



OPEN ACCESS

Man versus machine: comparison of naked-eye estimation and quantified capillary refill

Rani Toll John,^{1,2} Joakim Henricson,³ Chris D Anderson,^{4,5} Daniel Björk Wilhelms^{2,3}

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/emermed-2018-207948>).

¹Division of Cell Biology, Department of Clinical and Experimental Medicine, Faculty of Health Sciences, Linköping University, Linköping, Östergötland, Sweden

²Department of Emergency Medicine, Local Health Care Services in Central Östergötland, Region Östergötland, Linköping, Sweden

³Division of Drug Research, Department of Medical and Health Sciences, Faculty of Health Sciences, Linköping University, Linköping, Sweden

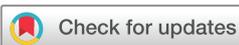
⁴Division of Cell Biology, Department of Clinical and Experimental Medicine, Faculty of Health Sciences, Linköping University, Linköping, Östergötland, Sweden

⁵Department of Dermatology and Venerology, Heart and Medicine Centre, Region Östergötland, Linköping, Sweden

Correspondence to

Dr Rani Toll John, Division of Cell Biology Department of Clinical and Experimental Medicine, Faculty of Health Sciences, Linköping University Linköping Östergötland Sweden ; ranitoll@hotmail.com

Received 6 July 2018
Revised 26 April 2019
Accepted 30 May 2019
Published Online First
15 July 2019



© Author(s) (or their employer(s)) 2019. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

To cite: Toll John R, Henricson J, Anderson CD, *et al.* *Emerg Med J* 2019;**36**:465–471.

ABSTRACT

Background Capillary refill (CR) time is traditionally assessed by 'naked-eye' inspection of the return to original colour of a tissue after blanching pressure. Few studies have addressed intra-observer reliability or used objective quantification techniques to assess time to original colour. This study compares naked-eye assessment with quantified CR (qCR) time using polarisation spectroscopy and examines intra-observer and interobserver agreements in using the naked eye. **Method** A film of 18 CR tests (shown in a random fixed order) performed in healthy adults was assessed by a convenience sample of 14 doctors, 15 nurses and 19 secretaries (Department of Emergency Medicine, Linköping University, September to November 2017), who were asked to estimate the time to return to colour and characterise it as 'fast', 'normal' or 'slow'. The qCR times and corresponding naked-eye time assessments were compared using the Kruskal-Wallis test. Three videos were shown twice without observers' knowledge to measure intra-observer repeatability. Intra-observer categorical assessments were compared using Cohen's Kappa analysis. Interobserver repeatability was measured and depicted with multiple-observer Bland-Altman plotting. Differences in naked-eye estimation between professions were analysed using ANOVA.

Results Naked-eye assessed CR time and qCR time differ substantially, and agreement for the categorical assessments (naked-eye assessment vs qCR classification) was poor (Cohen's kappa 0.27). Bland-Altman intra-observer repeatability ranged from 6% to 60%. Interobserver agreement was low as shown by the Bland-Altman plotting with a 95% limit of agreement with the mean of ± 1.98 s for doctors, ± 1.6 s for nurses and ± 1.75 s for secretaries. The difference in CR time estimation (in seconds) between professions was not significant.

Conclusions Our study suggests that naked-eye-assessed CR time shows poor reproducibility, even by the same observers, and differs from an objective measure of CR time.

INTRODUCTION

The capillary refill (CR) test is used in several globally implemented frameworks for assessment and resuscitation in patients with trauma (eg, ATLS), in paediatric emergency medicine (eg, APLS) and several paediatric triage systems.^{1–3} The test is conducted by pressing a finger on a tissue (most often the skin) to cause blanching, and optically estimating the time to return to original colour.⁴ Naked-eye assessment of the CR time was suggested for the initial evaluation of blood loss in wounded

Key messages

What is already known on this subject

- Capillary refill (CR) is frequently recommended for assessment of patients. However, subjectivity and a lack of standardisation in the execution and interpretation of the traditional CR time test may limit its clinical usefulness.
- Studies concerning agreement between observers or repeatability by the same observer using naked-eye assessment of the same CR test are few.

What this study adds

- In this study using videos, assessment of capillary refill time showed pronounced underestimation and overestimation compared to quantified CR time. Categorical assessment of fast, normal or slow was not more accurate.
- The agreement for the categorical assessment for naked eye was not more reliable.
- There was a low intra-observer repeatability and poor interobserver agreement by clinical staff in their naked-eye assessment of CR time.

soldiers in 1947, and was categorised into 'normal', 'definite slowing' or 'very sluggish'.⁵ In 1981, Champion *et al* arbitrarily defined an upper limit of 2 s as the normal CR time.⁶ Although the 2 s definition of upper limit of normal is most well known, other definitions have been suggested that take into account age and gender, advocating a 2.9 s upper limit for women and a 4.5 s cut-off for the elderly.⁷ In children, a range of 2–3 s has been suggested as normal.^{8–10}

Our understanding of the underlying physiology of CR response is incomplete, and critics claim that the subjectivity and a lack of standardisation in the execution and interpretation of the CR test may limit its clinical usefulness.^{11 12} Studies concerning reproducibility between observers or repeatability by the same observer using naked-eye assessment of repeated CR tests are scarce.^{13 14} There are few studies on CR using objective skin blood content quantification techniques, and these focus on the technique rather than the CR time itself or comparisons with naked-eye estimations.^{15–18} Consequently, both those who believe in the clinical value of the CR time test and those who do not, have essentially only had subjective, naked-eye data to consider when forming their opinion.

