Non-contact Vital Signs Monitoring in Paediatric Anaesthesia – Current Challenges and Future Direction

Nicole Grech\(^1\)*, Jean Calleja-Agius\(^2\), Stephen Sciberras\(^1\), Neil Micaleff\(^3\), Kenneth Camilleri\(^3\), Owen Falzon\(^3\)

**ABSTRACT**

Non-contact vital sign monitoring is an area of increasing interest in the clinical scenario since it offers advantages over traditional monitoring using leads and wires. These advantages include reduction in transmission of infection and more freedom of movement. Yet there is a paucity of studies available in the clinical setting particularly in paediatric anaesthesia. This scoping review aims to investigate why contactless monitoring, specifically with red-green-blue cameras, is not implemented in mainstream practise. The challenges, drawbacks and limitations of non-contact vital sign monitoring, will be outlined, together with future direction on how it can potentially be implemented in the setting of paediatric anaesthesia, and in the critical care scenario.

**KEYWORDS**

contactless; vital signs; paediatrics; anaesthesia; monitoring

**AUTHOR AFFILIATIONS**

\(^1\) Department of Anaesthesia and Intensive Care Medicine, Mater Dei Hospital, Malta
\(^2\) Department of Anatomy, Faculty of Medicine and Surgery, University of Malta
\(^3\) Centre for Biomedical Cybernetics, Faculty of Engineering, University of Malta

* Corresponding author: Department of Anaesthesia and Intensive Care Medicine, Mater Dei Hospital, Malta; nicole.grech@gov.mt

Received: 21 January 2023
Accepted: 8 May 2023
Published online: 8 November 2023
INTRODUCTION

Non-Contact Vital Signs Monitoring (NCVSM) is an area of science and healthcare that has garnered increasing interest in the past years due to its advantages both for patients and healthcare workers. Among the obvious advantages of monitoring vital signs without the use of leads and wires attached to patients is improved comfort and increased ability to mobilise (1). A more in-depth analysis highlights the potential for reduced transmission of multidrug resistant organisms (MDROs) through inadequately disinfected monitoring equipment; reduced issues of skin damage especially in vulnerable populations such as neonates, elderly frail people and burns patients; and reduced need for hospital staff to enter isolation rooms and waste precious personal protective equipment (PPE) just to replace a lead or sort out dislodged wires (2–6).

In children, especially those admitted to hospital, monitoring may prove difficult as paediatric patients may be stressed by the unfamiliar environment and the natural fear of upcoming procedures, or due to pain after such procedures. Children may not always cooperate with healthcare workers and the inability to keep monitoring leads and wires on for long enough to get a reliable reading of vital signs, is a common issue. Although inducing anaesthesia in a child who is not fully monitored is not desirable and increases risks of adverse events, it often ends up having to be done if the child is combative and will not stay still long enough for electrocardiography (ECG) leads and pulse oximeters to be appropriately positioned. Post-op, paediatric patients often wake up delirious and uncooperative. This may persist into the post-anaesthesia care unit (PACU), the intensive care unit (ICU), or the general wards. Once again, this renders paediatric patient care difficult and increases risks of deranged vital signs being detected at a much later stage (7–11).

The ability to monitor a child’s vital signs in a non-contact manner therefore has obvious advantages in that it does not further distress the child, and may provide reliable readings for healthcare professionals in the case where the child is uncooperative and keeps removing monitoring equipment. In toddlers and infants, it may be possible that they even put leads in their mouth, posing a risk to themselves and also cause aberrant readings. Contactless monitoring also allows children to move more freely and be monitored while performing enjoyable tasks such as playing. This allows more accurate vital signs to be obtained since the sympathetic response to stress is reduced (12).

