



Effectiveness of consumer-grade contactless vital signs monitors: a systematic review and meta-analysis

Chi Pham^{1,2} · Khashayar Poorzargar^{1,2} · Mahesh Nagappa³ · Aparna Saripella¹ · Matteo Parotto⁴ · Marina Englesakis⁵ · Kang Lee⁶ · Frances Chung^{1,2}

Received: 29 April 2021 / Accepted: 19 June 2021 / Published online: 9 July 2021
© The Author(s), under exclusive licence to Springer Nature B.V. 2021

Abstract

The objective of this systematic review and meta-analysis was to analyze the effectiveness of contactless vital sign monitors that utilize a consumer-friendly camera versus medical grade instruments. A multiple database search was conducted from inception to September 2020. Inclusion criteria were as follows: studies that used a consumer-grade camera (smartphone/webcam) to examine contactless vital signs in adults; evaluated the non-contact device against a reference medical device; and used the participants' face for measurement. Twenty-six studies were included in the review of which 16 were included in Pearson's correlation and 14 studies were included in the Bland–Altman meta-analysis. Twenty-two studies measured heart rate (HR) (92%), three measured blood pressure (BP) (12%), and respiratory rate (RR) (12%). No study examined blood oxygen saturation (SpO₂). Most studies had a small sample size (≤ 30 participants) and were performed in a laboratory setting. Our meta-analysis found that consumer-grade contactless vital sign monitors were accurate in comparison to a medical device in measuring HR. Current contactless monitors have limitations such as motion, poor lighting, and lack of automatic face tracking. Currently available consumer-friendly contactless monitors measure HR accurately compared to standard medical devices. More studies are needed to assess the accuracy of contactless BP and RR monitors. Implementation of contactless vital sign monitors for clinical use will require validation in a larger population, in a clinical setting, and expanded to encompass other vital signs including BP, RR, and SpO₂.

Keywords Contactless monitors · Vital signs · Camera · Photoplethysmography · Heart rate · Blood pressure

Abbreviations

HR	Heart rate
RR	Respiratory rate
BP	Blood pressure
SpO ₂	Hemoglobin oxygen saturation
PPG	Photo-plethysmography
Ippg	Imaging photo-plethysmography

COVID-19	Coronavirus Disease 2019
ECG	Electrocardiogram

1 Introduction

The COVID-19 pandemic has drastically impacted the healthcare system and has changed the way in which healthcare is delivered. Virtual care in the form of telemedicine

Chi Pham and Khashayar Poorzargar have contributed equally to this manuscript.

✉ Frances Chung
frances.chung@uhn.ca

¹ Department of Anesthesia and Pain Medicine, Toronto Western Hospital, University Health Network, University of Toronto, 399 Bathurst Street, McL2-405, Toronto, ON M5T 2S8, Canada

² Institute of Medical Science, Faculty of Medicine, University of Toronto, Toronto, ON, Canada

³ Department of Anesthesia and Perioperative Medicine, Schulich School of Medicine and Dentistry, London Health

Sciences Centre and St. Joseph Health Care, Western University, London, ON, Canada

⁴ Department of Anesthesia and Pain Medicine, Toronto General Hospital, University Health Network, University of Toronto, Toronto, ON, Canada

⁵ Library and Information Services, University Health Network, Toronto, Canada

⁶ Ontario Institute for Studies in Education, University of Toronto, Toronto, ON, Canada

transitioned from being a novelty to a necessity during the COVID-19 pandemic [1]. Current telemedicine methods are facilitated through telephone or video calls. Telephone visits have the advantage of being universally available to all patients but lack the ability to facilitate a physical examination. The current standard for virtual care is a video call, which allows health care providers to visually examine patients [2]. The measurement of vital signs is part of in-person medical visits, but it has not been integrated into the current telemedicine system.

Vital sign measurements that can be obtained by a consumer-grade camera, such as a webcam or smartphone camera, can offer patients a standard of care that closely resembles a visit to the clinic. These contactless monitors require the user to take a short video of their face using a mounted smartphone or webcam. The video will then be processed by a computer algorithm that can detect changes in light reflection off the facial skin and convert this to vital signs measurements. This process does not require the user to come into contact with the phone or webcam, if the buttons on the device are selected by another individual, which gives it the potential to act as electronic personal protective equipment in public clinics or hospitals. Given the ubiquity of smartphones, vital sign monitors that only require a camera are more accessible compared to existing self-monitoring alternatives such as smartwatches or automatic blood pressure cuffs. The primary objective of this systematic review and meta-analysis is to assess the effectiveness and accuracy of consumer-grade contactless vital signs measuring technologies such as smartphone cameras or webcams in comparison to standard medical devices.

2 Methods

2.1 Protocol and registration

This systematic review was performed in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [3]. This systematic review was registered in PROSPERO (CRD42020210938).

2.2 Study selection criteria

The inclusion criteria were as follows: (1) use of contactless vital signs measuring technologies in adults 18 years and older measuring vital signs [blood pressure (BP), heart rate (HR), respiratory rate (RR), and hemoglobin oxygen saturation (SpO_2)]; (2) compared to a reference medical device; and (3) using a consumer-grade camera (smartphone or webcam) with the participants' face as the region of interest (ROI). The exclusion criteria were as follows: (1) all

conference papers, case series, case reports and (2) non-English articles.

2.3 Data sources

We searched for articles published in the following electronic databases from inception (1946) to September 2020: MEDLINE (Ovid), PubMed-Not-MEDLINE (NLM), Embase, Cochrane Central Register of Controlled Trials, and Cochrane Database of Systematic Reviews. The Citations were searched to capture potentially missed articles. The search strategy was designed by an information specialist experienced in a systematic review (ME). The search terms included “non-contact”, “cell phones”, “webcam”, “camera”, “vital signs”, “heart rate”, “respiratory rate”, “blood pressure”, “oximetry” and all search terms are listed in Supplementary File.

2.4 Study selection and data extraction

The search yielded 16,927 articles after the duplicates were removed. After screening the titles and abstracts, 97 articles were assessed for full-text eligibility. An additional 37 articles were identified from citations. A total of 26 studies with 1,937 patients were included in the qualitative analysis. The PRISMA diagram of the search results is shown in Fig. 1.

Two reviewers (KP and AS) independently performed the title, abstract, and full-text screening. A third reviewer (CP) and a senior author (FC) were consulted on any discrepancy that was not resolved by discussion. Data were extracted by three reviewers (CP, KP, and AS) and managed using Excel. Two reviewers (KP and AS) divided the studies and extracted the data while a third reviewer (CP) independently extracted data from all studies for comparison. We extracted the vital sign(s) of interest, the number of participants, age, ethnicity, eligibility criteria, image technology, reference method(s), distance from camera to facial region, position of the camera, source of light, performance of new technology compared to the reference method, and reported limitations. Data on Bland–Altman bias and its standard deviation, as well as Pearson's correlation coefficient, were extracted from 14 and 16 studies respectively. WebPlotDigitizer was used to extract Bland–Altman values from Bland–Altman plots and Pearson's correlation was calculated using Bland–Altman values when possible.

