Published in final edited form as:

Ann Emerg Med. 2008 November ; 52(5): 504–511. doi:10.1016/j.annemergmed.2007.11.036.

Accuracy Evaluation of Emergency Department Tracking System Timestamps

Bradley D. Gordon, M.D., M.S., Thomas J. Flottemesch, Ph.D., and Brent R. Asplin, M.D., M.P.H.

Regions Hospital Department of Emergency Medicine, St Paul, MN, USA and the HealthPartners Research Foundation, Bloomington, MN, USA

Abstract

OBJECTIVES—Managers use timestamps from computerized tracking systems to evaluate emergency department (ED) processes. This study was designed to determine how accurately these timestamps reflect the actual ED events they purport to represent.

METHODS—An observer manually timestamped patient and provider movement events during all hours. The observed timestamps were then systematically matched to equivalent timestamps collected by an active tracking system (timestamps created by staff with keyboard/mouse) and a passive tracking system (timestamps created by sensor badge worn by staff members). The deviation time intervals between the matched timestamps were analyzed.

RESULTS—The observer noted a total of 901 events; 686 (76%) of these were successfully matched to active system timestamps and 60 (6.7%) were matched to passive system timestamps. For the active system, the median event was recorded 1.8 minutes before it was observed [IQR: 30.7 minutes before to 2.9 minutes after]. Protocol execution difficulties limited the study of the passive system (low number of successfully matched events). The median event was recorded by the passive system 1.1 minutes before it was observed [IQR: 1.3 minutes before to 0.9 minutes before] (n=60).

CONCLUSIONS—The timestamps recorded by both active and passive tracking systems contain systematic errors and non-normal distributions. The active system had much lower precision than the passive system, but similar accuracy when large numbers of active system observations are used. Medians should be used to represent timestamp and interval data for reporting purposes. Site specific data validation should be performed before using data in high-profile situations.

Key MeSH Subject Headings

Emergency Service, Hospital: organization & administration; Medical Records Systems, Computerized: organization & administration; Hospital Information Systems: organization &

Author correspondence may be sent to bradley.d.gordon@healthpartners.com.

Gordon and Asplin are members of the Department of Emergency Medicine, University of Minnesota Medical School, Minneapolis, MN, USA

Presented as a poster presentation at the Annual Meeting of the Society for Academic Emergency Medicine, San Francisco, CA: Saturday, May 20, 2005 9:00–11:00 am

Reprints are not available from the authors.

administration; Automatic Data Processing/instrumentation/methods; Patient Identification Systems: organization & administration; Time Factors

Introduction

Background

Healthcare in the United States has been indicted for mediocre quality despite high expenditures of resources.¹ As a result, healthcare leaders have been called to action to rapidly improve the healthcare system.² Measurement is central to the concept of quality, because an improvement in quality can only be demonstrated through careful measurement. The data derived from widespread use of emergency department (ED) computerized tracking systems provides ample opportunity for health services measurement.³ It is not clear, however, the degree to which operations data that are reported to managers are accurate. A primary goal of ED tracking systems are to reflect the actual state of the ED at any given point in time to help ED providers to keep track of the status of patients; recording event timestamps for later analysis is not fundamental to this goal. A secondary goal of these systems is to assist data collection for ED visits.

Additionally, two general types of computerized ED tracking systems exist: *active* and *passive*. Active tracking systems rely primarily upon a patient care provider using the graphical user interface of a computer to keep the tracking system's information up-to-date. Passive tracking systems (also known as 'real-time location systems') track the physical location of patients or providers via wireless location sensing technology. The distinctions between these two types of systems likely have significant impact upon the accuracy of the data they capture.

