems using turbine flow sensors, especially at high ventilations of $\dot{V}_E > 100$ L·min$^{-1}$. Moreover, the relative error in $\dot{V}_E$ for both systems is not constant, but increases with $\dot{V}_E$. Calibration measurements as described in section 4.2 can be used to relate the pressure drop over the flexible plastic flap to the flow through the sensor. However, for accurate flow measurements using these sensors, a full range calibration is needed for each flow sensor individually, since mechanical properties of the flap can vary between flow sensors. Moreover, repeated calibrations are necessary because mechanical properties of the flap can change over time due to wear and tear. Optimal calibration of the flow sensor can reduce the error of the flow sensor considerably, but only for stationary flows. Limitations of the flow sensor performance measuring instationary flows are determined by the frequency response [40].
Chapter 7

General discussion and conclusion

7.1 Discussion

The aim of this study is threefold. First of all, this study investigates the accuracy and precision of CPET equipment, in particular, when CPET systems are used to measure the exercise capacity of athletes. Secondly, we are interested in the absolute and relative importance of different sources of error, and for the clinical relevance of this study, whether same types of error play a role when equipment is used for patient care in clinical situations. Thirdly, the development of a quantitative method to study the accuracy and precision of CPET equipment is an important part of this study.

Stationary calibration measurements of a variable orifice flow meter used in the nSpire ZAN680 CPET system have been performed, as described in chapter 4. More calibration measurements of this variable orifice flow meter would have provided a better understanding of the stability, repeatability and variability between different flow meters. However, the measurements described in chapter 4 only partially help in understanding the accuracy and precision of CPET systems as a whole. The same holds for measurements of fixed tidal volumes at different frequencies using the metabolic simulator also described in chapter 4. These measurements show that the error in the measured tidal volume by the variable orifice flow meter varies as the frequency increases.

The accuracy and precision of CPET systems has been studied mainly using two methods. An error analysis method has been developed specifically to investigate the relative importance of input errors. Measurements using a metabolic simulator have been performed to study the error in CPET results of several commercially available systems.

Error analysis shows that the flow error $\delta V$, the delay time error $\delta t_{\text{delay}}$ and the error in temperature of expired air $\delta T_B$ cause significant error in $\dot{V}_{O_2}$ and $\dot{V}_{CO_2}$. The relative error is largely constant over the range of $\dot{V}_{O_2}$, showing
relative errors in the range of 9-12% and 5-7% for breath-by-breath and mixing chamber systems respectively. Relative errors in $\dot{V}CO_2$ are found in the range of 5-7% for breath-by-breath systems and approximately 4% for mixing chamber systems. These results show that mixing chamber systems are more accurate, although breath-by-breath systems are known for a better temporal resolution [37]. The validity of the error analysis method depends largely on the validity of input errors, as discussed in chapter 5.

Measurements using a metabolic simulator show significant differences between commercially available CPET systems. Relative errors in $\dot{V}O_2$ and $\dot{V}CO_2$ as high as 15% are found, whereas relative errors in $\dot{V}E$ are in the range of 0-10%.

Differences between our results and other validation studies [67, 68] are largely explained by differences in hardware and software settings. This shows that the specific settings of CPET equipment are important for the performance of the system. Since manufacturers regularly perform hardware and software updates, it is difficult to compare the results of different validation studies. Moreover, the validation of CPET equipment has not been standardized.

Although cardiopulmonary exercise testing is considered as a gold standard for exercise capacity [70], both our error analysis and the metabolic simulator measurements show that CPET results have substantial measurement errors that are in the range of 4-12%. This range of error hampers clinical application since these errors are well in the clinical relevant range. For example, differences for athletes between first and second place are in the range of 1%, hampering the clinical application of $\dot{V}O_{2,\text{max}}$ testing of athletes. Furthermore, errors in CPET results can influence patient care in clinical situations. CPET has been identified as as a core assessment for patients with heart failure [71] and risk stratification is based on key CPET parameters dividing patients in prognostic categories. Using such categorization, measurement errors in the range of 4-12% may hamper clinical application and limit standardization of cut-off values. Moreover, the effect of training on exercise capacity is difficult to quantify using CPET equipment, since training effects (10-15%) are in the range of measurement error [15, 16] and biological variability is in the order of 5% [11, 12, 13, 14].

7.2 Conclusion

Validation of CPET is a difficult process, since there is no gold standard for CPET. The Douglas bag method is often used to validate systems, but biological calibration studies with human subjects using the Douglas bag method still have to deal with biological variability issues [11, 12, 13]. Furthermore, the Douglas bag method has its own accuracy problems [19], such as the diffusion of gas through the collection bag and leakage of two-way-valves. Calibration measurements using a metabolic simulator are a good alternative without having to deal with biological variability issues. However, metabolic simulators do not produce warm and humid air which can cause significant accuracy problems in
the process of gas analysis. Error analysis has proved itself as a powerful method to study the relative and absolute importance of different sources of error.

Our error analysis method shows that the error in $\dot{V}_O_2$ is larger than in $\dot{V}_C_0_2$. Moreover, the error of breath-by-breath systems is larger than the error of mixing chamber systems. Differences are largely caused by the delay time error $\delta t_{delay}$. In general, the error of the flow sensor $\delta \dot{V}$ and the assumed temperature of the exhaled air $T_B$ are important for the accuracy and precision of CPET results.

Measurements using a metabolic simulator show that breath-by-breath systems are less stable for different values of minute ventilation than mixing chamber systems. Moreover, metabolic simulator measurements show that there are significant differences between the accuracy and precision of different commercially available CPET systems.

