An abnormal respiratory rate is often the earliest sign of critical illness. A reliable estimate of respiratory rate is vital in the application of remote telemonitoring systems, which may facilitate early supported discharge from hospital or prompt recognition of physiological deterioration in high-risk patient groups. Traditional approaches use analysis of respiratory sinus arrhythmia from the electrocardiogram (ECG), but this phenomenon is predominantly limited to the young and healthy. Analysis of the photoplethysmogram (PPG) waveform offers an alternative means of non-invasive respiratory rate monitoring, but further development is required to enable reliable estimates. This review conceptualizes the challenge by discussing the effect of respiration on the PPG waveform and the key physiological mechanisms that underpin the derivation of respiratory rate from the PPG.

Keywords: Photoplethysmography, Respiratory rate, Physiology, Remote monitoring

Introduction

The ability to assess respiratory rate remotely and track its trend over time is essential to the development of physiological telemonitoring. An abnormal respiratory rate is a sensitive early indicator of critical illness that often accompanies, and may precede, changes in other non-invasively monitored vital signs such as heart rate (HR), blood pressure (BP), or reduction in peripheral oxygen saturation (SpO₂) [1].

The ability to obtain multiple vital signs from a single, non-invasive, peripheral sensor is desirable. SpO₂ and heart rate are well-established products of photoplethysmogram (PPG) waveform analysis, and the addition of reliable respiratory rate estimation from PPG analysis would improve user acceptability by reducing the burden of wearing continuous vital sign monitoring hardware for potentially ambulatory patients, as only a single probe would be required if combined with wireless technology (for example, Bluetooth®).

This paper reviews the use of photoplethysmography in respiratory rate estimation. Whilst some authors distinguish between the terms ‘respiration’ and ‘ventilation’ to describe the metabolism of oxygen at a cellular level and the mechanics of breathing, respectively, this paper uses ‘respiration’ in the latter context to correspond with accepted terminology used widely in the literature relating to this topic. For the reader to understand key concepts, this review analyses and presents relevant respiratory and circulatory physiology contributing to the respiratory-induced variations in the PPG waveform.

Overview of physiology relating to the respiratory cycle

Gaseous exchange

Humans acquire oxygen and expel carbon dioxide via gaseous exchange between the alveoli of the lungs and the dense network of pulmonary capillaries. The respiratory cycle ensures a continuous supply of fresh gas to the alveoli and the maintenance of a concentration gradient for diffusion to take place. The total volume of gas inspired in one minute (minute volume) is a product of the volume taken with each breath (tidal volume) and the respiratory rate [2]. Not all entrained gas is accessible for exchange at the alveoli because a proportion, approximating 20–30% of tidal volume, will occupy the conducting airways at the end of inspiration—the ‘physiological dead space’. Through a series of well-conducted experiments using nitrogen gas and flow–volume analysis, this dead-space volume was shown to increase with larger tidal volumes (and therefore an increase in calibre of the conducting airways) but to a greater extent with higher respiratory rates. The latter was explained succinctly by the author: ‘the extent of boundary diffusion between alveolar and dead space gases is a function of time and hence the rate of respiration’ [3]. An understanding of the influence of dead space will provide an explanation as to why slow, deep breathing is
associated with proportionally greater gaseous exchange than rapid, shallow breathing for the same minute volume.

**Neurological control of the respiratory cycle**

The control of the respiratory cycle is orchestrated by the respiratory centre, which consists of ill-defined collections of neurones located in the medulla oblongata and pons of the brainstem. Four general neuronal groups within this region have an essential role in the control of the respiratory cycle: the dorsal respiratory group within the medulla initiates inspiration; the ventrolateral group within the medulla initiates expiration; and the pneumotaxic and apneustic groups within the pons modify the rate and depth of breathing.