2.5 Quality assessment

Two reviewers (KP and AS) independently assessed the quality of the methodology reporting with respect to the Guidelines for Reporting Reliability and Agreement Studies (GRRAS) [4]. We used a modified GRRAS reporting tool

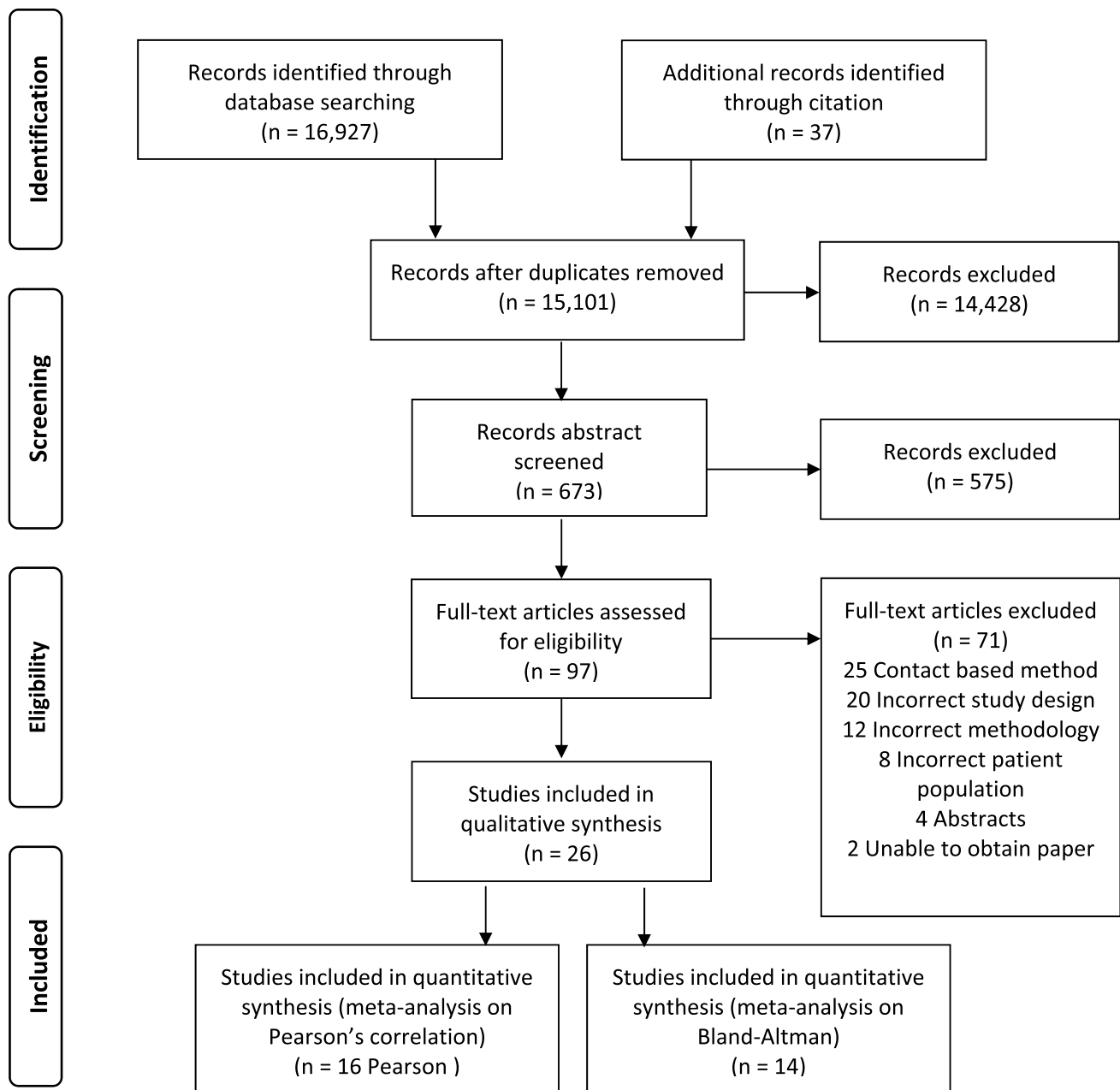


Fig. 1 PRISMA flow diagram. Flow diagram of the study selection process

based on a review of image-based, non-contact sensors, and vital sign monitoring methods [5].

The quality of the statistical analyses was assessed by assigning a rating, ranging from 0 to 3, based on a classification of statistical methodologies and reported measurements described in each study. The standard, with a rating of 3, was assigned if Bland–Altman plot/limits of agreement analysis, intraclass correlation coefficient, Lin's concordance correlation coefficient, and British standards reproducibility/repeatability coefficients were reported [5]. Studies meeting the standard but that did not account for

repeat measurements per subject [6] scored a rating of 2. Studies using other acceptable statistical methods, or a mixture of standard and inappropriate methods scored a rating of 1. Studies that used inappropriate statistical methods scored a rating of 0. The studies were given a score out of 3 for both population/set-ups reporting and statistical analysis. We gave an overall study reporting quality rating based on the aggregate number of points for each study, with a maximal score of 6. Studies with scores of 5 or 6 were assessed as 'Good', those with scores of 3 or 4 were 'Fair', and those with scores of 2 or less were

‘Weak’. We contacted authors via emails to inquire about missing data but did not receive any responses.

2.6 Statistical analysis

We used the comprehensive meta-analysis software, R program, and the “meta” package for our analysis. The continuous data are presented as the mean \pm standard deviation and discrete data are presented as a percentage. To determine the agreement between non-contact vital sign devices and vital signs, we conducted a meta-analysis on the final list of Pearson Correlation coefficient values using the Hunter-Schmidt method [7]. All vital signs were assessed separately and compared with other studies that used the same for the reference device in their comparison. A random-effect meta-analysis was used to pool the mean difference in the Bland–Altman and standard error meta-analyses. To assess the variance in outcomes of non-contact devices in comparison to standard medical grade instruments, we conducted a meta-analysis on Bland–Altman values from different studies [8]. To explore the heterogeneity and pooled estimate, we used “one study removed meta-analysis” and “cumulative meta-analysis” to further confirm our results. Additionally, meta-regression analysis was done, with age, male gender, year of publication, and sample size as the confounders. Publication bias was assessed by visual inspection of the funnel plot, Egger’s test, and Begg’s test. Two-tailed p value less than 0.05 was considered significant.

3 Results

3.1 Demographics and characteristics of studies

The demographics and characteristics of the studies are shown in Table 1. Participants’ age in all studies ranged from 18 to 89 years old. Ethnicity was reported in 65% of studies, which included participants from all skin colour categories of the Fitzpatrick scale (Table 1) [9–25]. Most studies had a small sample size (≤ 30 participants) and were done in a laboratory setting. Twenty-three studies (88%) reported testing their technology in healthy volunteers, two studies tested their technology in a patient population with cardiovascular disease [14, 26] and one study did not report their participant population [27].

4 Reference device

For vital sign measurements, twenty-two studies measured HR (92%), three measured BP (12%), and three measured RR (12%). No study examined SpO₂. The pulse oximeter was commonly used (54%) as a reference device. (Table 2)

Electrocardiogram (ECG) (5/26) [11, 14, 24, 26, 28], sphygmomanometer (4/26) [9, 12, 29, 30] HR monitor (1/26) [31], and respiratory monitor such as chest belts (4/26) [13, 17, 18, 32] were also used as reference devices. (Table 2).