The performance of CPET systems relies for a significant part on the specific software and hardware, which is difficult to validate separately. Moreover, manufacturers often release new versions of software and hardware, making making the comparison of different validation studies cumbersome [67, 68]. In general, validation of CPET equipment remains a difficult process, and systems should be considered as new (and recalibrated) when software or hardware changes have been made.
Chapter 8

Recommendations

This chapter discusses recommendations for future research and important considerations for future acquisition processes, as well as recommendations concerning the use of CPET in clinical practice.

8.1 Scientific recommendations

Important results of this study are obtained using a theoretical error analysis method. This method does not include errors caused by breath detection algorithms and the use of two way valves. Future research can be done to incorporate these errors in the error analysis method. Furthermore, individual calibrations of the gas sensors and flow sensor can be performed to obtain a better estimation for the input errors. This also involves incorporating response time corrections of gas signals [42, 43, 44, 45, 46, 47, 48, 49] and angular momentum corrections of signals from turbine flow sensors [41].

One of the most important limitations of the Vacumed metabolic simulator is the inability to produce warm and humid air. The Max II metabolic calibrator used by the Australian Sports Commission is able to produce warm and humid air and calibration services are offered commercially [17, 72]. Expired air is passed across a waterbath filled with hot water of approximately 60°C and the air exits fully saturated with water vapor at about 37°C. Future research could incorporate these techniques in metabolic simulator measurements.

Recently, a novel CPET method has been developed based on respiration chamber technology [73]. The Maastricht Instruments Omnical measures gas fractions of highly diluted expired air, so that fluctuations of gas fractions are minimal and the response time issues of gas analyzers can be neglected. General error propagations theory should be applied to this method to compare the accuracy to other CPET methods.
8.2 Considerations acquisition process

Measurements using the metabolic simulator show that there are large differences between the error in results of CPET equipment. For high values of minute ventilation $V_E > 100$, the two systems with variable orifice flow meters measured $V_E$ with larger error than systems with turbine flow sensors. Therefore, it is recommended to measure the ventilation of elite athletes using an CPET systems with turbine flow meter. Moreover, error analysis shows that mixing chamber systems are more accurate in measuring $V_{O_2}$ and $V_{CO_2}$. When temporal resolution is unimportant, a CPET system with mixing chamber is recommended.

8.3 Considerations clinical practice

Previous studies have shown that the outcome of CPET equipment can depend on system specific software and hardware settings [67]. Ideally, CPET systems are re-calibrated when software and/or hardware updates are performed. For example, this can be done using a metabolic simulator.

Error analysis calculations shows that fluctuations in the ambient oxygen concentration can cause significant error when $V_{O_2}$ is calculated. Sufficient ventilation of the exercise laboratory should prevent carbon dioxide accumulation and oxygen depletion. Although ventilation can cause a drop (or rise) in ambient temperature, this does not play a significant role according to error analysis calculations, especially not when ambient temperature is monitored continuously as most CPET systems do.

Whenever gas fractions are measured, one should know the relative humidity of the gas to determine dry gas fractions. Most CPET systems lead the sampled gas through a Permapure Nafion tube to dry the sample. The sampled gas will not be dried completely, but the vapor pressure of the sample equilibrates with the vapor pressure of ambient air [29]. Since we can measure the vapor pressure of the ambient air, we know the vapor pressure of the gas sample and dry gas fractions can be calculated (section 3.1). However, Permapure Nafion tubes become brown/yellow after many hours of operations and the drying function will be limited. Therefore, Cosmed advises their customers to replace Permapure Nafion tubing after 100 tests. Furthermore, the Permapure Nafion tubing should be dry before a new exercise test is started, since accumulation of water in the Permapure Nafion tubing and possibly the gas sensors due to previous tests will affect the measurement of gas fractions. Drying the Permapure Nafion tubing prior to testing with a ventilator is an easy solution to this problem. To prevent liquid water from entering the sampling tube, the sampling tube should not be mounted downwards on the flow sensor.
References


REFERENCES


Appendix A: Accuracy and precision

Measurements performed with small error are accurate and precise. However, there is a difference between accuracy and precision. Accurate measurements have a small systematic error, whereas precise measurements have a small random error. This is shown schematically in figure 1. Figure 1a shows a precise and accurate measurement, with a small systematic error and a small random error. Figure 1b shows a precise but accurate measurement, where the random error is small but the systematic error is large. Figure 1c is a schematic drawing of an accurate but imprecise measurement, with a small systematic and large random error. Finally, figure 1d shows an inaccurate an imprecise measurement, with large systematic and random errors.

Figure 1: Schematic drawing showing (a), (b), (c) and (d) [60].
Appendix B: Error propagation

Suppose quantity \( y \) is computed with the multi variable function \( f() \) from \( n \) variables, \( x_1, ..., x_n \)

\[
y = f(x_1, ..., x_n).
\]  

(1)

Since variables \( x_1, ..., x_n \) are the result of measurements and are subject to errors depending on the accuracy and precision of the measurement, a quantity \( x_i \) is stated as \( x_i \pm \delta x_i \). This means that the best estimate of the measurement is \( x_i \) and that it is almost certain that the true value lies between \( x_i + \delta x_i \) and \( x_i - \delta x_i \), where \( \delta x \) is called the error or uncertainty. Because variables \( x_1, ..., x_n \) are subject to uncertainties, the computed quantity \( y \) will also be subject to uncertainty \( \delta y \). The uncertainty \( \delta y \) can be calculated according to general error propagation theory [60]

\[
\delta y = \sqrt{\sum_{i=1}^{n} \left( \frac{\partial f(x_1, ..., x_n)}{\partial x_i} \delta x_i \right)^2}
\]  

(2)

and the error found is never larger than the ordinary sum

\[
\delta y \leq \sum_{i=1}^{n} \left| \frac{\partial f(x_1, ..., x_n)}{\partial x_i} \right| \delta x_i.
\]  

(3)