The dorsal respiratory group receives sensory input, via the glossopharyngeal and vagus nerves, from peripheral chemoreceptors and baroreceptors adjacent to the arch of the aorta and from stretch receptors within the lung. The dorsal respiratory group controls the basic rhythm of (involuntary) respiration through oscillatory bursts of activity from stimulatory neurones with a pacemaker potential. During high respiratory demand, the ventral group becomes active and augments expiration through a nervous relay that leads to contraction of the accessory muscles of respiration. However, during minimal respiratory demand, the ventral respiratory group remains inactive; instead, expiration is a passive function, resulting from the elastic recoil of the lungs at the end of inspiration.

The pneumotaxic centre modifies the basic rhythm of breathing by shortening the inspiratory phase and, consequently, the expiratory phase. The result of excitatory input from the pneumotaxic centre is an increase in the rate of breathing. A strong pneumotaxic signal may increase the rate of breathing to 30 to 40 breaths per minute, compared with 3 to 5 breaths per minute with a weak signal [4]. Stretch receptors in the lung may also function to curtail the inspiratory phase and limit lung overinflation by innervating the dorsal group via the vagus nerve (Hering–Breur inflation reflex) [2].

The apneustic centre promotes inspiration by inhibiting the input of the pneumotaxic centre and by direct excitatory input into the dorsal respiratory group of neurones.

**Physiological stimuli to increase minute volume**

The primary stimulus to an increase in respiratory depth and frequency is an increasing concentration of carbon dioxide (CO₂) within the blood. CO₂ traverses the blood–brain barrier and in combination with water forms carbonic acid. This fall in pH does not have direct effects on the respiratory centre but stimulates a separate group of neurones located in the medulla known as the central chemoreceptor or chemosensitive area, which inputs into the aforementioned neuronal groups within the respiratory centre. A stimulus to an increased respiratory rate can also occur in response to the sensing of rising concentrations of CO₂ or liberated hydrogen ions by the peripheral chemoreceptors, which are located within the carotid bodies adjacent to the carotid arteries and the aortic bodies adjacent to the aortic arch.

Oxygen binds to haemoglobin in the red blood cells at varying degrees of affinity, depending on the biochemical environment. This enables the delivery of oxygen to tissues at a wide range of partial pressures. Due to this wide variation, oxygen concentration is a lesser stimulus to an increase in respiratory rate, unless hypoxia (blood oxygen partial pressures of <8 kPa or 60 mmHg) ensues. Hypoxaemia is sensed by peripheral chemoreceptors, such as the carotid body, but unlike CO₂, does not directly stimulate the respiratory centre or central chemosensitive area.

**Respiration and heart rate variability**

Respiration has a direct effect on the variability of the heart rate. Stretch receptors within the lung send inhibitory projections to the cardiac vagal neurones (CVN) of the nucleus ambiguus in the medulla. The CVNs, which send efferent projections to the heart via the vagus nerve, reduce the heart rate by inhibitory control. Therefore, an increase in pulmonary stretch receptor activity leads to inhibition of the CVNs and an elevation of heart rate. This association of increased heart rate with inspiration is known as ‘respiratory sinus arrhythmia’ (RSA) [5, 6].

RSA is thought to represent a mechanism for preserving cardiac energy during expiration and optimizing gaseous exchange by matching increased pulmonary perfusion with inspiration [7], although recent experiments assessing pulmonary gas transfer efficiency at various heart rates using cardiac pacemakers have cast some doubt on this hypothesis [8]. The variation inter-beat interval (usually measured between two consecutive R peaks) as a result of RSA can be used to estimate respiratory rate from the electrocardiogram (ECG) [9].

**Respiration and cardiac output in the spontaneously breathing patient**

Following the perfusion of tissues at the capillary level, deoxygenated blood enters the right side of the heart via the peripheral and central venous systems. The primed right side of the heart pumps blood from the right ventricle through the pulmonary artery to the lungs, where gas exchange occurs. In diastole, oxygenated blood flows back to the left side of the heart via the pulmonary veins and returns to the systemic circulation via the aorta during the following cardiac systole.