4.1 Camera and experimental details

All studies used cameras that recorded in the visible light spectrum. The majority of studies used multiple camera channels [9–20, 22–25, 28–30, 32–35] and two studies used a single-channel [21, 31]. Most studies used a distance of 0.3 m (m) to 2 m between the camera and the participant [9, 11–13, 15–18, 29–31, 34].

Seventy percent of studies used 30 frames per second [9, 11, 13, 14, 16, 20, 21, 23, 25, 28, 31–34] and the remaining used 15 frames per second [15, 17–19, 22, 23, 29]. The resolution of the camera mostly ranged from 320 \times 240 pixels to 1280 \times 720 pixels [11, 13–18, 21–23, 25, 26, 28, 30–34]. Lighting conditions were variable; most used ambient lighting or sunlight [9, 11, 13, 15–20, 22–26, 28–30, 32–34] and three used lighting as an experimental condition [20, 23, 33]. Of the eighteen studies that reported duration of vital sign measurements, most recorded a time between 0.5 to 3 min [9, 12, 13, 15–20, 22–24, 28–35].

4.2 Position of participants

The majority of studies (92%) measured participants at rest in a sitting position, [9, 10, 12, 13, 15–26, 28–35]. Only three studies were done in a standing position, [23, 25, 35] and two in a supine position [14, 35] Five studies measured participants in motion and used motion as an experimental condition (Table 2) [10, 13, 19, 21, 28].

4.3 Quality assessment

Using the adapted Harford et al. 2019 GRRAS quality assessment tool, (Supplementary Table S1) four studies were assessed as ‘Good’, sixteen were assessed as ‘Fair’, and five were deemed ‘Weak’. Most studies did not report their study population and experimental set-up in detail. No study included an *a priori* description of their statistical analysis. The majority of studies used an inappropriate statistical analysis method to measure agreement between the contactless method and the reference device.

4.4 Meta-analysis on HR accuracy

Meta-analysis was performed on HR only as there were only three studies on BP and RR respectively. Sixteen studies on HR were included in the Pearson’s correlation meta-analysis with either direct extraction [10, 13, 15, 17–20, 24, 26, 29, 31] or manual calculation [22, 23, 30, 32, 33]. The

Table 1 Demographics and study reporting

Author, year (Country)	Vital signs	Types of study	Sample size (n)	Gender (%M)	Age (years)	Types of participants	Ethnicity	Setting
Yang, 2020 (China) [9]	BP	Prospective	85	51	61.6 ± 14.1	Volunteers	FP scale 3–4	Lab
Wang, 2020 (China) [10]	HR	Prospective	5	60	NA	NA	Fitzpatrick scale 1–6	Lab
Benedetto, 2019 (Italy) [11]	HR	Prospective	24	54	29.0 ± 2.8	Volunteers	Fitzpatrick scale 1–3	Lab
Luo, 2019 (Canada) [12]	BP	Prospective	1,328	NA	≥ 18	Volunteers	Fitzpatrick scale 4–5	Lab
Gonzalez Viejo C, 2018 (Australia) [29]	HR, BP	Prospective	15	NA	29 ± 5.17	Volunteers	NA	Lab
Rodríguez, 2018 (Spain) [19]	HR	Prospective	15	80	29 ± 3.44	Volunteers	Fitzpatrick scale 1–3	Lab
Sanyal, 2018 (Ireland) [20]	HR, RR	Prospective	25	60	25 ± 2.54	Volunteers	Fitzpatrick scale 1–6	Lab
Tang, 2018 (China) [21]	HR	Prospective	15	67	25 ± 2.87	Volunteers	Fitzpatrick scale 4–5	Lab
Coppetti, 2017 (Switzerland) [26]	HR	Prospective	108	68	64.3 ± 18.0	Patients (heart)	NA	Hospital
Wei, 2017 (China) [32]	HR, RR	Prospective	8	NA	26.5 ± 3.13	Volunteers	NA	Lab
Yan, 2017 (China) [24]	HR	Prospective	40	50	24.7 ± 5.2	Volunteers	Fitzpatrick scale 3–4	Lab
Huang, 2016 (Taiwan) [28]	HR	Prospective	6	67	24.8 ± 2.3	Volunteers	NA	Lab
Koprowski, 2016 (Poland) [33]	HR	Retrospective	10	NA	42.5 ± 8.08	Volunteers	NA	Lab
Cheatham, 2015 (USA) [35]	HR	Prospective	30	60	26 ± 5	Volunteers	NA	Lab
Couderc, 2015 (USA) [14]	HR	Prospective	11	73	65 ± 6	Patients (AF)	Fitzpatrick scale 1–3	Lab
Lin, 2015 (Australia) [34]	HR	Prospective	15	80	27 ± 4.31	Volunteers	NA	Lab
Li, 2014 (Finland) [31]	HR	Prospective	27	44	29.5 ± 5.2	Volunteers	NA	Lab
Xu, 2014 (China) [23]	HR	Prospective	21	NA	NA	Volunteers	Fitzpatrick scale 1–6	Lab
Holton, 2013 (Australia) [15]	HR	Prospective	18	89	41.5 ± 10.97	Volunteers	Fitzpatrick scale 1–6	Lab
Bousefsaf, 2013 (France) [13]	HR	Prospective	12	83	28 ± 16	Volunteers	Fitzpatrick scale 1–4	Lab
Lewadowska, 2012 (Poland) [16]	HR	Prospective	10	80	36.2 ± 15.35	Volunteers	Fitzpatrick scale 1–3	Lab
Sun, 2012 (China) [30]	HR	Prospective	10	100	29.2 ± 8.1	Volunteers	NA	Lab
Tsouri, 2012 (USA) [22]	HR	Prospective	45	NA	31.5 ± 6.12	Volunteers	Fitzpatrick scale 1–6	Lab
Zhu, 2012 (China) [25]	HR	Prospective	30	NA	22.5 ± 1.22	Volunteers	Fitzpatrick scale 1–6	Lab
Poh, 2011 (USA) [18]	HR, RR	Prospective	12	67	24.5 ± 3.97	Volunteers	Fitzpatrick scale 1–6	Lab
Poh, 2010 (USA) [17]	HR	Prospective	12	83	24.5 ± 3.97	Volunteers	Fitzpatrick scale 1–6	Lab

HR heart rate, RR respiratory rate, BP blood pressure, AF atrial fibrillation

Fitzpatrick scale 1–6—(Group A: Caucasian male (Fitzpatrick scale 1–3); Group B: Caucasian female (Fitzpatrick scale 1–3); Group C: African male (Fitzpatrick scale 6); Group D: African female (Fitzpatrick scale 6); Group E: Indian male (Fitzpatrick scale 4–5))