In the specific scenario of paediatric anaesthesia, NCVSM offers advantages even during maintenance of anaesthesia, apart from those already described relating to induction and emergence. Patient positioning in the operating theatre often dislodges leads and may cause monitoring equipment to become entangled, leading to potential damage to the internal structure of the wires. Once surgery has started and sterile surgical drapes are in place, access to monitoring equipment is very restricted. Attempts at manoeuvring and applying monitors under the drapes may lead to issues for the surgeon and displacement of surgical equipment, and is ultimately dangerous for the patient. Thus, any leads that get dislodged may lead to distorted readings of loss of monitoring capabilities. Use of diathermy often interferes with the quality of signals obtained, with ECG leads being particularly vulnerable to electrical interference (13). During long procedures, leads can put pressure on the fragile skin of small vulnerable children and in some interventions such as thoracic procedures they often need to be placed on the back, causing more pressure and potentially suboptimal monitoring (14, 15). Neurological complications such as acute brain injury, occurring for example during venoarterial extracorporeal membrane oxygenation (ECMO), may be detected better with non-invasive neuro-monitoring. Neuro-monitoring can be carried out via electroencephalography (EEG), somatosensory evoked potentials (SSEP) and near infrared spectroscopy (NIRS) (16). Cerebral desaturation measured using NIRS was associated with a poor short-term outcome in children undergoing ECMO (17). This is particularly important in an intensive care setting.

NCVSM can be carried out through vital sign extraction from red-green-blue (RGB) video recordings in real time (18). The beating heart produces rhythmically increased perfusion of the subcutaneous capillaries causing an increased pink tone in the skin, which is imperceptible to the naked eye. This occurs because the increased concentration oxygenated haemoglobin in these capillaries absorbs more of the blue spectrum of visible light and reflects red wavelengths. These changes can be extracted using multiple algorithms and they can be converted to a numerical value for heart rate (HR), and, subsequently, heart rate variability (HRV) over a period of time. This technique is termed reflectance photoplethysmography (rPPG) (19, 20). Commonly used algorithms include Eulerian video magnification (EVM), which essentially applies green filters to enhance the skin colour changes, and principal or independent component analysis (PCA and ICA respectively). The video is decomposed into different spatial frequency bands, where each spatial frequency band extracts the level of spatial detail required. All the bands are then filtered with the same time-domain filter to extract the motion detail required (e.g. heart rate perfusion), amplified and added to the original spatially filtered image frames, thus magnifying temporal changes in the video that may be imperceptible to the unaided eye (21–24). More recently, convolutional neural networks (CNNs) have gained huge popularity. These are very complex black box algorithms that can be trained to recognise a particular characteristic on a data set, often termed the “training” set, such as the skin colour changes in a set of data on healthy volunteers. Subsequently CNNs can be fed real world “test” data from which they should be able to detect the same changes that they learnt how to extract during training (25–27). The sensitivity and accuracy and reliability of these kind of systems can potentially be improved when used with machine learning in artificial intelligence (28).

Respiratory rate (RR) detection involves extraction of the rhythmic movements of the chest which translate to change in pixel intensity on RGB videos and images. To extract these signals, optical flow algorithms may be used as well as PCA or ICA with the focus being chest movements, and CNNs which can be trained to recognise the signal and are then applied to real world data (29–32).
One thing common to these algorithms is that a region of interest (ROI) must be selected, which is an area within the image, generally on the face of the patient, from which signals will be extracted. Other body parts may also be used such as hands and feet, but each poses its own challenges. In some algorithms, automatic recognition of the patients within the video frame and automatic selection of the best ROI for signal extraction is possible (33–35). While the technical details of data and image analysis are beyond the scope of this clinically oriented review, the limitations and how these affect clinical practise will be discussed.

In spite of the potential benefits to using NCVSM, this is not as yet routinely used in clinical practise. Experimental studies have been performed; however, there is a marked paucity of data on the use of these methods in terms of sensitivity and accuracy in the paediatric population, even more so in the perioperative period. The main aim of this review is to explore the limitations and challenges of NCVSM, in order to pave the way for increased uptake of this type of clinical monitoring in paediatrics.