The skin's dermal layer includes the red blood cell containing capillaries and sub-papillary plexa,



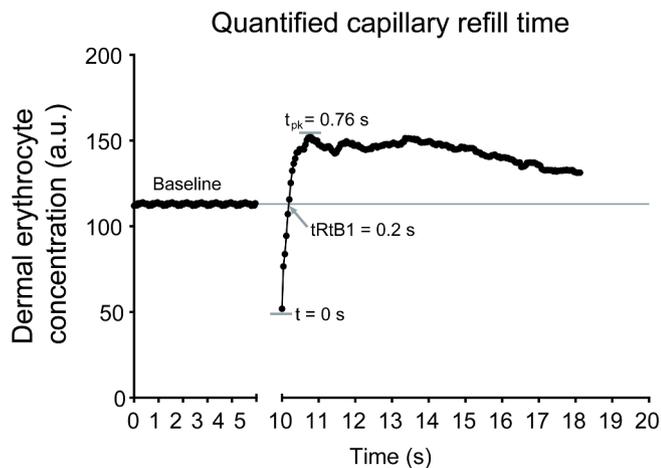


Figure 1 Delineation of a capillary refill curve where baseline represents the redness of skin before application of pressure. $T=0$ is the moment of pressure release during CR test. $tRtB1$ and tPk are shown. Each dot represents an image in which the concentration of erythrocytes has been quantified. CR, capillary refill; tPk , time to peak; $tRtB1$, time to Return to Baseline1.

which contribute to skin colour.¹⁹ Most of the capillaries are emptied when blanching pressure is applied and the refilling, indicated by return of original colour, occurs at various speeds depending on multiple factors, for example, age, skin temperature and sex. Polarisation spectroscopy¹⁶ can detect the change in concentration of red blood cells in the dermal layer, with a resolution of 25 frames per second (0.04s) for the whole of the observed period after blanching pressure. In analysis of this data, we have previously suggested an objective quantified capillary refill (qCR) time endpoint called time to return to baseline 1 ($tRtB1$) as an equivalent for the skin to 'regain its original colour'.^{4 5 15} From the data curve, we can also derive time to peak erythema (tPk), physiologically corresponding to the reactive hyperaemia observed in many cases after the return to baseline (figure 1).^{16 20 21} In this study, we asked participants to estimate time for return to original colour, but we do not know which of our measured values ($tRtB1$ or tPk) might be closest to their 'naked-eye' estimates.

The objectives of this study were thus to compare naked-eye assessment with the qCR time expressed as $tRtB1$ and tPk (in seconds); to measure the intra-observer repeatability and interobserver agreement in estimations of CR time among medical and non-medical staff and to investigate the agreement between 'man' and 'machine' in categorical estimations of the CR response.

METHODS

Study design

This was an observational study of agreement between observer and machine-derived estimates of CR time and categorical assessments. We also investigated intra-observer repeatability and interobserver agreement of CR time estimates, as well as potential differences between professions. The study was approved by the regional ethical review board of Linköping, Sweden, and all volunteers and observers gave written and oral consent prior to participation.

Study setting and selection of observers

The study was carried out at the ED of Linköping University Hospital, Sweden. Data were collected between September and

November 2017. Visual assessments were conducted by a convenience sample of 14 ED doctors (12 ± 8 years of experience), 15 ED nurses (12 ± 9 years of experience) and 19 secretaries (representing laymen). All observers watched a film of 18 CR videos in a set random order (in relation to CR time). Videos with data sets of quantified CR times ($tRtB1$ and tPk) were selected from our previously published material of healthy subjects.⁴ All tests had been performed on the finger pulp on 15 healthy volunteers (7 men) without medication, except for contraceptives. The calculated values of $tRtB1$ for the tests ranged from 0.16 to 9.76s and tPk ranged from 0.68 to 10.36s. The wide range of $tRtB1$ and tPk values among the volunteers was caused by differences in finger skin temperatures at the time of recording (range $30.2^\circ\text{C} \pm 3.5^\circ\text{C}$), as previously described.²²

Randomisation was performed by a person who did not participate in the assessments and did not have any knowledge about the contents of the videos. Each film was shown to the observers on the same computer screen in a windowless room with consistent lightning.