Peripheral venous pressure (PVP) is usually 8–10 mmHg higher than right atrial pressure, which is usually between 0 and 4 mmHg. This pressure gradient ensures continuous venous return from the peripheral circulation back to the heart. Inspiration augments venous return by two distinct mechanisms:

1. Expansion of the chest wall causes radial traction on the thoracic organs and a reduction in pressure within the heart and blood vessels external to the lungs (i.e. a reduction in intrathoracic pressure). Right atrial pressure decreases to ~2 mmHg on average, widening the pressure difference between the peripheral and central circulations and increasing blood flow from the peripheries to the thorax.

2. The lowering of the diaphragm during inspiration increases pressure external to the intra-abdominal veins, which also augments the return of blood to the heart by increasing the pressure in abdominal vessels.
Photoplethysmography

Photoplethysmography is an optical technique that exploits the wavelength-dependent variation in light absorption coefficient for different tissues. It is typically performed using a pulse oximeter, which usually employs light emitting diodes at two wavelengths: red (660 nm) and near-infrared (940 nm). An increase in blood volume within a tissue will result in an increase in the optical path length, and thus a decrease in the intensity of transmitted light. Most manufacturers invert the waveform, such that a reduced transmission of light is illustrated as a positive deflection. In addition, post-processing of the waveform amplitude, described later, results in a qualitative output, so the waveform is unitless, and comparisons cannot be made between different subjects.

Transmission-mode photoplethysmography uses a light source situated opposite a receiving photodetector, in which the latter senses the light transmitted through the tissues. The requirement for opposing sensors limits the potential attachment sites to distal extremities such as the fingertip, toe, or ear lobe. Reflectance-mode photoplethysmography uses a photodetector adjacent to the emitting light source and is, therefore, not as restricted in the anatomical sites at which it may be employed. Despite the differences in the way that these two methods detect light, the PPG waveforms are similar [12]; in each case, the light arrives at the detector after multiple scattering by the red blood cells [13, 14].

The reflectance method may be more sensitive to the oscillations in venous pressure during the respiratory cycle because the transmission method usually involves using a clip or similar device to secure the probe, which can result in confounding effects of probe pressure on the PPG signal due to external venous compression [12]. However, these effects are likely to remain constant throughout the respiratory cycle. The reflectance method limits this artefact by using an adhesive sticker to secure the probe, but the trade-off may be a reduction in overall signal quality due to the possibility of inferior contact with the skin.

The PPG waveform

The exact origin of the PPG waveform is not clear. Blood flow at capillary level is generally continuous and unlikely to produce a pulsatile waveform, and the low compliance of the distal arterioles is such that changes in vessel diameter are likely to be minimal. An alternative explanation is that the waveform represents pulsatile venous distension from cutaneous arterio-venous anastomosis [15, 16]. At first sight, the pulsatile component of the PPG waveform is similar in morphology to the waveform obtained from arterial blood pressure monitoring, revealing the inherent similarities between blood pressure and tissue perfusion (figure 1). In general, an increase in vessel diameter within the monitored region results in a reduction in transmitted light and a corresponding positive deflection of the (inverted) PPG waveform. Thus, vasodilated or easily distensible vessels, such as veins, give rise to relatively higher PPG amplitudes than vessels of smaller diameter. It should be noted that although increased PPG amplitude may represent increased tissue perfusion, it is not necessarily synonymous with a high blood pressure – peripheral vasoconstriction will cause a reduction in PPG amplitude [17] but may concurrently increase systolic blood pressure.

Components in the PPG waveform

Pulse oximetry relies on the assumption that the PPG signal at both the red and infrared wavelengths consists of two components only: a baseline (d.c.) component and a pulsatile (a.c.) component at the cardiac frequency, which will generally be between 0.5 Hz and 4 Hz, corresponding to 30 beats min$^{-1}$ and 240 beats min$^{-1}$, respectively. This cardiac-synchronous pulsatile component is further assumed to depend solely on the arterial inflow of blood into the monitored region; hence peripheral arterial oxygen saturation (SpO$_2$), the ratio of oxygenated haemoglobin to total haemoglobin in arterial blood, is derived from the ratio of the amplitudes of the pulsatile component at the two wavelengths, usually via a look-up table [18].

