Hong Kong Journal of Emergency Medicine

Capillary refill time in adults has poor inter-observer agreement

成人的微絲血管再充填時間在觀察員之間的吻合很差

B Anderson, AM Kelly, D Kerr, D Jolley

Objectives: Capillary refill time (CRT) has been taught as a rapid indicator of circulatory status but to be a useful clinical test, CRT needs to be reproducible when performed by another health care worker. No inter-rater agreement studies have been reported for adult patients. The aim of this study was to determine the inter-observer reliability of CRT in a sample of adult emergency department (ED) patients. ***Methods:*** This prospective observational study included clinically stable ED patients with a variety of conditions from two community EDs. A doctor and a nurse each measured CRT by estimation to the nearest half-second using a standard method on each patient. They were blinded to each other's measurements. The primary outcome of interest was inter-rater agreement. Secondary outcome was agreement in classification as normal or abnormal according to accepted definitions. Data was analysed using bias plot analysis, correlation, absolute percent agreement and kappa analysis. ***Results:*** Totally, 209 patients were enrolled; 51% were female and 86% were Caucasian. Median CRT was 2 seconds (95% CI 2-2.35 seconds). The mean difference between measurements by the different observers was 0 second, however the 95% limits of agreement were very wide (-1.7 to +1.9 seconds). Agreement was 70% for classification of 'normal' or 'abnormal' using the 2-second definition of normal, with a kappa of 0.38. ***Conclusion:*** Inter-observer agreement in measurement of CRT was poor in adult subjects with wide limits of agreement. This is a serious threat to the appropriateness of this test for use in clinical practice. (*Hong Kong j.emerg.med.* 2008;15:71-74)

目的：微絲血管再充填時間被教導為循環狀態的一個快速指標，但不是一個有用的臨床測試。微絲血管再充填時間當由另一位醫療工作人員進行時，結果需要是一致的。在成年病者中，並沒有評估員之間的吻合研究報告。本研究旨在確定在一個急症室成年病者樣本中，微絲血管再充填時間在觀察員之間的可靠性。**方法：**這前瞻性觀察研究包括兩間社區急症室，有各種各樣情況但臨床上穩定的病者。一位醫生及一位護士各自為每一位病者量度微絲血管再充填時間，使用標準的方法及評估至最近的半秒。他們不知道對方的量度結果。所關注的主要結果是評估員之間的吻合程度，次要的結果是根據公認的定義分類為正常或不正常的吻合程度。使用偏性標圖分析，相關、絕對百分比吻合及卡巴分析進行數據的分析。**結果：**登記了 209 名病者，51% 為女性及 86% 為白種人。微絲血管再充填時間的中位數為 2 秒（95% 置信區間 2-2.3 秒）。不同觀察員間量度的平均誤差為 0 秒，然而 95% 吻合界限很廣闊（-1.7 至 +1.9

Correspondence to:
Anne-Maree Kelly, MD, FACEM
Sunshine Hospital, Joseph Epstein Centre for Emergency Medicine Research, Furlong Road, St Albans 3021, Australia
Email: Anne-Maree.Kelly@wh.org.au

Debra Kerr, BN, MBL
Bronwyn Anderson, Medical Student

Monash Institute of Health Services Research, Clayton, Australia
Damien Jolley, MSc(Epidemiol), AStat

秒）。使用 2 秒為正常的定義，分類為「正常」或「不正常」的吻合程度為 70% ，而卡巴為 0.38 。**總結**：成年人士量度微絲血管再充填時間的觀察員間吻合程度很差，並有很闊的吻合界限。對在臨床上使用這測試的合適度有嚴重的威脅。

Keywords: Capillaries, observer variation

關鍵詞：微絲血管、觀察員異差

Introduction

Measurement of capillary refill time (CRT) is a simple test that has been advocated for use in assessing the haemodynamic status of patients.[1-3] Despite this, the clinical utility of CRT has been questioned. This is in part due to concerns about inter-observer reliability of CRT measurement which is essential for a test to be a useful tool in routine clinical practice. Paediatric studies have reported the inter-observer reliability of CRT measurement as only fair to moderate.[4-6] To date, the inter-observer reliability of CRT measurement has not been quantified in adults.

The aim of this study was to determine the inter-observer reliability of CRT in a sample of clinically stable, adult emergency department (ED) patients.

Methods

This was a prospective observational cohort study conducted in the ED of Western and Sunshine Hospitals in Melbourne, Australia. Western and Sunshine Hospitals are community teaching hospitals with annual ED census of 32,000 and 70,000 patients respectively.

Patients were eligible for inclusion in the study if they were being treated in either ED between the 1st of May 2006 and the 31st of August 2006, were over 18 years of age and able to provide informed written consent. They were excluded if they were medically unstable, had an altered mental status or were unable to communicate in English and a suitable translator was unavailable.