Study protocol

Observers watched each video once and estimated both CR time in seconds and categorical assessment as 'fast', 'normal' or 'slow'. The terms fast, normal or slow were chosen after consideration of different definitions of deranged capillary refill times in the literature.^{5 7 23} The time frame in which previous studies define normal lies within 2s up to 4.5s.^{4 7 24} In our estimation of correct classifications, quantitative cut-offs for the categorical evaluations of fast were defined as <2 s, normal as 2–3s and slow as >3 s. Observers had 5s to deliver their estimates after viewing each video. No information was given to the doctors or nurses about guidelines and reference values of CR time prior to the assessment. Secretaries were assumed to have less familiarity with the CR test, and were instructed to estimate time (in seconds) to when the colour of the blanched area had returned to the same colour as before the applied pressure and to give their categorical estimation of fast, normal and slow without information about our cut-off limits. All observers were shown two CR tests videos as practice prior to the tests included in the analysis.

The participants' estimates were then assessed against measurements of the $tRtB1$ and tPk . The $tRtB1$ and tPk values calculated from the videos ranged from 0.16 to 9.76s and 0.68 to 10.36s, respectively. One slow video ($tRtB1$ 7.28s and tPk 10.2s, videos 3 and 15), one normal video ($tRtB1$ 2.40s and tPk 7.68s, videos 6 and 13), and one fast video ($tRtB1$ 0.36s and tPk 3.36s, videos 5 and 10) were each shown twice without the observers' knowledge to test intra-observer repeatability.

Data analysis

GraphPad Prism V.7.04 was used for statistical analysis. A Kruskal-Wallis test was conducted to compare naked-eye estimations and $tRtB1/tPk$ values. In doing this, $tRtB1$ and tPk were considered to have been conducted an equal number of times to the number of naked-eye time estimates for each profession. Cohen's Kappa was used to compare the categorical assessments made by the observers to the classifications based on $tRtB1$. Differences between professions in CR time estimation was tested with analysis of variance (ANOVA). A p value <0.05 was regarded as statistically significant.

To calculate interobserver agreement, a modification of the Bland-Altman plot,²⁵ multiple observer Bland-Altman plot, was used.²⁶ This method retains the capacity to evaluate consistency

of agreement over different magnitudes of continuous measurements using a single plot. The limits of agreement with the mean represent how different an individual observer estimate compares with the mean measurement of all observers. The differences between each observer and the overall mean for each of the 18 videos were calculated according to the profession of a given participant. Systematic differences between observers were investigated using ANOVA, calculating mean square residuals by profession prior to constructing the plot.

RESULTS

We recruited 51 observers, and complete data sets were available from 48 observers (14 doctors, 15 nurses and 19 secretaries). Of the three excluded observers, one secretary had misunderstood how to estimate the CR time, and two observers (one doctor and one nurse) failed to fill in the CR time in seconds and only completed the categorical evaluation.

Accuracy of naked-eye estimation compared to qCR time

Most observers assessed CR time in whole seconds rather than half seconds. There was a significant difference ($p < 0.05$) between naked-eye assessment (in seconds) and qCR time expressed as tRtB1 for all videos, except for two (secretaries: video 13; doctors: video 16). The significant difference between naked eye and tRtB1 in video 6 was not seen for the identical video 13. Figure 2A-C allows a comparison at an overview level of the naked-eye assessment (in seconds) and the qCR values.

Naked-eye assessments in general overestimated short tRtB1 values (< 1 s) and underestimated prolonged tRtB1 values (> 5 s). The naked-eye time deviated more from tPk than tRtB1, with the deviation being even more pronounced with prolonged CR times. Specifically, there was a significant difference ($p < 0.05$) between naked-eye assessment (in seconds) and qCR time expressed as tPk for all videos, except for two: for video 8, estimates of secretaries and a nurse did not differ significantly, and for video 11, there was no significant difference between tPk and observer estimates for all professions. The data support our original suggestion that tRtB1 is the qCR point with closest correlation to the naked eye estimations, and we therefore used only tRtB1 in our analysis of the categorical estimates.

Categorical assessments

Using our categorical definitions, the interobserver agreement of participants' assessment of fast, normal or slow was 51% (Cohen's kappa 0.27), corresponding to a poor agreement according to Fleiss' Kappa Benchmark Scale.²⁷

A wide intra-observer and interobserver variability in naked-eye categorical classifications compared with categorical classifications based on tRtB1 was seen (table 1). Percent of correct categorical estimates (fast, normal or slow) compared with qCR time classification are highlighted by green colour in table 1. There was a low consistency in naked-eye categorical evaluations for identical videos. For example, 57% of the doctors concluded fast in film 5 and 79% assessed it as fast the second time in the identical film 10.