Importance

As more emphasis is placed upon quality improvement initiatives, there are now larger consequences for failing to achieve high quality performance. Much of the data used to evaluate the success of ED operations is derived from ED computerized tracking systems. The accuracy of these data must be established so that we may achieve reliable and accurate reporting methods for our quality management programs. For example, a current national quality benchmark is completion of an ECG within 10 minutes of arrival into the ED for patients with chest pain or suspected of myocardial infarction.^{4, 5} If the timestamps used to measure a department's success rate with this measure were inaccurate by one minute, the change in the success rate could be significant. Failure to successfully achieve quality goals could have marked impacts to a department's financial performance and community reputation. Calls for further research in ED operations emphasize the importance of understanding how these tracking systems produce data for further analysis.⁶

There appears to be limited formal evaluation of the accuracy of tracking system data. Within a study designed to evaluate ED waiting times and patient satisfaction, an informal analysis of active tracking system timestamp data was performed, concluding that the data are accurate to within 1–5 min of the actual events.⁷ In a more formal evaluation of timestamp accuracy in the operating room environment, it was determined that the passive

tracking system was more accurate than the active tracking system's manually documented timestamps.⁸ Another study performed to assess the accuracy of other data elements from ED tracking systems concluded that data must be independently validated before they are used for policy or research purposes.⁹

We aimed to decrease the gap of knowledge that exists around data accuracy derived from tracking systems. The results of this study will provide providers and administrators with useful guidance regarding the appropriate use of tracking system data in the management of their organizations.

Goals of this investigation

The primary aim of this study was to assess the time accuracy of patient and provider movement timestamp data generated by ED computerized tracking systems. Our goal was to describe the accuracy and precision of the timestamp data via direct comparison to a gold standard measurement. Secondary aims of this study were to compare the accuracy between 1) types of events and 2) types of computerized systems used to generate the data (active tracking systems and passive tracking systems).

Methods

Computerized tracking systems capture events over time. To measure the time accuracy of the events, we placed an observer in the ED to independently record the time of specific events. Then we compared the observed event time to the event times recorded by the two tracking systems.

Study design

This study was a prospective observational study designed to capture the actual time events in the ED occurred, and compare that time to the data captured by our active and passive tracking systems.

The study was designed to minimize the risk to data integrity. ED staff members were not informed of the study's plan to compare the observer data to the active system. This planned comparison was hidden from ED staff; because we felt behavior change a likely result and this change would adversely affect the integrity of the data in the active tracking system. ED staff members were informed of the study's plan to compare observer data to the passive tracking system, because the passive system requires only the activation and wearing of a badge and these actions were not likely to be altered.

This study was prospectively reviewed and approved by the Institutional Review Board (IRB) under a 'minimal risk' category with special consideration of a deceptive technique, and was granted a waiver of informed consent. As this study involved a deception technique, procedures to debrief study subjects were planned prior to study execution. Data collected on three physicians were removed prior to analysis because they were aware of the complete study protocol. Following the completion of data analysis phase, the full study protocol was disclosed to all ED staff.

Setting

Regions Hospital is a urban teaching hospital located in St. Paul, Minnesota, with an approximate annual patient visit volume of 66,000. At the time of the study, the ED had 35 designated treatment areas.

Data collection

The timestamps compared in this study were those that could be readily obtained from all three systems without disruption of the standard care process. Table 1 describes the definition of each type of event included in this study.

The three data collection systems were:

Observer System: A trained observer manually recorded the time of each event using a portable computer and data collection system implemented in the Microsoft Access database package. (Microsoft Corporation, Redmond, WA)

Active Tracking System: The Epic ASAP[®] Emergency Department Module was utilized by ED staff to track patients through their ED visit. (Epic Systems Corporation, Verona, WI)

Passive Tracking System: The infrared-based Executone Integrated Locating System (ILS) tracked provider entry and exit times to each sensor area within the ED. (Executone Information Systems, Inc., Milford, CT)

Time on all three systems (observer, active and passive) was coordinated and maintained with Internet network timeservers to ensure time synchronization among the different timestamp sources.

To minimize observer time inaccuracy, the observer's graphical user interface was designed so that a button was pressed to record an event at the exact time, followed by time spent completing data entry regarding that event. The observer's training consisted of an initial orientation to the basic ED care process, including instruction on how to recognize each person's role (e.g. nurse or physician) and how to record each timestamp. The study was piloted over 2 four-hour periods, with direct observation by the principal investigator, to ensure observer training was successful and recorded timestamps were accurate.

The study focused on one specific area of the emergency department for specific reasons: These rooms typically receive a diverse cohort of patients, Emergency Severity Index (ESI) acuity classification II–V, both walk-in and ambulance arrivals. Additionally, this area also has a convenient work area for the observer, has visibility of the maximum number of rooms, and is minimally intrusive to staff and patient flow.