Table 2 Camera and experiment details

Author, Year	Reference medical device (brand)	Camera details, manufacturer	No. channels	Light source	Position of participant	Distance (m)	Duration per subject (min)	Resolution (pixels)	Frame rate (fps)
Yang, 2020 [9]	Sphyg (mercury, Jiangsu Yuyue)	Smartphone (iphone 6 plus front camera)	Multi	Ambient and LCD light	Sitting	0.5	1	NA	30
Wang, 2020 [10]	Pulse oximeter (Heal Force)	Smartphone (P20), Huawei	Multi	NA	Sitting, moving	NA	NA	NA	NA
Benedetto, 2019	ECG (procomp Infiniti T7500M)	Webcam (Logitech HD Pro C920–1080 HD), Logitech	Multi	Ambient and LCD light LCD screen	Sitting	0.6	20	1280×720	30
Luo, 2019 [12]	Sphyg (CNAP® Monitor 500/ Biopac)	Smartphone (iphone 6 Plus front camera), Apple	Multi	LED	Sitting	0.4–0.6	2	NA	NA
Gonzalez Viejo C, 2018 [29]	Sphyg (OMRON HEM-790IT), finger pulse (Arduino®)	Tablet (Android) Google	Multi	White lighting	Sitting, moving	0.3–0.45	1.5–3	1920×1080	15
Rodríguez, 2018 [19]	Pulse oximeter (MF36 BIOPAC)	Webcam (HD Pro C910), Logitech	Multi	Ambient	Sitting	0.3	1	1920×1080	15
Sanyal, 2018 [20]	Pulse oximeter (Biosync B-50DL)	Smartphone (G2), LG	Multi	Ambient 2and with flash	Sitting	0.5	<1 (40 s)	NA	30
Tang, 2018 [21]	HR monitor	Webcam (USB-30W04MTRL21 Model), ELP	Single	Fluorescent lamp	Sitting, Moving	2	1	640×480	30
Coppetti, 2017 [26]	Pulse oximeter (Dräger Infinity Delta XL; KG), ECG (Philips Intellivue X2; Koninklijke Philips NV.)	Smartphone (iphone 4/ iphone 5), Apple	NA	Ambient	Sitting	1.5	NA	(1) iphone 4-640×4802) iphone 5–1280×960	NA
Wei, 2017 [32]	Pulse oximeter (HKG-07A Hefei Huake Info Tech Co) Respiratory sensor (HKH-11B breathing apparatus)	Smartphone (iphone 4 s), Apple	Multi	Ambient	Sitting	NA	4–6	640×480	30
Yan, 2017 [24]	ECG (12 lead, (GE Series 2000,))	Smartphone (iphone 6 s) Apple	Multi	Ambient	Sitting	0.3	<1 (40 s)	NA	NA

Table 2 (continued)

Author, Year	Reference medical device (brand)	Camera details, manufacturer	No. channels	Light source	Position of participant	Distance (m)	Duration per subject (min)	Resolution (pixels)	Frame rate (fps)
Huang, 2016 [28]	ECG (R1 Blue Comfortex +, Sigma sport)	Smartphone (Xperia Z1), Sony	Multi	Ambient	Sitting, Moving	NA	1	640×480	30
Koprowski, 2016 [33]	Pulse oximeter (PO 80, Beurer)	1) webcam (CI170), Logitech 2) webcam (CAM69U), Gembird	Multi	Ambient, artificial (LED, filament & fluorescent bulb)	Sitting	NA	4	640×480	30
Cheatham, 2015 [35]	Pulse oximeter (Model CMS-50E Contec) Polar H7 Belt monitor (wear-link wireless tech)	Smartphone (ipod Touch 5th gen), Apple	NA	NA	Supine, sitting, standing	NA	5	NA	NA
Couderc, 2015 [14]	ECG (12 lead, Mortara H12p recorder, Mortara Instrument, Milwaukee, WI)	Webcam (RGB lifecam Cinema), Microsoft	Multi	NA	Supine	1	NA	1280×720	30
Lin, 2015 [34]	Pulse oximeter (Model CMS60D, Contec)	Webcam (HD-5000), Microsoft	Multi	Ambient light	Sitting	0.5	3	640×480	30
Li, 2014 [31]	HR monitor (Polar S810 HR)	Tablet (ipad), Apple	Single	Indoor and computer screen light	Sitting	0.35	<1 (30 s)	780×580	30
Xu, 2014 [23]	Pulse oximeter (SM-110, Santa-medical)	(1) smartphone (Android TM 2.3), Huawei U8652 (2) smartphone (iphone 4), Apple	Multi	Ambient light, sunlight (outdoor)	Sitting, Standing	NA	1.5	(1)800×480 (2)1280×720	(1)15 (2)30
Holton, 2013 [15]	Pulse oximeter (Rossmax SB220)	Webcam (C910 HD Pro), Logitech	Multi	Fluorescent light, daylight	Sitting	0.6	1.5	640×480	15
Bousefsaf, 2013 [13]	PPG (BVP-Flex/Pro, Thought Tech Ltd.), chest belt respiration sensor (Resp-Flex/Pro, Thought Tech)	Webcam (Lifecam Cinema), Microsoft	Multi	Sunlight	Sitting, moving	1	<1 (35 s)	320×240	30

Table 2 (continued)

Author, Year	Reference medical device (brand)	Camera details, manufacturer	No. channels	Light source	Position of participant	Distance (m)	Duration per subject (min)	Resolution (pixels)	Frame rate (fps)
Lewadowska, 2012 [16]	ECG (ascard imgrey v.201, Aspel) Thermal cam (FLIR Therma Cam SC3000)	Webcam (9000 pro), Logitech	Multi	Daylight, incandescent light, fluorescent light	Sitting	0.5	< 1 (30 s)	640×480	30
Sun, 2012 [30]	Pulse oximeter (Model: P871RA) Sphyg (Model: M6, Omron)	Webcam (9000 Pro), Logitech	Multi	Ambient	Sitting	0.2	3	320×240	NA
Tsouri, 2012 [22]	Pulse oximeter	Webcam, Logitech	Multi	Ambient	Sitting	NA	1	320×240	15
Zhu, 2012 [25]	Pulse oximeter	Webcam	Multi	Daylighting or fluorescent lighting	Sitting/ standing	NA	NA	640×480	30
Poh, 2011 [18]	PPG and chest belt respiration sensor (Flexcomp Infniti, Thought Tech)	Webcam (isight camera), Apple	Multi	Ambient	Sitting	0.5	1	640×480	15
Poh, 2010 [17]	PPG and chest belt respiration sensor (Flexcomp Infniti, Thought Tech)	Webcam (isight camera), Apple	Multi	Ambient	Sitting	0.5	1	640×480	15

Specifications for only the non-contact camera used in each study were recorded
Sphyg sphygmomanometer, Tech Technologies

random-effects meta-analysis on Pearson’s values (Fig. 2) showed that the pooled weighted correlation coefficient was 0.962 (95% CI 0.905 to 0.985; p value < 0.00) (Predictive Intervals: -0.0839 to 0.99). The overall inter-study heterogeneity (I^2) is 93%, while the between-study variance is (τ^2) 0.873 ($\tau = 0.934$). The weights of all studies were evenly spread despite overall low participant counts (< 50). To explore the impact/influence of one study on the pooled correlation coefficient and study’s effect on the heterogeneity, we conducted the “one study removed meta-analysis” (Supplementary Figure S1). With this analysis, the pooled correlation coefficient varied from highest value of 0.96 (after removal of Coppetti et al. [26]) to lowest value of 0.95 (after removal of Yan et al. [24]). After removal of Coppetti et al. [26] and Yan et al. [24], heterogeneity decreased to 79%, suggesting that these two studies contributed the maximum heterogeneity to the pooled correlation coefficient (Supplementary Table S2). The results were further confirmed by cumulative meta-analysis, which did not change the final inference of our results. We identified age, male gender, year of publication, and sample size as the confounders and conducted the meta-regression analysis. Meta-regression by age, male gender, year of publication, and sample size showed that the results were similar over these confounders since the slope of effect size was non-significantly altered (Supplementary Table S3, Figure S4). On visualization of the Funnel plot there was no publication

bias (Supplementary Figure S3). The absence of publication bias was further confirmed by Egger’s test ($p = 0.595$) and Begg’s test ($p = 0.458$).