Intra-observer repeatability

One slow CR test (tRtB1 7.28 s, videos 3 and 15), one 'normal' CR test (tRtB1 2.4 s, videos 6 and 13), and one 'fast' CR test (tRtB1 0.36 s, videos 5 and 10) were each shown twice without the observers' knowledge.

The intra-observer repeatability in the naked-eye time estimation was highly variable. The proportion of identical

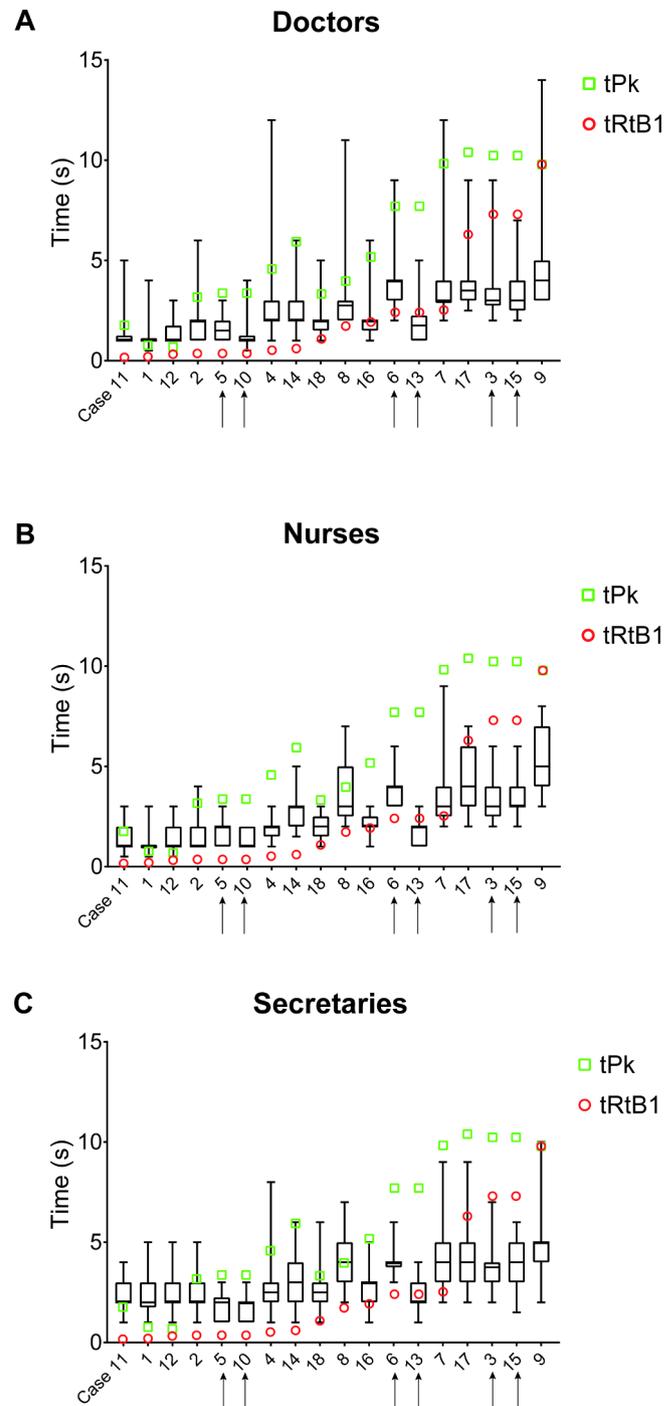


Figure 2 (A-C) Box plots of the naked-eye estimation of CR time in seconds, including the 5th to 95th percentile of the values. qCR time in tRtB1 (red circle) and tPk (green square) are shown for each video. The videos with the fastest tRtB1 values are plotted to the left and the slowest to the right on the x-axis. The number of each video indicates the order in which they were shown to the observers in the film. Videos 5 and 10, 6 and 13, and 3 and 15 (marked with arrows) are the videos that were shown twice. CR, capillary refill; qCR, quantified CR; tPk, time to peak; tRtB1, time to Return to Baseline 1.

time estimates for the repeated tests ranged from 6% to 60% (table 2). Differences in estimations of the same video of up to 3 s were seen. The observers' second estimations differed from the first without apparent pattern, with the exception of the normal

Table 1 Categorical evaluation of the 18 videos with fast, normal or slow tRTB1 times, including the three videos shown twice