**METHODOLOGY**

This review involves an analysis of studies carried out since 2018 to date (May 2022) that involve monitoring of HR and RR by RGB cameras in real world hospital settings involving paediatric populations. The results obtained from these studies were analysed from a medical rather than a technical perspective, focusing on the limitations that may be contributing to the restricted use of these methods in everyday practise.

Boolean operators were used to search multiple search engines, namely Google Scholar, Medline, IEEE Explore, Cochrane Database, SCOPUS and CINAHL. In the case of Google Scholar, since over seventeen thousand results were obtained, the first hundred results were analysed. The search terms entered were the following: (contactless OR non-contact OR noncontact OR wireless OR RGB OR camera) AND (infant OR child OR baby OR paediatric OR neonate OR newborn) AND (heart OR heartbeat OR respiration OR breathing) NOT (wearable OR radar)

---

**Fig. 1 PRISMA flow diagram summarising search strategy and results (39).**
Inclusion criteria were: studies published in peer-reviewed journals since 2018 and available in the English language; studies that include human subjects, performed on children less than eighteen years of age and carried out in clinical settings such as clinics or hospitals; inclusion of the use of RGB cameras and the extraction of HR or RR from RGB video data.

Exclusion criteria included studies published outside of peer review journals, including conference papers and symposia; studies published before 2018, not published in English and for which an English translation is not available; studies involving animals or simulators but no human subjects; involving only patients over eighteen years of age; performed outside of clinical settings; not involving RGB cameras or extraction of HR or RR data.

Studies involving RGB cameras were exclusively chosen for this review because these are the most widely available and are very cost effective. Some studies have in fact successfully extracted vital parameters using smartphone cameras, which in this day and age tend to be of very high quality and are overall very affordable (36, 37). Other types of cameras are used to extract vital signs from video data, such as depth cameras, structured light cameras and thermal cameras and although these have been successful in several trials, prohibitively high costs are a reason why they are not widely used. Therefore, these types of cameras were excluded from our review (38).

The PRISMA flow diagram as shown in Figure 1 summarises the search results obtained and the final full text papers included in this review.

The papers which have been included were analysed and the results obtained from the studies were categorised according to accuracy of vital signs obtained from video recordings and the various limitations posed by the clinical setting in which they were studied. This was done to attempt to understand what is holding back from the mainstream use of these NCVSM devices in clinical practice.

RESULTS

Sixteen studies are included in this review. The included studies took place within hospitals, most commonly in neonatal intensive care units (NICUs) on children aged less than one year. Two of the included studies are reviews themselves and focus on NICUs, but do not include the most recent studies and are limited to NICU babies only (40, 41). Of the experimental studies, thirteen of them include a total of one hundred and fifty patients (42–55). This is a small sample size and the focus is mainly on just one clinical setting, albeit a complex one. No reports were identified in which children were studied specifically in the operating theatre setting or in the peri-operative period, and few studies included children who were older than the infant period, that is, above a year of age (56).

Table 1 summarises the studies included, the vital sign monitored, the population studied and its size, brief details of algorithms used and results obtained.

The first, and indeed, the most important factor to consider in any monitoring system is the accuracy of the system. In medicine, accurate readings of vital signs are essential in assessing and managing patients, even more so in anaesthesia when patients are unable to report symptoms (57). Studies that involve RGB cameras extract readings from videos taken in clinical scenarios and compare the vital signs readings extracted from these videos with those obtained by gold standard contact methods, most commonly the multiparameter monitors available in the hospital setting (55). A direct comparison between the results obtained by different research teams is difficult since some quote the root mean square error (RMSE), some the mean absolute error (MAE) for their dataset, and others quote other parameters such as the Pearson correlation coefficient (PCC). However, error values as low as an MAE of 1.8 beats per minute were reported for respiratory rate by Lucy et al. and MAE of 1.8 beats per minute was reported by Khanam et al. for their study involving heart rate extraction (55). Lucy et al. studied a group of five children known to have pneumonia and extracted their respiratory movements using colour component analysis in their video frames and then applied Fourier transformation to obtain respiratory rates (46). Khanam et al. used a deep neural network called YOLO 3 to automatically select the best ROI in their video frames obtained from seven NICU patients and to subsequently also extract colour changes corresponding to heart rates (55). These accuracy rates are excellent, since any deviation in heart rate or respiratory rate lower than two beats or breaths each minute is clinically not very significant. In paediatric surgery, small variations in parameters are usually not a cause for concern and do not prompt administration of drugs or any other changes (58). However, issues arise since this accuracy rate is not always maintained throughout long periods of time, with periods of time that involve changes in ambient illumination or movement of the patient affecting outcomes.