It is clear from figure 2 that the amplitude of the PPG waveform at either wavelength also varies in synchrony with the respiratory cycle (here, as is usually the case, this is shown for the red wavelength). This phenomenon can be described either as amplitude modulation of the cardiac-synchronous pulsatile waveform or respiratory-induced intensity variation (RIIV) in the baseline. Which description is adopted depends on the choice of cut-off frequency for the filters used to separate the

---

1Multiple scattering by the red blood cells prevents the Beer-Lambert law, which is valid for purely absorptive media from being used to derive $SpO_2$.
baseline and pulsatile components. A low-pass filter extracts the baseline component (frequencies below the cut-off frequency being let through) and a high-pass filter extracts the pulsatile component (frequencies above the cut-off frequency being let through). If the cut-off frequency is chosen to be 0.5 Hz (as with most pulse oximeters), the baseline component will include the RIIV, for all values of respiratory rate. If the cut-off frequency is chosen to be 0.1 Hz instead, the respiratory-induced variation will appear as amplitude modulation of the pulsatile component, for respiratory rates above six breaths min\(^{-1}\). This distinction is further complicated by the fact that most equipment manufacturers include some form of automatic gain control (AGC), or auto-centring [18], to ensure that the amplitude of the pulsatile component is minimally unaffected by anatomical variations such as finger width (larger fingers will therefore have relatively more light shone through them). The AGC circuitry is driven by monitoring the variations in the amplitude of the baseline component; if the latter includes the respiratory-induced variations (as with a cut-off frequency of 0.5 Hz), then the AGC will attempt to ‘compensate’ for these variations, thereby making it impossible to extract the RIIV reliably from the PPG waveform.

In this paper, we will assume a cut-off frequency of, say, 0.5 Hz, and hence describe respiratory-induced changes in light transmission (or reflection) as baseline variations (RIIV). Baseline fluctuations in the PPG signal with a frequency in the 0.2–0.45 Hz region, synchronous with ventilation, were first reported by Bernardi et al. [19]. There are several possibilities as to the origin of these respiration-synchronous amplitude changes. Firstly, inspiration results in a momentary reduction in stroke volume and, therefore, a corresponding reduction in cardiac output, which should in turn have an effect on the pulsatile component of the PPG waveform. Next, there will also be tissue blood volume changes during the respiratory cycle due to the transmitted changes in thoracic pressure. However, it has also been argued that sympathetically mediated vasoconstriction from ventilatory-synchronous discharge of the autonomic nervous system also plays a part. The evidence for this comes mostly from a study monitoring the PPG waveform obtained from a finger probe placed...
distal to an artery occluded by a blood pressure cuff inflated above systolic pressure. A variation in the baseline of the PPG waveform was evident during a deep inspiratory gasp but not during normal respiration [20]. The occlusion of the vasculature in the limb excluded the possibility that the change in waveform was transmitted intravascularly from mechanical changes in the thoracic cavity. The link suggests a coupling of the respiratory system to the autonomic nervous system, which can affect the PPG waveform independently of the changes resulting from the mechanical effects of the respiratory cycle. Whether a direct autonomic relay between the respiratory system and the peripheral vasculature exists or whether autonomic outputs stimulated by the respiratory system are the result of baroreceptor detection of a change in blood pressure is unclear. Finally, it should be noted that there are also baseline fluctuations in the PPG waveform which are independent of ventilation. Lower frequency (0.1 Hz) baseline fluctuations are evident in humans as part of a separate vascular response to the sympathetic nervous system [21, 22]. These fluctuations are often referred to as Mayer waves [23] and are thought to represent the baroreflex mediated oscillation of arterial blood pressure. Very low frequency (0.01–0.08 Hz) sympathetically mediated variations in the baseline may also be apparent as a vascular response in the regulation of body temperature [24].