Video 11 Fast (0.16 s)			Video 1 Fast (0.2 s)			Video 12 Fast (0.32 s)			Video 2 Fast (0.36 s)			Video 5 Fast (0.36 s)			Video 10 Fast (0.36 s)			Dr (%)		
Nurse (%)	Dr (%)	Sec (%)	Nurse (%)	Dr (%)	Sec (%)	Nurse (%)	Dr (%)	Sec (%)	Nurse (%)	Dr (%)	Sec (%)	Nurse (%)	Dr (%)	Sec (%)	Nurse (%)	Dr (%)	Sec (%)	Fast	Normal	Slow
87	72	86	99	61	71	80	61	23	53	56	57	53	89	79	80	89	89	Normal	7	0
13	28	14	7	33	29	20	39	79	40	44	43	40	11	21	13	11	11	Slow	0	0
0	0	0	0	6	0	0	0	0	7	0	0	0	0	0	0	0	0			
Video 4 Fast (0.5 s)			Video 14 Fast (0.6 s)			Video 18 Fast (1.08 s)			Video 8 Fast (1.72 s)			Video 16 Fast (1.92 s)			Video 6 Normal (2.4 s)			Dr (%)		
Dr (%)	Nurse (%)	Sec (%)	Dr (%)	Nurse (%)	Sec (%)	Dr (%)	Nurse (%)	Sec (%)	Dr (%)	Nurse (%)	Sec (%)	Dr (%)	Nurse (%)	Sec (%)	Dr (%)	Nurse (%)	Sec (%)	Nurse (%)	Sec (%)	Dr (%)
7	47	50	7	0	17	27	22	57	0	0	6	29	20	33	0	0	7	7	7	7
Normal	79	47	39	86	61	67	61	57	71	33	33	71	73	67	21	21	23	23	23	23
Slow	21	0	11	7	20	22	0	14	29	67	61	0	7	0	79	79	80	79	80	80
Video 13 Normal (2.4 s)			Video 7 Normal (2.5 s)			Video 17 Slow (7.28 s)			Video 3 Slow (7.28 s)			Video 15 Slow (7.28 s)			Video 9 Slow (9.76 s)			Dr (%)		
Dr (%)	Nurse (%)	Sec (%)	Dr (%)	Nurse (%)	Sec (%)	Dr (%)	Nurse (%)	Sec (%)	Dr (%)	Nurse (%)	Sec (%)	Dr (%)	Nurse (%)	Sec (%)	Dr (%)	Nurse (%)	Sec (%)	Nurse (%)	Sec (%)	Dr (%)
36	47	56	0	0	0	0	6	7	0	0	17	0	0	0	0	0	0	0	0	0
Normal	64	47	43	47	39	20	11	43	50	60	44	50	27	61	7	7	7	7	7	17
Slow	0	0	0	57	61	60	67	79	50	40	39	73	50	39	93	93	83	83	83	83

The actual tRTB1 time is shown in brackets. Respective answers are reported as percentage of observers, according to profession. The 'correct' (see Methods section) subjective evaluation is highlighted in green. Figures without highlighting indicate deviations from the 'correct' categorical evaluation according to tRTB1 time. Identical tests are shown in videos 5 and 10, 6 and 13 and 3 and 15. The table is organised according to incrementally increasing tRTB1 times. The number of each video corresponds to the place in the sequence in which they were shown to the observers.

tRTB1 , time to Return to Baseline¹

Table 2 Intra-observer repeatability shown as the percentage of answers having identical time estimations

Intra-observer repeatability - Naked-eye time estimation			
tRtB1	0.36 s	2.40 s	7.28 s
Doctors	36%	6%	38%
Nurses	60%	7%	27%
Secretaries	58%	11%	37%

tRtB1, time to return to baseline1.

video, where a majority of observers estimated a faster time for the second observation (see online Supplementary figure S1).

Using categorical estimation, the percentage of intra-observer repeatability (table 3) was slightly better than actual time estimations (table 2). Observers agreed with their prior classification 14%–71% of the repeated videos. Similar to the naked-eye time estimations, the repeatability was lowest in the normal range.

Interobserver agreement

There were no significant differences (in seconds, $p < 0.05$) between professions (ANOVA) in the naked-eye assessments, except for three videos (videos 1, 11 and 12), where a difference between the estimation by secretaries and the other professions was noted.

The 95% limits of agreement of the mean ranged between ± 1.98 s for doctors, ± 1.6 s for nurses and ± 1.75 s for secretaries (figure 3A–C). The doctors had the largest limits of agreement with the mean, but on closer analysis, this was due to two individuals with consistently outlying values. If these individuals were to be excluded, the limits of agreement for the mean of the remaining 12 doctors decreased to ± 0.9 s.

DISCUSSION

In this study, we found a poor correlation between naked-eye assessment of the CR time and qCR time measures in both laymen and clinical staff. Further, we observed poor naked-eye intra-observer repeatability and interobserver agreement by clinical staff in their assessment of CR time. The use of a categorical evaluation of time measurement did not improve agreement between naked-eye estimations and machine-derived classifications.