One major issue encountered in several studies that rely on RGB cameras to extract heart rates is concerning changes in ambient lighting (47, 49, 55). If low levels of light are available, not enough wavelengths may be reflected off the skin to obtain a significant result, and variations in ambient light during a video clip will cause worsening accuracy rates with values that begin to differ significantly from the gold standard monitoring (59).

During some surgeries, relative darkness is necessary, such as during laparoscopy when the laparoscopy monitor is the main source of light. Shadows cast by staff members moving around the patient also provide challenges and are noted to increase error values (42). One possible solution for this is to have a narrow beamlight source such as an LED or a bulb shining on the ROI chosen for the particular patient (49). The ROI selected in a theatre setting must allow for surgical drapes and equipment such as forced air blankets that often cover much of the body, and this itself is another practical issue (45). However, uncovering a small part of the body far away from the operative site, such as the hands or feet in the case of abdominal surgery, or the forehead, will make this possible. Having a dedicated light source will add an extra piece of equipment to an often already cluttered anaesthetic space, and will also incur the extra cost associated with this particular equipment. However, overhead bulbs can be small and unobtrusive if placed strategically.
Tab. 1 Summary of all studies included in this review with brief details of population studied, clinical setting and algorithms used with results obtained. Key to abbreviations: FFT – Fast Fourier Transform; PFF – Principal Flow Field; MSD – Micromotion and Stationary Detection; EEMD – Ensemble Empirical Mode Decomposition.