The CRT of each patient was measured twice, once each by a doctor and a nurse, using a standard technique. Each observer was asked to blanch the participant's right index finger using five seconds of moderate pressure and to count the length of time to the return of normal colour. These were recorded as a value in seconds, with half-second intervals allowed. They were estimated without a timing device, mimicking the clinical scenario in which CRT is routinely assessed. Measurements were taken within 5 minutes of each other, ambient temperature was unchanged between measurements and no therapy was administered between measurements. Observers were blinded as to each other's CRT measurement.

The primary outcome of interest was inter-observer agreement between two clinician observers for measurement of CRT. Secondary outcome was the proportion with agreement in classification as 'normal' or 'abnormal' for each of the 2-second and Schriger definitions[7,8] of normal CRT. While the 2-second definition is the more well-known, the Schriger definitions take into account age and gender and may be more accurate. They advocate a 2.9 second upper limit of normal for women and a 4.5 second cut-off for the elderly.

Data was analysed using bias plot analysis, with mean difference and 95% limits of agreement and Pearson correlation. Agreement in classification was analysed by descriptive statistics and kappa analysis. The study was approved by the relevant institutional research ethics committees.

Results

A total of 209 ED patients were enrolled in this study. The sample contained slightly more females (51%) than males and the majority of participants were Caucasian (85%, n=178). The age and gender

distribution of the study participants is summarised in Figure 1. Median CRT for the sample, based on the average of the clinician's estimates, was 2 seconds (95% CI 2-2.35 seconds).

The mean difference between CRT measurements by the two clinicians was 0 second, with 95% limits of agreement of -1.7 to +1.9 seconds. A bias plot of agreement between the clinicians is shown in Figure 2. The correlation coefficient was 0.47 (95% CI 0.35-0.57).

Using the 2-second definition of the upper limit of normal, the clinicians agreed in their assignment of

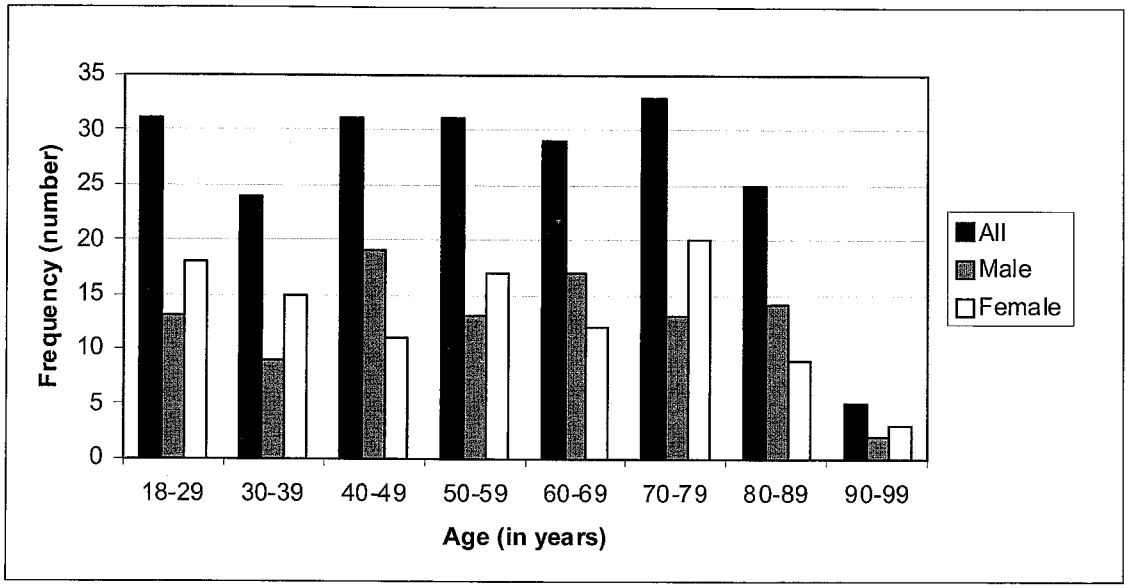

Figure 1. Age and gender distribution of the study population.