Fourteen studies were included in the Bland–Altman standard error of mean meta-analysis [10, 11, 13, 15, 17, 19, 20, 22, 24, 26, 30, 32, 33, 35]. The pooled estimate of the mean difference between non-contact and standard method for HR detection was 0.36, with the pooled 95% confidence interval ranging from -1.22 to 1.95 (Fig. 3).

4.5 Qualitative analysis on BP and RR accuracy

Three studies examined the accuracy of contactless BP monitors [9, 12, 29]. Luo et al. recorded average prediction bias \pm error SDs of 0.39 ± 7.30 mmHg for systolic BP and -0.20 ± 6.00 mmHg for diastolic BP compared to a sphygmomanometer (CNAP® Monitor 500/Biopac) [12]. Yang et al. recorded -0.4 ± 6.7 mmHg for systolic BP and 1.2 ± 7.0 mmHg for diastolic BP compared to a mercury sphygmomanometer [9]. Both studies validated the ‘Anura’ application and found that the contactless monitor falls within the key accuracy threshold of 5 ± 8 mmHg set by the Association for the Advancement of Medical Instrumentation (AAMI) standard [9, 12]. Gonzalez et al. indicated that their HR and BP contactless monitor had a high pooled correlation (0.97) compared to a sphygmomanometer (Omron

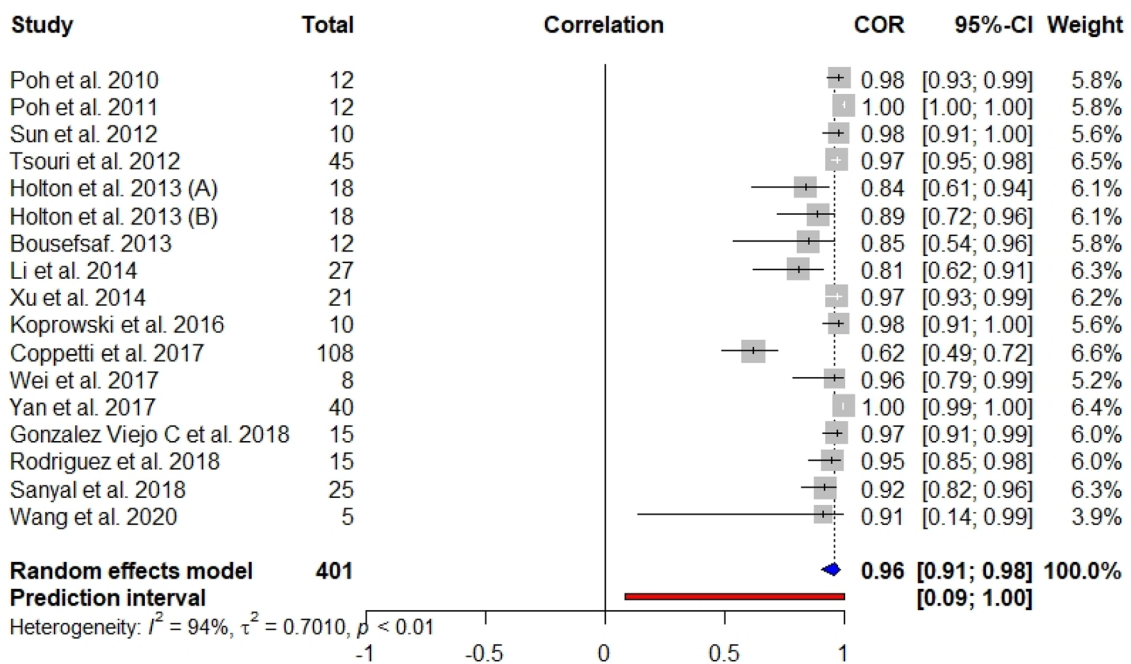


Fig. 2 Forest plot of Pearson’s correlation coefficients of heart rate measurements at rest in contactless vital signs monitor compared to a reference medical device. Forest plot of random-effects meta-analysis on Pearson’s values showing pooled weighted correlation coef-

ficient of 0.962 (95% CI 0.905 to 0.985; p value < 0.00) (Predictive Intervals: -0.0839 to 0.99). The overall inter-study heterogeneity (I^2) is 93%, while between-study variance is (τ^2) 0.873 ($\tau = 0.934$)

Bland-Altman Meta-analysis

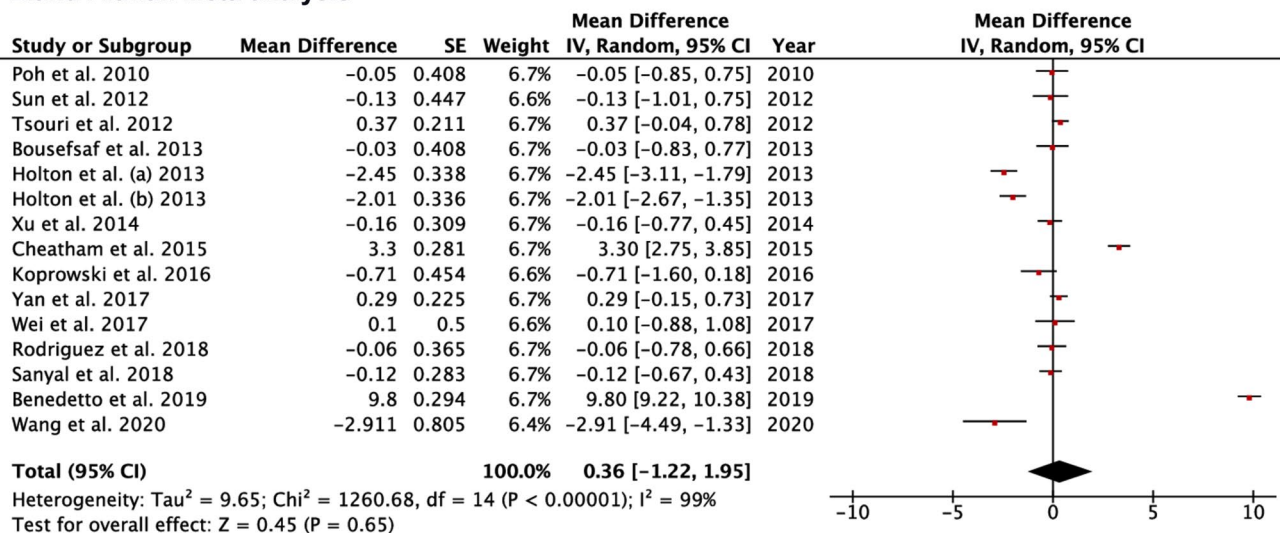


Fig. 3 Forest plot of standard error of the mean of heart rate measurements at rest in contactless vital signs monitor compared to a reference medical device. Forest plot of pooled estimate of the mean

difference between non-contact and standard method for HR detection was 0.36, with the pooled 95% confidence interval ranging from -1.22 to 1.95

Hem-790IT) but did not explicitly provide data for BP measurement alone [29].