<table>
<thead>
<tr>
<th>Year</th>
<th>Authors</th>
<th>Population</th>
<th>Location</th>
<th>Algorithm Used</th>
<th>Parameter extracted</th>
<th>Error Value</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>2018</td>
<td>Cobos Torres et al.</td>
<td>9 preterm infants</td>
<td>NICU</td>
<td>FFT</td>
<td>HR and RR</td>
<td>PCC 0.94 for HR and 0.86 for RR</td>
<td>Movement, lighting</td>
</tr>
<tr>
<td>2019</td>
<td>Chaichulee et al.</td>
<td>15 preterm infants</td>
<td>NICU</td>
<td>Deep learning framework</td>
<td>HR and RR</td>
<td>98.8% accuracy results</td>
<td>Lighting</td>
</tr>
<tr>
<td>2019</td>
<td>Sun et al.</td>
<td>5 preterm infants</td>
<td>NICU</td>
<td>Conventional optical flow and deep learning optical flow</td>
<td>RR</td>
<td>Cross-correlation coefficient of 0.70 for conventional optical flow and 0.74 for deep learning</td>
<td>Image resolution, background noise, lighting</td>
</tr>
<tr>
<td>2019</td>
<td>Gibson et al.</td>
<td>10 premature infants</td>
<td>NICU</td>
<td>EVM and FFT</td>
<td>HR and RR</td>
<td>Mean difference of 4.5 bpm for HR and mean bias of 0.8 bpm for RR</td>
<td>Lighting, camera movement,</td>
</tr>
<tr>
<td>2019</td>
<td>Villarroel et al.</td>
<td>30 premature infants</td>
<td>NICU</td>
<td>Convolutional neural network</td>
<td>HR and RR</td>
<td>MAE 2.3 bpm for HR and 3.5 bpm for RR</td>
<td>Background noise, motion</td>
</tr>
<tr>
<td>2019</td>
<td>Paul et al.</td>
<td>19 neonates</td>
<td>NICU</td>
<td>Short time Fourier transform</td>
<td>HR</td>
<td>Segments of 3 bpm of difference obtained</td>
<td>Motion, light, ROI tracking</td>
</tr>
<tr>
<td>2019</td>
<td>Rossol et al.</td>
<td>18 infants</td>
<td>NICU</td>
<td>EVM, PFF, MSD</td>
<td>RR</td>
<td>RMSE 6.36 bpm</td>
<td>Lighting and movement but MSD is robust to these</td>
</tr>
<tr>
<td>2021</td>
<td>Chen et al.</td>
<td>9 neonates</td>
<td>NICU</td>
<td>EVM with majority voting</td>
<td>HR</td>
<td>MAE 3.39 bpm at rest and 4.34 bpm during movement</td>
<td>Unexpected head movement</td>
</tr>
<tr>
<td>2021</td>
<td>Khanam et al.</td>
<td>7 neonates</td>
<td>NICU</td>
<td>CNN with automatic ROI selection and EEMD to reduce noise</td>
<td>HR and RR</td>
<td>MAE of 1.8 bpm for HR and 2.13 bpm for RR</td>
<td>Lighting, subject and camera movement</td>
</tr>
<tr>
<td>2021</td>
<td>Wieler et al.</td>
<td>28 term neonates</td>
<td>Maternity hospital nursery</td>
<td>Publicly available software, FFT and manual ROI selection</td>
<td>HR</td>
<td>RMSE 20.4 bpm</td>
<td>Higher birth weight, movement</td>
</tr>
<tr>
<td>2021</td>
<td>Lorato et al.</td>
<td>17 infants</td>
<td>Medium care unit</td>
<td>MATLAB software with labelling of motion in videos</td>
<td>RR</td>
<td>MAE 3.31 bpm on testing and 5.36 on validation dataset</td>
<td>Severe motion (although small motions compensated), clothing, suckling</td>
</tr>
<tr>
<td>2021</td>
<td>Lucy et al.</td>
<td>5 children with pneumonia</td>
<td>Hospital</td>
<td>FFT and denoising algorithms</td>
<td>RR</td>
<td>MAE of 1.8 bpm</td>
<td>Motion, illumination changes</td>
</tr>
<tr>
<td>2021</td>
<td>Paul et al.</td>
<td>1 baby</td>
<td>NICU</td>
<td>Feature maps to identify pulsatile signals, FFT</td>
<td>HR</td>
<td>Not mentioned</td>
<td>Shadows, reflective materials, movement</td>
</tr>
<tr>
<td>2021</td>
<td>Nagy et al.</td>
<td>7 infants</td>
<td>NICU</td>
<td>CNN with detection of ROI and severe motion periods – 2 different algorithms</td>
<td>HR and RR</td>
<td>MAE 7.08 bpm and 6.19 bpm for HR when non-ideal conditions and ideal conditions were present respectively; MAE 5.08 bpm and 2.03 bpm for RR when non-ideal vs ideal conditions present respectively</td>
<td>Lighting, movement</td>
</tr>
</tbody>
</table>

Patient movement is another major impediment mentioned in most studies. This is especially noticeable in studies involving the neonatal population since infants often exhibit uncoordinated and random movement. Movement may lead to the ROI selected moving out of the video frame as well as impeding the focus on a particular ROI if there is constant movement (41, 44, 46, 49–51, 54). In an operating theatre setting, intra-operatively, movement is often not an issue since patients are often paralysed, and almost all are heavily sedated enough that gross movement will not occur. This is, of course, completely different when the child is awake especially pre- and post-op, when distress may be significant and children often refuse to stay still for any appreciable length of time. This may be a valid reason...
why these methods are still not in mainstream use and only advancements in the algorithms used for data analysis can overcome this issue. Many motion stabilisation algorithms are already being trialled and show promising results, especially in terms of accuracy and reliability (53, 54).