It is self-evident to most clinicians that different observers, not only in regard to the CR test, often disagree in clinical assessments based on naked-eye observation.²⁴ Previous studies on the reliability of the CR test have partially addressed this by showing a lack of interobserver agreement, but neither performance on the task to actually determine ‘return to normal’ skin colour, nor the intra-observer repeatability for a group of observers on a standardised set of cases has been assessed previously.^{11 28 29}

We have added the use of an objective technique to determine restoration of skin redness and applied this as an external reference for the performance of naked-eye observations of the

Table 3 Intra-observer repeatability shown as the percentage of answers having identical estimations

Intra-observer repeatability - Categorical estimation			
Category	Fast	Normal	Slow
Doctors	64%	14%	71%
Nurses	60%	13%	26%
Secretaries	61%	17%	61%

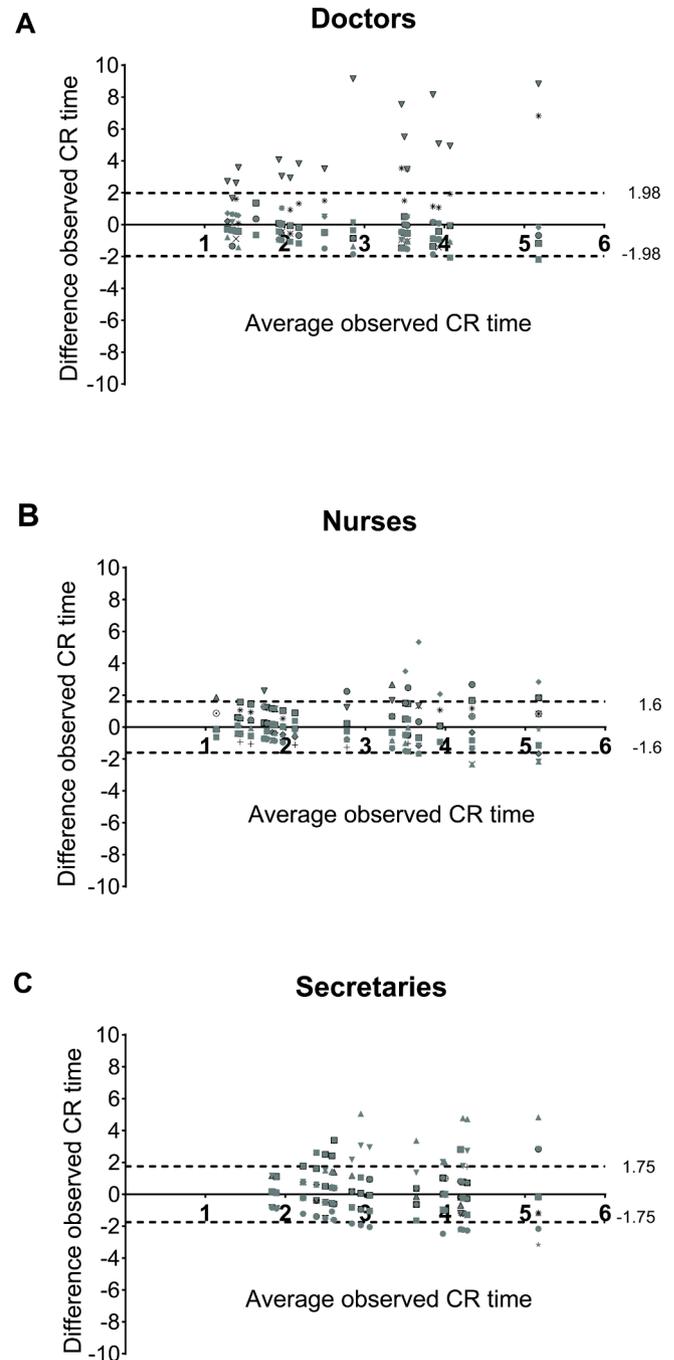


Figure 3 (A–C) Multiple observer Bland-Altman plots for interobserver variability of naked-eye-assessed CR time in seconds shown by profession, with each figure/form representing one observer. The 95% limits of agreement with the mean of all observers are shown as dotted lines. CR, capillary refill.

capillary refill process. Specifically, we have previously suggested tRtB1 as a quantitative endpoint that corresponds to the clinical ‘return to original colour’ with the rationale being that the technique allows determination of the exact point in time at which skin redness rendered by dermal erythrocyte concentration is restored to the same level as prior to the application of blanching pressure.^{15 30 31} Since naked-eye observation may more closely reflect the hyperaemia seen after blanching than the actual return to baseline, we also compared the CR time for naked-eye assessments to the maximal redness achieved (tPk). The tPk measure, however, correlated even less well to ‘return

to normal colour', suggesting that the observers did not try to estimate the same phenomenon (hyperaemia) that was shown by the qCR curve. Arguably, the findings in this study represent differences in sense of time, rather than an inability to assess the visual dynamics of the CR reaction. For this reason, we also asked the observers to categorise the refill responses as fast, normal or slow. These categorical classifications based on naked-eye observations were then compared with classifications based on qCR values, showing a slightly better agreement than time-based estimations. Categorical estimation only improved intra-observer repeatability slightly.