Three studies examined the accuracy of contactless RR monitors [18, 20, 32]. Two studies found that RR was less accurate than HR measurement [18, 20] one study found that RR and HR measurements had comparable accuracy [32]. Sanyal et al. recorded a correlation value of 0.66 for RR compared to 0.92 for HR when comparing their contactless RR monitor against self-reported measurement [20]. Poh et al. recorded a correlation value of 0.94 for RR compared to 1.00 for HR when comparing their contactless RR monitor against a chest belt respirator sensor [18]. Wei et al. calculated their correlation value per measurement, in which RR ranged from 0.90–0.98 (mean=0.96, median=0.96) and HR ranged from 0.91–0.98 (mean=0.95, median=0.96), showing comparable accuracy between the contactless method and a breathing apparatus sensor (Model: HKH-11B, Hefei Huake Info Technology Co.) [32].

4.6 Reported limitations in video recordings

Motions that range from head movements, back and forth body movement to light exercises have been reported to affect the accuracy of the contactless method [9–11, 14, 16, 34]. Since only normotensive participants were used to train the predictive models on vital signs, more errors were reported in the lower and higher range of HR and BP measurement for three studies [9, 11, 12]. Two studies reported racial homogeneity as a potential limitation yet found that skin tone did not influence the predictive model's accuracy [9, 12]. As for camera specification, lack of automatic face

tracking [15, 32], low frame rate, and resolution [11, 16, 33] decrease the accuracy of vital sign prediction. Poor lighting [9, 16, 30, 33, 34] and facial obstruction [20] were reported as a limitation for capturing the participant's face.

4.7 Commercially available applications

Our review included contactless vital sign monitors that can be used by the public. At present, four applications are currently available on the Apple App store for consumer use. The 'Anura' application (Nuralogix Inc., Canada) measures BP and was deemed accurate compared to a sphygmomanometer by two studies [9, 12]. The 'Cardio' application (Cardio Inc. Cambridge, USA), was validated in two studies with contradicting results [24, 26]. One study found the application to be accurate using the iPhone 6 with a better camera of higher resolution and frame rate for video recording [24] while another study found inaccuracy using the iPhone 5/4 [26] when compared to ECG for measuring HR. The remaining applications, 'Whats My Heart Rate' (Vitrox Technologies, Malaysia) [26] and 'Cardio Buddy' (Azumio Inc., USA) [35] were reported as not accurate or effective in measuring HR compared to a pulse oximeter.

5 Discussion

To date, this is the first systematic review and meta-analysis in the literature to evaluate the accuracy of consumer-friendly contactless vital sign monitors that utilizes a smartphone or webcam. We found that contactless vital sign

monitors were accurate when compared to standard medical devices in HR measurement. The mean absolute error of the estimated heart rate using the contactless vital sign monitors is 0.36 bpm, while the best performance by the contactless vital sign monitor is 1.95 bpm. The consumer-friendly contactless vital sign monitors showed a high correlation of 0.962 with standard medical devices.

Almost all contactless devices measured HR, while only 3 studies measured BP [9, 12, 29] and RR [18, 20, 32]. Due to the small number of studies on RR and BP, we were not able to perform a meta-analysis on these parameters. Two out of three studies which measured BP using the same application found that the contactless monitor achieved the key accuracy according to the AAMI standard [9, 12]. Two out of three studies which measured RR found that RR contactless monitor had lower accuracy when compared to their respective reference devices [18, 20]. We did not find any study using this new technology on SpO₂. SpO₂ is a vital sign that is essential for medical evaluations and should be added to future contactless monitoring technology. To date, only studies that used high performing, expensive, non-consumer grade cameras have attempted to measure SpO₂ [36, 37].

More than half of the studies used the pulse oximeter as a reference medical device for measuring HR while other studies used ECG [11, 14, 24, 26, 28] or an HR monitor [31]. The pulse oximeter uses PPG (Photo-Plethysmography) technology, which is similar to imaging PPG (iPPG) technology implemented in contactless devices. The similarity between both optical methods supports the pulse oximeter's use as a reference device [38, 39]. Melanin concentration and skin pigmentation can affect PPG technologies in recording pulse rate [40], which raises the importance of testing camera detection in a wide range of skin types. Our review shows that the current contactless technology has been validated in all skin types since 65% of studies included all possible skin types.

Most studies used cameras that captured videos at 30 frames per second, while resolution ranged from 320×240 pixels to 1280×720 pixels. Three studies observed that higher resolution correlates with higher measurement accuracy, and suggested that at minimum a frame rate of 15 fps should be used in all contactless monitors [11, 15, 16]. As camera technology improves with higher specifications, the accuracy of these technologies will also improve as time continues.

The majority of studies used healthy volunteers in a laboratory setting to validate their contactless technology. Only two studies tested contactless technology on patients with atrial fibrillation [14] and heart problems in a hospital setting [26]. Couderc et al. noted that contactless HR measurements using video plethysmography is useful for at-home monitoring of patients with atrial fibrillation, highlighting the utility of this new technology in chronically ill patients

[14]. In order to validate the use of this new technology in clinical settings, studies will need to include patients with comorbidities to assess its performance on a wide range of vital signs. Validation of the contactless technology in clinical settings will be crucial to factor in movement, variable lighting, and patient perceptions on ease of use. To date, the American National Standards Institute (ANSI)/Association for the Advancement of Medical Instrumentation (AAMI)/International Organization for Standardization (ISO) 81,060–2:2013 standard has not been developed for contactless, camera-, vital sign measuring technologies [9], which makes it difficult to deem if currently available technologies are adequate for clinical use. Further work is needed in this area.

A variety of technological barriers still restrict the accuracy of contactless technology and must be resolved before widespread implementation. Motion [9–11, 14, 16, 34], poor lighting [16, 34], and manual face tracking [15, 31, 32] were reported as a limitation for accuracy. Only four contactless vital sign monitors are currently available for consumer use, three of which measures HR [24, 26, 35] and one measures BP [9, 12].

It is important to note that the majority of studies were conducted on healthy participants, and the accuracy of contactless monitors with abnormal vital signs has yet to be determined. To truly assess accuracy, the contactless monitors must be referenced against ground truth measurements, such as ECG for HR and mercury column manometer for BP.

