Are we on the same page? Assessing inter-rater reliability of video recorded cardiac arrest review

Authors: Joseph Landers BS, Molly McCann-Pineo PhD MS, Katherine T Carroll BA, Timmy Li PhD, Ghania Haddad MD, Elizabeth A Young MPH, Lance Becker MD, Daniel Rolston MD MPSH, Daniel Jafari MD MPH

Affiliations

Categories: Cardiac arrest, resuscitation, research, quality improvement Keywords: cardiac arrest, video review, interrater reliability, quality improvement

Background:

Audio and video recording of cardiac arrest (CA) is a valuable research, education, and quality improvement tool. Ensuring inter-rater reliability (IRR) is crucial. Current studies often poorly describe or fail to take steps to ensure IRR. We aim to evaluate IRR of data in our CA data registry obtained by our reviewers from video recordings of Emergency Department (ED) CAs. We hypothesized that our reviewers have moderate to high IRR.

Methods:

This is a retrospective study of a random sample of CA videos from a quaternary care academic ED with an annual average of 112 CAs. All adult CA patients with video recordings available were reviewed by one of two expert video reviewers, who are board certified emergency medicine and critical care physicians. 21 CAs were randomly selected for blind evaluation by the other expert reviewer. Reviewers assessed the recordings for metrics using a standardized scoring sheet, including number of intubation attempts, emergency medical services (EMS) reported initial rhythm, initial ED rhythm, and total duration of cardiopulmonary resuscitation (CPR) interruption. EMS reported initial rhythm and initial ED rhythm were reported as either a shockable, non-shockable, documented unknown, or undocumented. The total duration of CPR interruption was defined as the sum of all time intervals (in seconds) in which chest compressions were interrupted without return of spontaneous circulation (ROSC) beginning after patient arrival. IRR was evaluated utilizing Cohen's Kappa and Intraclass

coefficient (ICC). Kappa statistics were produced for categorical variables, number of intubation attempts, EMS and hospital rhythm. ICC was produced for total interruption time. This study was deemed exempt by the institutional review board.

Results:

21 pairs of CA video reviews were assessed for IRR. IRR was excellent when assessing the total interruption of CPR time, (ICC=0.97, 95% CI 0.94, 1.00). IRR was moderate when assessing the number of total intubation attempts that occurred throughout the recording (k=0.78, CI 95% 0.56, 1.00). IRR was weak when assessing the categorical variables of stated EMS rhythm (k=0.45, 95% CI 0.17, 0.73) and initial ED rhythm (k=0.56, 95% CI 0.28, 0.84).

Discussion:

Our results showed excellent IRR of the continuous metric, total CPR interruption time but weak to moderate IRR of discrete variables such as EMS rhythm, initial ED rhythm, and number of intubation attempts.