Different skin tones reflect light wavelengths differently and may confound the interpretation of rPPG signals. Therefore, studies that include members of the population with different skin tones are important to develop robust algorithms following camera calibration on the skin (60). There is a paucity of data regarding the skin tones of children studied, and only one study identified in this review by Paul et al. specifically mentions the effect of skin tone on vital signs extraction (51). Including children with different skin tones should not be an issue especially in the operating theatre setting, since changes in light and patient movement can be practically abolished in anaesthesia, allowing investigation of skin tone as the sole confounding factor.

DISCUSSION

This concise review aims to provide a brief overview of NCVSM, specifically heart rate and respiratory rate, by use of RGB cameras. RGB cameras are advantageous as they are inexpensive and widely available, with many studies successfully obtaining heart rate and respiratory rate measurements by using regular smartphone cameras. Other means of monitoring vital signs certainly exist, such as different types of radar, ultrasound Doppler and thermal imaging cameras, but these come with added costs and complexity and in some cases such as radar, potential health safety concerns about prolonged use abound, in view of possible increases in thermal energy transfer (61, 62).

Studies based on the paediatric population make up a significant proportion of non-contact vital sign studies carried out in clinical areas; however the studies are few and include small populations. Very few studies include patients with specific pathologies, and no study was identified that actually took place in the operating theatre or patients with specific pathologies, and no study was identified in this review. There is a paucity of data regarding the skin tones of children studied, and only one study identified in this review by Paul et al. specifically mentions the effect of skin tone on vital signs extraction (51). Including children with different skin tones should not be an issue especially in the operating theatre setting, since changes in light and patient movement can be practically abolished in anaesthesia, allowing investigation of skin tone as the sole confounding factor.

Studies based on the paediatric population make up a significant proportion of non-contact vital sign studies carried out in clinical areas; however the studies are few and include small populations. Very few studies include patients with specific pathologies, and no study was identified that actually took place in the operating theatre or peri-operative environment, highlighting a need to explore this area further. Although these technologies show promise, there are still a number of limitations to tackle such as issues of patient or camera movement and changes in lighting. Many studies have in fact focused on excluding periods with a lot of noise signals such as during clinical intervention and accepting brackets of time in which the patient may not be monitored (43, 50). This, however, is often unacceptable in paediatric anaesthetic practise. A large lacuna exists regarding how sensitive, accurate and reliable these systems are or can be, especially in a paediatric setting.

The theatre environment lends itself very well to such studies. The availability of camera-based monitoring would be of great advantage to vulnerable populations such as infants and patients with skin conditions or burns and so on, especially during prolonged procedures or surgeries where monitoring equipment such as ECG leads must be placed on the back, potentially leading to excessive pressure being applied to the patient leading to irritation and, in severe cases, ulceration (3). Patients undergoing surgery are also ideal candidates for data collection since they do not move, and illumination can be kept constant by dedicated light sources if the theatre lighting is inadequate (49).

One further point to consider which has not been deeply tackled in any of the studies identified as part of this review is the issue of privacy and data protection (63). RGB video data is easily identifiable and therefore data protection policies must be in place to ensure ethical handling of collected videos (64). The fact that a paediatric population is being studied and particularly at a time when the patient is most vulnerable in the peri-operative period makes such issues even more important to tackle and consent from the child’s legal guardian or the children themselves if they are deemed to have capacity to consent is of utmost importance (65). This also makes sharing of data with other scientist teams more difficult and possibly ethically questionable and may affect the ability of different teams to build on each others’ work and achieve better results.

CONCLUSION

NCVSM is an exciting new area of medicine which offers unique advantages in clinical practise to both the patient and the attending clinicians. Although interest in the field is growing and progress is being made, there is still a long way to go in terms of obtaining adequate accuracy and sensitivity levels and overcoming practical obstacles to be able to incorporate NCVSM into everyday practise to give reliable and replicable results. The need for further studies within the hospital environment is highlighted in this review.

FUNDING


REFERENCES