The research assistant was trained to passively observe this four-room area of the ED for randomly assigned four-hour blocks of time, stratified over all hours of the day. Four-hour time periods allowed the observer to remain focused during the observation time and provided some additional scheduling flexibility. Review of ED census logs indicated that approximately 34 shifts would be required to capture at least 400 matched observations. An

additional eight shifts were added to the schedule to ensure that sufficient numbers of matched events occurred, and to allow equal sampling of each shift.

To minimize bias due to time of day or heavy workloads, the observer was trained to capture all of the specified events for those rooms over the shift period. The study pilot demonstrated this approach was the most feasible and least intrusive, compared to other possible methods of data collection.

Each treatment room observed in this study had a unique infrared sensor inside the room and a solid opaque door and in-room curtain. These barriers decrease the chance of false positive room entry detection. Each provider was issued an infrared transmitting badge upon hiring; however, actual use was variable. To maximize the use of the badges, the research assistant had a number of extra badges on hand to loan to ED providers during their shift, should they not have one to wear.

In our active tracking system, a typical patient is greeted and immediately entered into the system, with registration occurring later in the process. A typical patient has these events: Patient Arrived; Patient Roomed; Begin Physician Exam; Disposition Entered; Admit/Discharge Event. In our system, there are also timestamps when providers assign themselves to a patient's treatment team. Providers assign themselves to the treatment team, and are instructed to do so when they are initiating patient care. An exception: attending physicians are frequently assigned by mid-level providers, for convenience, rather than when they are entering the room. Therefore, we expected higher variability with the attending timestamps. Patients are discharged from the system by the primary nurse, and patients are admitted from the ED by the ED clerk when they are physically transferred to the floor. In this study, we observed and attempted matching to the Patient Roomed, Admit/Discharge, and treatment team assignment events.

Emergency department staff members were introduced to the study through mass email announcing the study, as well as through a single page study information sheet. The observer had copies of this sheet to present to those who asked for information. Per the study protocol, these communications only referred to the evaluation of the passive tracking system.

Data processing

Data extracts were pulled from the active and passive systems corresponding to all of the four-hour time periods when the observer was present in the ED. Wide time margins were built into the queries to ensure all possible timestamp data were included.

Both systems generated a considerable number of possible matches with events recorded by the observer. To match this large volume of data with a consistent approach, an internally developed computer program managed the matching process.

The algorithm identified candidate matches on the basis of ED room number, patient medical record number and ED staff member name. The algorithm used wide brackets of time to select candidates for matching (active: twelve hour bracket, passive: four hour bracket). Wide time brackets were chosen to maximize the number of tracking system

timestamps available to the matching algorithm, and thereby minimize the chance of failing to match a widely erroneous timestamp (i.e. false negatives). The research assistant supervising the matching was blinded to the observed event time thereby minimizing systematic bias in this phase of the protocol. Therefore, the matching process was structured to provide matches independent of the outcome of interest: time difference between events.

The end result of this process was a list of observed timestamps that were matched to 0, 1 or 2 other systems' timestamps.

Data analysis

The primary unit of analysis was the deviation time interval of each directly observed timestamp to its matched computerized system timestamp. We hoped to identify a series of straightforward adjustments to the active and passive timestamps would correct for any bias and make them more reflective of reality.

The deviation time intervals were evaluated for normality (Shapiro-Wilk tests) and summarized using descriptive statistics (mean, standard deviation, 95% confidence interval, median and inter-quartile range (IQR)).

For each active tracking system event type, and for the passive tracking system events, one sample Wilcoxon rank sum tests were performed to determine the statistical significance of deviations from zero, with a significance level of 0.05. Statistical and graphical analysis was performed using the R statistical package (R Foundation for Statistical Computing, Vienna, Austria).