Contactless, camera-based vital sign monitors are accessible and can greatly advance virtual health care. Tools for remote patient home monitoring are relevant and necessary as a way to encourage social distancing measures and as a bedside tool during the COVID-19 pandemic [41]. Effective and widely spread contactless camera-based monitoring can free-up hospital beds for very sick patients by allowing health care providers to monitor moderately sick patients at home and reduce peer-to-peer contact in clinical settings [41, 42]. Compared to other self-monitoring alternatives like smart-watches and automatic blood pressure cuffs, the ubiquity of smartphones and webcams make this technology accessible for the general population. Although existing vital signs monitors (automatic blood pressure cuffs, SpO₂ monitors) range from \$50–100, many individuals do not possess such devices nor have access to them. While novel smartwatches can measure all vital signs, they are expensive (\$200–\$500) and features such as SpO₂ measurement have not been clinically validated for most brands.

Overall, contactless vital sign monitors need further development before widespread implementation and measurement of SpO₂ is essential as part of vital signs. It is crucial that contactless vital sign monitors are tested for accuracy in a large and diverse population before being used in a clinical setting.

6 Limitations

One of the main limitations to the findings of this meta-analysis is the absence of studies which used randomized controlled trials with a large sample size. Although no evidence of bias was noticed, the pooled analysis found high heterogeneity, primarily attributable to differences in the patient population, varying technology, study outliers [10, 18, 26, 35] and possibly due to the use of different algorithms/methods in developing the contactless technology over a period of time. Our meta-analysis may be underpowered for pooled estimates increasing the chances of type II errors. Some of these limitations and other unknown confounding factors which are inherent to the prospective observational studies may have contributed to bias in pooled estimates and its dispersion. Despite these limitations, our meta-regression analysis offers a comprehensive analysis of the available evidence on the accuracy of contactless monitors to measure vitals. In future studies, it will be useful to evaluate the specific technologies for measuring specific vitals over different populations.

7 Recommendations

Studies that aim to validate their contactless technology should use a medical grade reference device; aim to measure a wide range of vital signs, and test their device in a diverse patient population in order to determine the technology's practical application in clinical scenarios. We recommend using the Bland–Altman analysis to test for the agreement between contactless devices and medical devices since this method has been widely used in agreement research [43]. A modified Clark Error Grid analysis can also be used to quantify clinical accuracy and the consequences for clinical decision-making of HR, RR and BP measurements [44]. This analysis has been used in studies to validate the accuracy of wearable vital sign monitors [45, 46]. The use of Pearson's correlation without another statistical test, such as the Bland–Altman analysis, is inappropriate since the conclusions drawn from Pearson's correlation coefficients alone do not create a full picture to showcase the variance of the devices being compared [47]. Future regulations on contactless monitors should cover standardization of reporting of results [48], and guidelines on how to validate these technologies for clinical use, such as a comparison to specific medical devices and running experiments with adequate sample size.

8 Conclusions

This systematic review and meta-analysis showed that contactless vital sign monitors which utilize a consumer-friendly camera, such as a webcam or smartphone camera,

are accurate for measuring HR when compared to a medical device. More studies on contactless BP and RR are needed to assess their accuracy. In order for contactless vital sign monitors to be implemented for clinical use, they must be validated in a larger and more diverse population, and expanded to encompass all important vital signs, including blood oxygen saturation.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10877-021-00734-9>.

Author contribution CP This author designed the study, collected, analyzed, and interpreted the data, and prepared the manuscript. KP This author designed the study, collected, analyzed, and interpreted the data, and prepared the manuscript. MN This author analyzed, interpreted the data, and prepared the manuscript. AS This author collected, interpreted the data, and prepared the manuscript. MP This author prepared the manuscript. ME Design and execution of database search strategies and prepared the manuscript. KL This author prepared the manuscript. FC This author designed the study, interpreted the data, and prepared the manuscript.

Funding Funding was provided by the ResMed Research Chair of Anesthesia, Sleep, and Perioperative Medicine Fund.

Declarations

Conflict of interest Frances Chung: Reports research support from University Health Network Foundation, Up-to-date royalties, consultant to Takeda Pharma, STOP-Bang proprietary to University Health Network. All other authors have no conflicts to disclose.

Research involved in animal and human rights Additional declarations for articles in life science journals that report the results of studies involving humans and/or animals.

References

- Hollander JE, Carr BG. Virtually perfect? Telemedicine for Covid-19. *N Engl J Med*. 2020;382(18):1679–81. <https://doi.org/10.1056/NEJMp2003539>.
- Wosik J, Fudim M, Cameron B, Gellad ZF, Cho A, Phinney D, et al. Telehealth transformation: COVID-19 and the rise of virtual care. *J Am Med Inform Assoc*. 2020;27(6):957–62. <https://doi.org/10.1093/jamia/ocaa067>.
- Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med*. 2009;6(7):e1000097.
- Kottner J, Audigé L, Brorson S, Donner A, Gajewski BJ, Hróbjartsson A, et al. Guidelines for reporting reliability and agreement studies (GRRAS) were proposed. *Int J Nurs Stud*. 2011;48(6):661–71.
- Harford M, Catherall J, Gerry S, Young JD, Watkinson P. Availability and performance of image-based, non-contact methods of monitoring heart rate, blood pressure, respiratory rate, and oxygen saturation: a systematic review. *Physiol Meas*. 2019;40(6):06. <https://doi.org/10.1088/1361-6579/ab1f1d>.
- Bland JM, Altman DG. Agreement between methods of measurement with multiple observations per individual. *J Biopharm Stat*. 2007;17(4):571–82.