These findings show the difficulty in achieving reproducibility in a seemingly simple visual assessment even among clinical staff in an ED. Further underscoring the difficulty in being precise in naked-eye CR assessment, we found no obvious difference in precision of laymen and healthcare professionals in estimating CR time or performing categorical assessments. No studies have compared naked-eye assessments with an objective method quantifying skin redness. Studies of technical devices measuring CR time on fingers show similar qCR times as for naked-eye assessments according to literature but direct comparisons with naked-eye assessments have not been performed.^{17 32} The lack of consistency both between naked-eye observations and qCR, and between individuals and repeat assessments with the same individuals over time suggests that the naked-eye method for determining CR time estimation, as it is currently performed, is unreliable.

The usefulness of the CR test for the assessment of circulatory status in patients is, indeed, a recurring topic of debate.^{11 12 33 34}

Taken together with previous studies on the questionable reliability of the CR test, the findings of this study indicate a need to revisit the role of CR time estimation in several globally implemented concepts for the assessment and monitoring of emergency patients, such as ATLS, APLS and paediatric early warning score. The findings also underscore the need for more research on the underlying physiology of this test, in order to establish reliable and reproducible methods for measurements that are coupled to relevant outcomes for specific patient groups and situations.

This study has some limitations. At present, there is no technical criterion standard for quantified assessment of the CR test. Polarisation spectroscopy which focuses on red blood cells in the dermis is a possible candidate to quantify return to normal colour in the CR time test by quantifying tRtB1 or other components derived from the qCR curve. Since there are no accepted, validated definitions of return to normal colour and no criterion standard for qCR time, we were unable to check the veracity of our chosen endpoints. The naked-eye assessments were made based on videos, which arguably may differ from some clinical situations. The study was performed in only one centre and the sample size was not large, but the conclusion on the poor performance of naked-eye assessment is still clear. Possible effects of changing the order of the videos were not studied. The videos were not augmented by other clinical information as would be the case in clinical real life, but the objective of this study was to isolate the actual skill involved in CR time estimation. Another limitation was that observers gave the time estimates and categorical estimates at the same time after watching the video, which raises the possibility that the observers' own estimation in seconds may have influenced the categorical estimation rather than being independent. A large proportion of the observers changed naked-eye time estimates and categorical estimates in repeated observations, which indicated lack of consistency with the qCR-determined categories.

In conclusion, the ability of both healthcare professionals and laymen to assess the time for return to normal colour was poor when compared with qCR time measures. Intra-observer repeatability was low, as well as interobserver agreement. CR time measurement should ideally be performed by a quantitative method rather than by naked-eye assessment.

Contributors DBW and RTJ conceived the study. DBW, RTJ and JH designed the trial. DBW obtained research funding. DBW and CDA supervised the conduct of the trial. RTJ and JH undertook recruitment of observers and managed the data with assistance from DBW. DBW and RTJ drafted the manuscript, and all authors contributed substantially to its revision. RTJ takes responsibility for the paper as a whole.

Funding This work was supported by two grants from Region Östergötland to author DBW (LIO-532001 and LIO-700271).

Competing interests RTJ and DBW have no conflicts of interest to declare. The bioengineering method TiVi is marketed by the company WheelsBridge AB. No financial support from WheelsBridge AB was involved in the conduct of the study. JH is employed by the Östergötland County Council but has a royalty agreement with WheelsBridge AB. CDA has a full-time academic position but also limited involvement in WheelsBridge AB.

Patient consent for publication Not required.

Ethics approval This study was reviewed and approved by the Regional Ethical Vetting Board in Linköping, Sweden (permit number M200-07).

Provenance and peer review Not commissioned; externally peer reviewed.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