Sensitivity analysis

As previously stated, ED census logs were used as exploratory data to understand the frequency of room turnover of specific areas of the emergency department. This allowed us to understand how many events a single observer might be able to capture over a given period of time. Additionally, we estimated a range of variability in the resultant dataset to perform a sample size analysis. Our analysis also assumed a maximum deviation time interval ranging from 30 to 90 minutes. The study was powered such that the 95% confidence interval of the mean deviation time interval would approximate 10% of the mean. The run-in period previously mentioned demonstrated that the number of observations was consistent with the exploratory data analysis, and no changes were made to the study protocol.

Results

Protocol Implementation

The research protocol was implemented during the summer of 2005, over a two-month period. Forty-two blocks of time were observed, resulting in 901 total observations. Each observation period was four hours in length. These observations were successfully matched according to the above algorithm to 686 (76%) active tracking system timestamps, and to 60 (6.7%) passive system timestamps.

As providers expressed no knowledge of the full extent of the study during the study debriefing process, it appears that the partial blinding technique was successful.

Main Results

Evaluation of Normality—We studied the data distribution for its basic characteristics, including normality because there was no substantial pre-existing data to guide our analysis procedure. Using Shapiro-Wilk test of normality, we determined that the distributions of both the active and passive timestamp deviation time intervals were non-normal ($p < 0.005$). Therefore, results are reported as medians and IQRs.

Timestamp Comparisons – Active Tracking System—Overall, events were recorded 1.8 minutes (median, IQR: 30.7 min before to 2.9 min after) before they were observed. We also stratified the analysis by event type to find if there were systematic differences in accuracy by event type. Nonparametric statistical testing demonstrated all active tracking system medians were significantly different than zero. These results are reported in Table 2 and displayed graphically in box plot format in Figure 1. This analysis indicated the active tracking system did not significantly differ from the recorded system with regards to patient movement (e.g. a patient entering and a patient leaving a room), but providers did tend to prematurely record their actions in the active tracking system. Table 2 also includes the mean of each event type in order to illustrate the degree and direction of distribution skew when compared with the median.

Passive Tracking System—Unfortunately, we were unsuccessful at matching most of the observer's timestamps to their equivalent passive tracking system timestamps. Post-study data analysis indicates that two major factors contributed to the outcome. First, due to user error when activating the passive tracking system's recording function, passive system timestamps were not recorded to disk for approximately 10 of the 42 observing blocks. Retraining occurred following the detection of this error, and further data loss did not occur. While unfortunate, this user error does not completely explain the 6.7% match rate. If the same rate of passive timestamp matching would have occurred during the shifts with lost data as existed when data was present (1.87 average passive matches per shift) we would expect an overall 8.8% successful match rate. Provider usage of passive system tracking system badges was sporadic and we were unable to successfully execute our plan to loan temporary badges to providers during observation periods. Re-execution of the study was not possible due to resource and system constraints.

In analyzing the deviation time intervals from the 60 matched events, we found a comparatively high degree of precision. The events were recorded 1.1 min (median, IQR: 1.3 before to 0.9 before) before they were observed. This result was significantly different than zero. Seventy-two percent were recorded within ± 2 minutes of the observed event. The data are graphically summarized in Figure 2. Sub-group analysis by provider type was not performed on the passive tracking system data due to the low number of successfully matched events.

Limitations

The passive tracking arm of the study was significantly limited by low numbers of successfully matched timestamps. However, we do find that more than 50% of the matched timestamps were observed to occur approximately one minute after they were recorded by the passive system. The low variability of these deviation time intervals is likely due to the lack of the influence of human behavior on the recording process.

At first glance, it would seem that passive systems are insulated from erratic human behavior because the passive tracking system matched events were much more precise than the active tracking system matched events. However, if staff members choose not to wear a passive tracking system sensor badge, the reports from such systems are systematically biased by absent data.

This lack of use of the sensor badge partially caused one of the major limitations in this study. This limitation has been cited as the main operational problem in the recently published study evaluating a radio-frequency based passive tracking system in the OR.⁸ Organizations implementing this type of information system should expect these same challenges. From our experience, an organization can improve compliance by: ensuring badges have low-battery warning systems, easy access to badge replacement/repair, proper devices to attach badges to important objects (such as stethoscope or identification badge) and devices that manage their own power such that users don't need to remember to turn-on and turn-off at the end of a work period. The key to badge compliance likely lies in creating direct value for the provider wearing the badge. If other work tasks were saved by wearing a badge, (e.g. assigning patients automatically rather than having to manually perform in tracking system) it is likely more providers would choose to wear them. In many installations, however, the value from wearing a badge is delivered to others, rather than to the provider wearing the badge. (e.g. clerk determining the provider's location in order to route a phone call).