Effect of mainly ‘thoracic’ versus mainly ‘abdominal’ contribution to inspiration on baseline fluctuation in PPG

In the study by Johansson and Stromberg [25], participants were trained in varying degrees of thoraco-abdominal (T-A) separation, such that inspiration was achieved predominantly either by expansion of the thorax (T) or by lowering of the diaphragm (A), at different tidal volumes. Tidal volume was measured with a spirometer, and inductive plethysmography belts around the thorax and abdomen were used to record changes in volume. Multiple regression analysis was then performed to correlate tidal volume and T-A separation with the forearm PPG signal and determine which scenarios resulted in the greatest detectable variations in the PPG waveform baseline—the RIIV. Large tidal volumes and predominantly thoracic breathing were reported to have most influence on the RIIV peak-to-peak value.

Association of peripheral venous filling pressures with RIIV in the PPG waveform

Respiratory-synchronous variation in central venous pressure is a well-recognized phenomenon and is illustrated in figure 2. This variation in central venous pressure typically follows changes in arterial pressure in the spontaneously breathing patient. Whether the change in tissue blood volume which gives rise to respiratory-induced changes in the PPG waveform is a consequence of the forward change (transmitted reduction in arterial pressure from low stroke volume) or a backward change (from transmitted reduction in venous pressure due to negative intrathoracic pressures) is unclear [26].

A series of studies by Nilsson et al. [27] analysed the relationship between RIIV and peripheral venous filling pressures. Spontaneously breathing patients were studied by correlating the magnitude of change in peripheral venous pressure (PVP) in forearm veins with the magnitude of RIIV in the PPG waveform in 16 healthy subjects, while varying respiratory rate, tidal volume, and thoraco-abdominal (T-A) separation. The greatest changes in PVP and RIIV corresponded to periods of high tidal volumes, low respiratory rates, and predominantly thoracic breathing. Interestingly, despite a high correlation between PVP and RIIV, this study demonstrates that RIIV usually leads changes in PVP, which suggests that the RIIV in the PPG waveform may be a covariate, rather than secondary to the reverse transmission of changes in filling pressures from within the peripheral (forearm) venous system. However, variation between subjects in this study was large.

A later study by the same group, which included a spontaneously breathing group of subjects, corroborates the finding of their previous study by demonstrating that respiratory-synchronous changes in arterial blood pressure (ABP) again preceded a change in CVP (median −0.2 s), and that RIIV changes in the PPG waveform preceded changes in PVP (by approximately 0.5 s). However, this result did not reach significance [28].

Summary of physiological changes associated with RIIV

The mechanics of respiration, in particular inspiration, can lead to a reduction in tissue blood volume by two distinct mechanisms. The first is by a reduction in cardiac output causing a reduction in arterial blood flow and, therefore, tissue perfusion; the second is by a reduction in intra-thoracic pressure transmitted through the venous system, which siphons the blood from the vascular bed within the tissue.

The implication from the studies to date is that RIIV is closely associated with a reduction in arterial pressure, rather than a backward transmission of reduced venous pressure. Whether the association with reduced arterial pressure is mediated through forward transmission of reduced stroke volume or through an autonomic relay to the skin microvasculature that accompanies the reduction in blood pressure is yet to be elucidated.

Derivation of respiratory rate from the PPG waveform

A robust, reliable, and patient-acceptable method of remotely monitoring respiratory rate is essential to the development of telemonitoring technology. PPG-based estimation of respiration rate is desirable, given that heart rate and oxygen saturation can also be derived from the same probe, but the analysis of respiratory-induced variations in the waveform requires frequency analysis of the PPG baseline, which is often removed in the processing performed by commercial devices. As summarized in §3.5, the cardiovascular and respiratory physiology underpinning the PPG signal is complex, and the exact origin of the PPG waveform remains unclear. However, this does not matter if the parameter of interest is respiratory rate. The PPG waveform is likely to provide better estimates of this parameter than ECG-derived estimates using RSA analysis because variations in the PPG waveform
are at least in part influenced by the mechanics of respiration and are not solely dependent on an intact autonomic nervous system. This dissociation from an RSA dependent estimation of respiratory rate is important because RSA diminishes in the elderly, the critically ill, and those with diseases causing autonomic neuropathy (of which diabetes mellitus is the most prevalent worldwide). Moreover, these patient subgroups will form a large sector of the target population for remote continuous vital sign monitoring as part of health surveillance or post-discharge monitoring.