7. Hunter JE, Schmidt FL. *Methods of meta-analysis: correcting error and bias in research findings*. Sage; 2004.
8. Tipton E, Shuster J. A framework for the meta-analysis of Bland-Altman studies based on a limits of agreement approach. *Stat Med*. 2017;36(23):3621–35.
9. Yang D, Xiao G, Wei J, Luo H. Preliminary assessment of video-based blood pressure measurement according to ANSI/AAMI/ISO81060-2: 2013 guideline accuracy criteria: anura smartphone app with transdermal optimal imaging technology. *Blood Press Monit*. 2020;25(5):295–8. <https://doi.org/10.1097/MBP.0000000000000467>.
10. Wang D, Yang X, Liu X, Jing J, Fang S. Detail-preserving pulse wave extraction from facial videos using consume-level camera. *Biomed Opt Express*. 2020;11(4):1876–91.
11. Benedetto S, Caldato C, Greenwood DC, Bartoli N, Pensabene V, Actis P. Remote heart rate monitoring-assessment of the facereader rPPg by Noldus. *PLoS ONE*. 2019;14(11):0225592.
12. Luo H, Yang D, Barszczyk A, Vempala N, Wei J, Wu SJ, et al. Smartphone-based blood pressure measurement using transdermal optical imaging technology. *Circ Cardiovas Imaging*. 2019;12(8):008857.
13. Bousefsaf F, Maaoui C, Pruski A. Continuous wavelet filtering on webcam photoplethysmographic signals to remotely assess the instantaneous heart rate. *Biomed Signal Process Control*. 2013;8(6):568–74.
14. Couderc J-P, Kyal S, Mestha LK, Xu B, Peterson DR, Xia X, et al. Detection of atrial fibrillation using contactless facial video monitoring. *Heart Rhythm*. 2015;12(1):195–201.
15. Holton BD, Mannapperuma K, Lesniewski PJ, Thomas JC. Signal recovery in imaging photoplethysmography. *Physiol Meas*. 2013;34(11):1499.
16. Lewandowska M, Nowak J. Measuring pulse rate with a webcam. *J Med Imaging Health Informat*. 2012;2(1):87–92.
17. Poh M-Z, McDuff DJ, Picard RW. Non-contact, automated cardiac pulse measurements using video imaging and blind source separation. *Opt Express*. 2010;18(10):10762–74.
18. Poh M-Z, McDuff DJ, Picard RW. Advancements in noncontact, multiparameter physiological measurements using a webcam. *IEEE Trans Biomed Eng*. 2010;58(1):7–11.
19. Rodríguez AM, Ramos-Castro J. Video pulse rate variability analysis in stationary and motion conditions. *Biomed Eng Online*. 2018;17(1):1–26.
20. Sanyal S, Nundy KK. Algorithms for monitoring heart rate and respiratory rate from the video of a user's face. *IEEE Journal of translational engineering in health and medicine*. 2018;6:1–11.
21. Tang C, Lu J, Liu J. Non-contact heart rate monitoring by combining convolutional neural network skin detection and remote photoplethysmography via a low-cost camera. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops 2018*. p. 1309–15.
22. Tsouri GR, Kyal S, Dianat SA, Mestha LK. Constrained independent component analysis approach to nonobtrusive pulse rate measurements. *J Biomed Optics*. 2012;17(7):077011.
23. Xu S, Sun L, Rohde GK. Robust efficient estimation of heart rate pulse from video. *Biomed Opt Express*. 2014;5(4):1124–35.
24. Yan BP, Chan CK, Li CK, To OT, Lai WH, Tse G, et al. Resting and postexercise heart rate detection from fingertip and facial photoplethysmography using a smartphone camera: a validation study. *JMIR Mhealth Uhealth*. 2017;5(3):e33.
25. Zhu H, Zhao Y, Dong L. Non-contact detection of cardiac rate based on visible light imaging device. *Optics and Photonics for Information Processing VI: International Society for Optics and Photonics*; 2012. p. 849806.
26. Coppetti T, Brauchlin A, Müggler S, Attinger-Toller A, Templin C, Schönrrath F, et al. Accuracy of smartphone apps for heart rate measurement. *Eur J Prev Cardiol*. 2017;24(12):1287–93.
27. Wan X, Wang W, Liu J, Tong T. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. *BMC Med Res Methodol*. 2014;14(1):1–13.
28. Huang R-Y, Dung L-R. Measurement of heart rate variability using off-the-shelf smart phones. *Biomed Eng Online*. 2016;15(1):1–16.
29. Gonzalez Viejo C, Fuentes S, Torrico DD, Dunshea FR. Non-contact heart rate and blood pressure estimations from video analysis and machine learning modelling applied to food sensory responses: A case study for chocolate. *Sensors*. 2018;18(6):1802.
30. Sun Y, Azorin-Peris V, Kalawsky R, Hu S, Papin C, Greenwald SE. Use of ambient light in remote photoplethysmographic systems: comparison between a high-performance camera and a low-cost webcam. *J Biomed Optics*. 2012;17(3):037005.
31. Li X, Chen J, Zhao G, Pietikainen M (2014) Remote heart rate measurement from face videos under realistic situations. *Proceedings of the IEEE conference on computer vision and pattern recognition* 4264–71.
32. Wei B, He X, Zhang C, Wu X. Non-contact, synchronous dynamic measurement of respiratory rate and heart rate based on dual sensitive regions. *Biomed Eng Online*. 2017;16(1):1–21.
33. Koprowski R. Blood pulsation measurement using cameras operating in visible light: limitations. *Biomed Eng Online*. 2016;15(1):1–15.
34. Lin J, Rozado D, Duenser A. Improving video based heart rate monitoring. *Stud Health Technol Inform*. 2015;214:146–51.
35. Cheatham SW, Kolber MJ, Ernst MP. Concurrent validity of resting pulse-rate measurements: a comparison of 2 smartphone applications, the polar H7 belt monitor, and a pulse oximeter with bluetooth. *J Sport Rehabil*. 2015;24(2):171–8.
36. Humphreys K, Ward T, Markham C. Noncontact simultaneous dual wavelength photoplethysmography a further step toward noncontact pulse oximetry. *Rev Sci Ins*. 2007;78(4):044304.
37. Wieringa FP, Mastik F, van der Steen AF. Contactless multiple wavelength photoplethysmographic imaging: A first step toward “SpO2 camera” technology. *Ann Biomed Eng*. 2005;33(8):1034–41.
38. Jubran A. Pulse oximetry. *Intensive Care Med*. 2004;30(11):2017–20.
39. Kamshilin AA, Nippolainen E, Sidorov IS, Vasilev PV, Erofeev NP, Podolian NP, et al. A new look at the essence of the imaging photoplethysmography. *Sci Rep*. 2015;5(1):1–9.
40. Fallow BA, Tarumi T, Tanaka H. Influence of skin type and wavelength on light wave reflectance. *J Clin Monit Comput*. 2013;27(3):313–7.
41. Abbo AR, Miller A, Gazit T, Savir Y, Caspi O. Technological Developments and strategic management for overcoming the COVID-19 challenge within the hospital setting in Israel. *Rambam Maimonides Med J*. 2020;11(3):0026.
42. Bokolo AJ. Exploring the adoption of telemedicine and virtual software for care of outpatients during and after COVID-19 pandemic. *Ir J Med Sci*. 2021;190(1):1–10. <https://doi.org/10.1007/s11845-020-02299-z>.
43. Zaki R, Bulgiba A, Ismail R, Ismail NA. Statistical methods used to test for agreement of medical instruments measuring continuous variables in method comparison studies: a systematic review. *PLoS ONE*. 2012;7(5):e37908.
44. Clarke WL, Cox D, Gonder-Frederick LA, Carter W, Pohl SL. Evaluating clinical accuracy of systems for self-monitoring of blood glucose. *Diabetes Care*. 1987;10(5):622–8. <https://doi.org/10.2337/diacare.10.5.622>.
45. Breteler MJM, KleinJan EJ, Dohmen DAJ, Leenen LPH, van Hillegersberg R, Ruurda JP, et al. Vital signs monitoring with wearable sensors in high-risk surgical patients: a clinical validation study. *Anesthesiology*. 2020;132(3):424–39. <https://doi.org/10.1097/aln.0000000000003029>.
46. Breteler MJMM, Huizinga E, van Loon K, Leenen LPH, Dohmen DAJ, Kalkman CJ, et al. Reliability of wireless monitoring using a wearable patch sensor in high-risk surgical patients at a step-down unit in the Netherlands: a clinical validation study.

- BMJ Open. 2018;8(2):e020162. <https://doi.org/10.1136/bmjopen-2017-020162>.
47. Van Stralen K, Dekker F, Zoccali C, Jager K. Measuring agreement, more complicated than it seems. *Nephron Clin Pract*. 2012;120(3):c162–7.
 48. Antink CH, Lyra S, Paul M, Yu X, Leonhardt S. A broader look: camera-based vital sign estimation across the spectrum. *Yearb Med Inform*. 2019;28(1):102.

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