The passive system studied here was installed in mid-1990s with the goal of tracking provider location (not patients). Technology improvements in system architecture (radio-frequency identification tags) are occurring and likely provide enhanced features. Such features may include badge-related improvements (listed above), as well as software improvements (e.g. data filtering for meaningful events). However, these systems will continue to be hampered if staff members are not motivated to wear their badges.

While a majority of the active timestamps were successfully matched, a significant proportion (24%) were not. Analysis of unmatched timestamps revealed no identifiable patterns. It is not clear if these unmatched timestamps would have influenced the results.

Not all treatment areas were included in this study. The area observed primarily serves acute and sub-acute conditions with most patients not requiring critical care services. This makes the results less generalizable to every treatment area in an ED because, for example, providers in a resuscitation area are frequently present before the patient arrives and active tracking systems may have difficulty recording this type of event.

The study did not evaluate other types of events, such as: inpatient bed request time, electrocardiogram completion time, etc. While these events are important in measuring the care process, this study was performed on events that were observable by the research assistant without disturbing the care providers.

Finally, because this study was performed at one center, specific estimates of accuracy are likely not valid when applied to other environments. However, many of the same assumptions and practices that occur in our setting likely hold true in other settings.

Discussion

Our initial assumptions of maximum deviation time intervals of 30–90 minutes somewhat approximated the results of the active tracking system, as the bulk of the events were recorded within 60 minutes of the actual event occurrence. However, we found the resultant timestamp data are both heavily skewed and contain a significant number of outlying observations. This calls into question the use of means when reporting time interval data from such systems. Due to the skew of the data, the mean is not a reliable measure of central tendency. Given an adequate number of observations, the median is a more meaningful measure of typical performance for purposes of decision support.

While the active system appeared to accurately and precisely capture patient movement, providers consistently updated the system prior to actually performing the action. This caused the event to be prematurely recorded and calls into question the utility of such timestamps when assessing performance and analyzing quality improvement initiatives.

A determination of the causes of variability between event types was not performed in this study. However, we speculate that because each event type has a different workflow, variations between these workflows modify the degree of data accuracy in the active tracking system. The authors observe in daily practice that the textual description of a timestamp often doesn't reflect the actual use of that event in an active tracking system. For example, our tracking system records a "Disposition Decision Time", which seems to reflect the time the provider decided on a planned disposition (admit, discharge, etc). In practice, it records when a provider took the time to enter a disposition. That provider may have actually decided to admit an intubated, hypotensive patient from the first moment of interaction, but didn't mark the tracking system as 'admit' or order a bed until later in the care process. Further qualitative research into these types of differences in workflow would likely provide a more clear understanding of why the degree of error varies by event type.

Protocol implementation problems severely limited the number of matched observed timestamps to the passive tracking system data. Therefore, conclusions from this arm of the study are limited.

Conclusions

In summary, tracking systems have become an integral component in managing busy healthcare environments. Tracking systems provide a massive amount of detailed data about the care delivery process. Yet, research into understanding the accuracy and precision of this

new source of administrative data has been limited. This work contributes a real-world accuracy assessment at one center.

We conclude that time interval data from active tracking systems is most accurate when medians are used for summarization. Additionally, events of different types are recorded with differing degrees of accuracy.

Managers using reports with this data need to know these limitations when interpreting tracking system reports. It would be wise to implement local validation of data to determine the degree of inaccuracy that exists. This would be particularly important when many groups of comparison result in low numbers of observations in each group (such as in staff performance profiling).

Passive tracking systems likely improve on the precision of the times recorded, particularly when they can be linked to meaningful events in the workflow and staff participate in the use of the sensing devices.