Reflectance mode PPG has theoretical advantages over transmission mode PPG for the estimation of respiratory rate by potentially avoiding external venous compression. The choice of anatomical site is important, as avoidance of distal extremities may reduce the influence of the sympathetic nervous system on terminal vascular beds. Shelley et al. [18] investigated whether the finger, ear or forehead is best suited to acquire the respiratory signal from the PPG waveform in a study carried out in spontaneously breathing and mechanically ventilated patients. Results from spectral analyses indicate that variations in the PPG waveform acquired from the ear were more than 18 times larger in magnitude than the variations in the PPG waveform acquired from a finger in the ventilated patient and 12 times larger in the spontaneously breathing patient. The authors suggest that this difference can be explained by the proximity of the head to the chest and that the vessels in the head have fewer sympathetic innervations and so are not as likely to express an attenuated signal due to vasoconstriction, when compared with vessels in the extremities [18]. This hypothesis is supported by the work of Awad et al. [29], who tested the sympathetic response of the finger and ear to a cold-water immersion stimulus (4°C) in 12 healthy volunteers. The authors report a 48% reduction in the amplitude of the PPG waveform acquired from a contralateral finger, but just a 2% reduction in the amplitude of the PPG waveform acquired from the ear.

A later study by Nilsson et al. [30] compared several anatomical sites (the shoulder, wrist, forearm and finger), using a combination of reflectance and transmission methods of photoplethysmography, for the synchronous measurement of heart rate, respiratory rate and oxygen saturation. The authors conclude that no single site is superior for acquiring accurate estimates of all of the desired vital signs because the best site for detecting the heart rate (the finger) tended to be the worst site for detecting respiratory rate. However, reflectance PPG from the forearm resulted in significantly more accurate estimation of respiratory rate than the other anatomical sites included in the study. No single site is optimal for all PPG-derived vital signs, but proximity to the chest is more likely to return a signal that demonstrates RIIV. The largest respiratory induced variations in the PPG baseline are evident at low respiratory rates, high tidal volumes, and during ‘thoracic’ rather than ‘abdominal’ breathing.

Declaration of interest: The Centre of Excellence in Personalised Healthcare, funded by the Wellcome Trust and EPSRC under grant number WT 088877/Z/09/Z, and the NIHR Biomedical Research Centre Programme supported this work.

References

Cardiovascular Research, 70, 12–21.

Spectral analysis of AC and DC components of the pulse photoplethys-
mograph at rest and during induction of anaesthesia. International 

and thoraco-abdominal separation on the respiratory induced varia-
tion of the photoplethysmogram. J Journal of Clinical Monitoring and 

not the sole determinant of respiratory induced variations in the reflec-
tion mode photoplethysmographic signal. Physiological Measurement, 
24, 925–937.

in the reflection mode photoplethysmographic signal. Relationships 
to peripheral venous pressure. Medical & Biological Engineering & 

can be monitored by photoplethysmography with high 
sensitivity and specificity regardless of anaesthesia and ven-
tilatory mode. Acta Anaesthesiologica Scandinavica, 49, 
1157–1162.

& Shelley, K.H., 2001, Different responses of ear and finger pulse ox-
imeter wave form to cold pressor test. Anesthesia & Analgesia, 92, 
1483–1486.

2007, Combined photoplethysmographic monitoring of respiration 
rate and pulse: a comparison between different measurement sites in 
spontaneously breathing subjects. Acta Anaesthesiologica Scandinavica, 
51, 1250–1257.