REFERENCES

- 1 American College of Surgeons CoT. *ATLS advanced trauma life support for doctors student's course manual*. 8th ed: American College of Surgeons, 2008:3–366.
- 2 Advanced Life Support Group. *Oxford*. England: Wiley-Blackwell Publishing Ltd, 2005.
- 3 Group ALS. Circulation assessment. *Acute medical emergencies*. 2nd ed. England: Oxford/Wiley-Blackwell Publishing Ltd, 2010:55–65.
- 4 Pickard A, Karlen W, Ansermino JM. Capillary refill time: is it still a useful clinical sign? *Anesth Analg* 2011;113:120–3.
- 5 Beecher HK, Simeone FA. The internal state of the severely wounded man on entry to the most forward hospital. *Surgery* 1947;22:672–711.
- 6 Champion HR, Sacco WJ, Carnazzo AJ, et al. Trauma score. *Crit Care Med* 1981;9:672–6.
- 7 Schriger DL, Baraff L. Defining normal capillary refill: variation with age, sex, and temperature. *Ann Emerg Med* 1988;17:932–5.
- 8 King D, Morton R, Bevan C. How to use capillary refill time. *Arch Dis Child Educ Pract Ed* 2014;99:111–6.
- 9 Samuels MWS. Structured approach to the seriously ill child. In: *Advanced paediatric life support*. 1st ed. England: Oxford/Wiley-Blackwell Publishing Ltd, 2016:115–26.
- 10 Fleming S, Gill P, Jones C, et al. Validity and reliability of measurement of capillary refill time in children: a systematic review. *Arch Dis Child* 2015;100:239–49.
- 11 Lobos AT, Menon K. A multidisciplinary survey on capillary refill time: Inconsistent performance and interpretation of a common clinical test. *Pediatr Crit Care Med* 2008;9:386–91.
- 12 Sheridan DC, Baker SD, Kayser SA, et al. Variability of Capillary Refill Time among Physician Measurements. *J Emerg Med* 2017;53:e51–e57.
- 13 Espinoza ED, Welsh S, Dubin A. Lack of agreement between different observers and methods in the measurement of capillary refill time in healthy volunteers: an observational study. *Rev Bras Ter Intensiva* 2014;26:269–76.
- 14 Brabrand M, Hosbond S, Folkestad L. Capillary refill time: a study of interobserver reliability among nurses and nurse assistants. *Eur J Emerg Med* 2011;18:46–9.
- 15 John RT, Henricson J, Nilsson GE, et al. Reflectance spectroscopy: to shed new light on the capillary refill test. *J Biophotonics* 2017.
- 16 O'Doherty J, Henricson J, Anderson C, et al. Sub-epidermal imaging using polarized light spectroscopy for assessment of skin microcirculation. *Skin Res Technol* 2007;13:472–84.
- 17 Bordoley A, Irwin R, Kalyani V, et al. *DIGIT: Automated, temperature-calibrated measurement of capillary refill time*, 2012.
- 18 Blaxter LL, Morris DE, Crowe JA, et al. An automated quasi-continuous capillary refill timing device. *Physiol Meas* 2016;37:83–99.
- 19 Krogh A. The anatomy and physiology of the capillaries. *New Haven, CT: Yale University, Press* 1922:304.

- 20 Zhai H, Chan HP, Farahmand S, *et al.* Tissue viability imaging: mapping skin erythema. *Skin Res Technol* 2009;15:14–19.
- 21 Nilsson GE, Zhai H, Chan HP, *et al.* Cutaneous bioengineering instrumentation standardization: the Tissue Viability Imager. *Skin Res Technol* 2009;15:6–13.
- 22 John RT, Henricson J, Nilsson GE, *et al.* Reflectance spectroscopy: to shed new light on the capillary refill test. *J Biophotonics* 2018;11:e201700043.
- 23 Brierley J, Carcillo JA, Choong K, *et al.* Clinical practice parameters for hemodynamic support of pediatric and neonatal septic shock: 2007 update from the American College of Critical Care Medicine. *Crit Care Med* 2009;37:666–88.
- 24 Otieno H, Were E, Ahmed I, *et al.* Are bedside features of shock reproducible between different observers? *Arch Dis Child* 2004;89:977–9.
- 25 Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1:307–10.
- 26 Jones M, Dobson A, O'Brian S. A graphical method for assessing agreement with the mean between multiple observers using continuous measures. *Int J Epidemiol* 2011;40:1308–13.
- 27 Fleiss JL. *Statistical Methods for Rates and Proportions*. 2nd ed. New York: John Wiley, 1981:121–8.
- 28 Klupp NL, Keenan A-M. An evaluation of the reliability and validity of capillary refill time test. *Foot* 2007;17:15–20.
- 29 Eisenhut M. Factors influencing capillary refill time. *J Pediatr* 2007;151:e17.
- 30 Henricson J, Toll John R, Anderson CD, *et al.* Diffuse reflectance spectroscopy: Getting the capillary refill test under one's thumb. *J Vis Exp* 2017;130:e56737–e37.
- 31 John RT, Henricson J, Junker J, *et al.* A cool response-The influence of ambient temperature on capillary refill time. *J Biophotonics* 2018;11:e201700371.
- 32 Kviesis-Kipge E, Curkste E, Spigulis J, *et al.* Real-time analysis of skin capillary-refill processes using blue LED. *ResearchGate* 2010;7715.
- 33 Gorelick MH, Shaw KN, Baker MD. Effect of ambient temperature on capillary refill in healthy children. *Pediatrics* 1993;92:699–702.
- 34 Anderson B, Kelly AM, Kerr D, *et al.* Capillary Refill Time in Adults Has Poor Inter-Observer Agreement. *Hong Kong Journal of Emergency Medicine* 2008;15:71–4.