Acknowledgments

Dr. Gordon is supported by a National Library of Medicine Early Career Development Award in Medical Informatics (K22LM008573-01)

The authors wish to give thanks to Joel Holger M.D. for project development review

References

1. Kohn, LT.; Corrigan, J.; Donaldson, MS. To err is human: building a safer health system. Washington, D.C.: National Academy Press; 2000.
2. Institute of Medicine (U.S.). Committee on Quality of Health Care in America. Crossing the quality chasm: a new health system for the 21st century. Washington, D.C.: National Academy Press; 2001.
3. Husk G, Waxman DA. Using data from hospital information systems to improve emergency department care. *Acad Emerg Med*. 2004; 11(11):1237–1244. [PubMed: 15528590]
4. Antman E, Anbe D, Armstrong P, et al. ACC/AHA guidelines for the management of patients with ST-elevation myocardial infarction; A report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Committee to Revise the 1999 Guidelines for the Management of patients with acute myocardial infarction). *J Am Coll Cardiol*. Sep; 2004 44(3):E1–E211. [PubMed: 15358047]
5. Diercks DB, Peacock WF, Hiestand BC, et al. Frequency and consequences of recording an electrocardiogram >10 minutes after arrival in an emergency room in non-ST-segment elevation acute coronary syndromes (from the CRUSADE Initiative). *The American journal of cardiology*. Feb 15; 2006 97(4):437–442. [PubMed: 16461033]
6. Beach C, Haley L, Adams J, et al. Clinical Operations in Academic Emergency Medicine. *Acad Emerg Med*. Jul 1; 2003 10(7):806–807. 2003. [PubMed: 12837658]
7. Hedges JR, Trout A, Magnusson AR. Satisfied Patients Exiting the Emergency Department (SPEED) Study. *Acad Emerg Med*. 2002; 9(1):15–21. [PubMed: 11772664]
8. Marjamaa RA, Torkki PM, Torkki MI, et al. Time accuracy of a radio frequency identification patient tracking system for recording operating room timestamps. *Anesthesia and analgesia*. Apr; 2006 102(4):1183–1186. [PubMed: 16551921]
9. Svenson JE, Pollack SH, Fallat ME, et al. Limitations of electronic databases: a caution. *J Ky Med Assoc*. 2003; 101(3):109–112. [PubMed: 12674902]

Deviation time intervals of active tracking system recorded events

Deviation time intervals of active tracking system recorded events (zoomed to 240 minute range)

Figur 1.

Figure 2.

Table 1

Event type descriptions

Event Type	Observer	Passive Tracking	Active Tracking
Patient placed in room	New patient is placed in room	Not applicable **	Patient roomed time
Primary nurse enters room	Primary nurse enters room to start care	First instance of primary nurse entering room	Primary nurse assigned to patient
Mid-level * enters room	First mid-level provider enters room to start care	First instance of primary mid-level entering room	Mid-level provider assigns self to patient
Attending physician enters room	First attending physician enters room to start care	First instance of attending physician entering room	Attending physician assigns self, or is assigned by midlevel provider, to patient ***
Patient leaves room for disposition	Patient is discharged or admitted or expires	Not applicable **	Patient disposition time

* Mid-level providers: ED and Rotating residents, Physician Assistants and Medical Students

** The passive tracking system in our ED only tracked staff members, it did not track patients.

*** The knowledge that mid-levels sometimes assign attending providers gave us *a priori* knowledge that this event would have decreased accuracy.

Table 2

Active Tracking System Event Results, by Event Type

Event Type*	N	<i>Time difference of observed timestamp from tracking system timestamp**</i>		
		Mean (min)	Median (min)	IQR (min)
Patient enters room	150	6.4 before	0.26 after	1.4 before to 2.8 after
Attending first enters room	108	23.7 before	30.0 before	57.7 before to 1.3 before
Mid-level first enters room	171	41.7 before	10.2 before	67.7 before to 3.0 before
Nurse first enters room	129	15.5 before	11.3 before	47.8 before to 4.7 after
Patient leaves room	121	16.7 after	4.2 after	1.4 after to 9.1 after

* While we recorded timestamps from social worker staff, the data were not analyzed further, due to a limited number of successfully matched events (n=7).

** Interval computed as: tracking timestamp minus observer timestamp. Therefore, “6.4 before” indicates the event was recorded in the tracking system 6.4 minutes before the same event was observed.