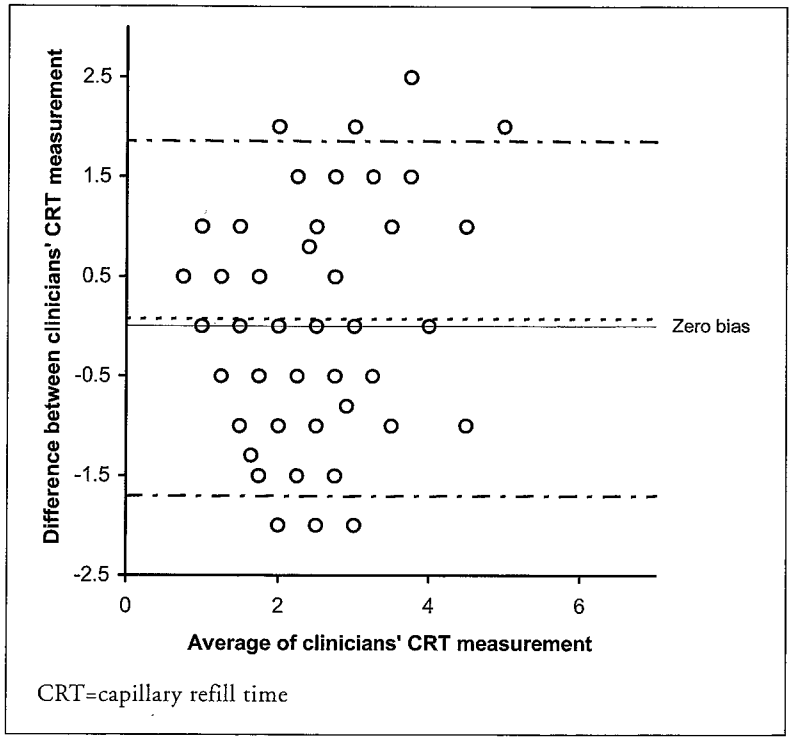

Figure 2. Bias plot of agreement between the clinicians.

'normal'/ 'abnormal' on 70% of occasions with a kappa for agreement of 0.38. Using the Schriger definitions, for the proportion of the sample to which they are applicable, the clinicians agreed in their assignment of 'normal'/ 'abnormal' on 92% of occasions with a kappa for agreement of 0.48.

Discussion

The mean difference in CRT measurements between clinicians was 0 second, however the 95% limits of agreement were -1.7 to +1.9 seconds. These are unacceptably wide in view of the average CRT of 2 seconds. Using the 2-second definition as the upper limit of normal, clinicians agreed in 70% of cases with respect to their classification of normal/abnormal, with a poor kappa value. Using the Schriger's age and gender-based definitions of CRT,[7] agreement in classification of CRT as normal/abnormal was better (92%) with a moderate kappa value.

This is the first study of inter-observer reliability to be performed on adults. Paediatric studies, using the 2-second definition, have reported kappa statistics of 0.49-0.54 and correlation coefficients of 0.55-0.71.[4,5] Our results suggest slightly worse agreement. This may be because our cohort was older and thus more likely to have vascular disease or be taking medications that impact on vascular tone. Taken together, the reported poor inter-observer agreement in CRT, not just in the number recorded but classification as normal or abnormal according to published definitions, raises serious concern about the appropriateness of this test for use in clinical practice.

This study has some limitations that should be considered when interpreting the results. We used a convenience sample of patients. Due to consent issues, our sample was under-represented with patients with more serious medical conditions such as shock. Information on participant diagnosis was not obtained. We used several doctors and nurses in the study. Observers did not receive formal training in CRT measurement, other than the instructions described in Methods. This was an active decision of the research team, to better replicate clinical practice, but differences in technique might have contributed to poor inter-observer reliability of CRT in this study. The order of CRT measurements by observers was also not standardised. Furthermore, international generalisation of these results cannot be assured, as data was only collected on an Australian population.

Conclusion

The inter-observer agreement in measurement of CRT was poor with wide limits of agreement. This is a serious threat to the appropriateness of this test for use in clinical practice.

References

1. Beecher HK, Simeone FA, Burnett CH, Shapiro SL, Sullivan ER, Mallory TB. The internal state of the severely wounded man on entry to the most forward hospital. Surgery 1947;22:672-81.
2. Champion HR, Sacco WJ, Hannan DS, Lepper RL, Atzinger ES, Copes WS, et al. Assessment of injury severity: the triage index. Crit Care Med 1980;8(4):201-8.
3. Champion HR, Sacco WJ, Carnazzo AJ, Copes W, Fouty WJ. Trauma score. Crit Care Med 1981;9(9): 672-6.
4. Raju NV, Maisels MJ, Kring E, Schwarz-Warner L. Capillary refill time in the hands and feet of normal newborn infants. Clin Pediatr (Phila) 1999;38(3): 139-44.
5. Gorelick MH, Shaw KN, Baker MD. Effect of ambient temperature on capillary refill in healthy children. Pediatrics 1993;92(5):699-702.
6. Otieno H, Were E, Ahmed I, Charo E, Brent A, Maitland K. Are bedside features of shock reproducible between different observers? Arch Dis Child 2004;89 (10):977-9.
7. Schriger DL, Baraff LJ. Capillary refill--is it a useful predictor of hypovolemic states? Ann Emerg Med 1991; 20(6):601-5.
8. Schriger DL, Baraff L. Defining normal capillary refill: variation with age, sex, and temperature. Ann Emerg Med 1988;17(9):932-5.