Published in final edited form as:

Res Nurs Health. 2014 February ; 37(1): 75–84. doi:10.1002/nur.21578.

Can Real Time Location System Technology (RTLS) Provide Useful Estimates of Time Use by Nursing Personnel?

Terry L. Jones [Assistant Professor] and

University of Texas at Austin School of Nursing, 1710 Red River Street, Austin, Texas 78701

Cara Schlegel [Director, Seton Family of Hospitals, Doctoral Student]

Diagnostics and Therapeutics Informatics, Seton family of Hospitals, University of Texas at Austin School of Nursing, Austin, Texas

Abstract

Accurate, precise, unbiased, reliable, and cost-effective estimates of nursing time use are needed to insure safe staffing levels. Direct observation of nurses is costly, and conventional surrogate measures have limitations. To test the potential of electronic capture of time and motion through real time location systems (RTLS), a pilot study was conducted to assess efficacy (method agreement) of RTLS time use; inter-rater reliability of RTLS time-use estimates; and associated costs. Method agreement was high (mean absolute difference = 28 seconds); inter-rater reliability was high (ICC = 0.81–0.95; mean absolute difference = 2 seconds); and costs for obtaining RTLS time-use estimates on a single nursing unit exceeded \$25,000. Continued experimentation with RTLS to obtain time-use estimates for nursing staff is warranted.

Keywords

real time location systems; radiofrequency identification technology; time and motion studies; instrument development and validation; nursing staffing; quality assurance; patient safety

Declining hospital reimbursement and a national nursing shortage have been linked to lean staffing practices and rising concerns regarding the time available for bedside nurses to complete important care activities (Heinz, 2004; Norrish & Rundall, 2001). Time scarcity among bedside nurses results in the omission of nursing care activities and is associated with negative patient and nurse outcomes (Jones, 2013; Kalisch, Landstrom, & Williams, 2009; Schubert et al., 2008). This missed care or implicit rationing of care is routinely experienced by almost all bedside nurses. Concerns related to time scarcity among bedside nurses led to a proliferation of nurse staffing research over the past two decades (Heinz, 2004; Kane, Shamliyan, Mueller, Duval, & Wilt, 2007; Numata et al., 2006; Thungjaroenkul, Cummings, & Embleton, 2007) and a national initiative to increase available time at the bedside for direct care (Rutherford, Lee, & Greiner, 2004). Despite the increased attention to time use among bedside nurses, challenges in time-use estimation persist.

Conventional Time-Use Estimation Methods

The four major approaches to quantification of time use among social scientists include: (1) continuous direct observation of time and motion; (2) real time activity sampling (referred to as work sampling when the activities of interest are associated with a particular job-related role such as bedside nursing); (3) self-report; and (4) derived estimates from administrative databases. These approaches differ in accuracy, intrusiveness, specificity, precision, bias, and efficiency.

Continuous Direct Observation

Direct observation of time-use behavior arguably remains the gold standard for accuracy in quantification of time use (Bratt et al., 1999; Larson, Aiello, & Cimiotti, 2004; Ver Ploeg et al., 2000). A designated observer follows a subject of interest in real time and records the duration of time spent on activities of interest and/or in locations of interest. The accuracy gained through direct observation does not come without significant costs, which include human resources for observation and data entry, intrusion during interactions and activity flow, and potential for changes in time-use behaviors based on perceived social desirability (Larson et al., 2004; Bratt et al., 1999; Weigl, Müller, Zupanc, & Angerer, 2009). Continuous direct observation of time use is often considered cost-prohibitive for large-scale projects.

Work Sampling

Time-use estimates generated by work sampling reflect the proportion of time spent on activities or at locations rather than actual time duration. Estimation of time use through work sampling is achieved through recording predefined activities as they are performed and/or the worker's presence at predefined locations at a number of randomly selected times or established intervals (Robinson, 2010). Each recorded activity/location is considered an occurrence. The sum of occurrences for a given activity or location is divided into the sum of occurrences across all activities/locations to obtain a time-use estimate for that activity/location (Finkler, Knickman, Hendrickson, Lipkin, & Thompson, 1993; Pape, 1992).

Traditionally, the frequency of occurrences has been obtained through direct observation by trained observers shadowing one or more subjects of interest. Observers are equipped with a checklist of activities/locations and a stopwatch or pager programmed to alarm at predetermined or randomly determined times. At the predetermined intervals or at the sound of the alarm, observers record the observed activity/location on the checklist. Because work sampling involves intermittent rather than continuous recording of time-use behavior, it may be possible to enhance efficiency of data collection by having one observer shadow multiple subjects simultaneously.

In work sampling, accuracy and precision are a function of the number of sample points and the level of detectable time proportion desired (e.g., activities/locations occurring at a frequency of 5%, 10%, 15%, etc.). The number of sample points is determined by the number of subjects shadowed, the duration of the data collection period (i.e., days, weeks, or months) and the interval between sampling points (i.e., every 5, 10, or 15 minutes).

Estimation of time use with a high degree of precision and confidence through work sampling often requires significant human resources, particularly if detection of activities/locations associated with infrequent time use is desirable (Pelletier & Duffield, 2003). For example, Finkler et al. (1993) reported a 20% or greater difference in time-use estimates obtained by continuous direct observation (8 residents observed for 24 hours) and work sampling (15-minute sampling intervals and 892 total observations) for 8 of 10 activities recorded. Almost all of the activities (9 of 10) occurred at a frequency of <20%. The number of observations required to achieve time-use estimates with 10% precision varied significantly based on frequency of occurrence: 1,532 for activities consuming >20% of time; 3,682 for activities consuming 10% of time; 7,007 for activities consuming 5% of time; and 21,822 for activities consuming 2% of time.

More recently, personal digital assistants (PDAs) have been used for self-report and direct entry of time-use data in response to randomly timed alarms for work sampling studies (Ogunfiditimi, Takis, Paige, Wyman, & Marlow, 2013; Robinson, 2010; Upenieks, Akhavan, Kotlerman, Esser, & Ngo, 2007). The use of PDAs can reduce the resources required for data collection and data entry compared to direct observation, but the issue of potential response bias persists (Donaldson & Grant-Vallone, 2002; Robinson, 2010). Moreover, an alarming PDA may be perceived by the bedside nurse as intrusive and disruptive to workflow, particularly when the time interval between alarms is shortened to enhance precision.

Self-Report

Due to the intensity of resources required for direct observation, self-report measures of time use are often used instead. The most commonly used self-report time-use measures include time diaries (TD), experiential sampling methodology (ESM), and stylized respondent reports (SRRs). TDs typically require respondents to keep a chronological record of their activities over a predetermined period of time. The record can be completed in real time or retrospectively. The most common approach involves free-form entries, allowing respondents to use their own description of activities and include actual start and stop times. TDs completed in real time are considered to have minimal recall error, and the recording of actual start and stop times enhances the accuracy and precision of time-use estimates (Otterbach & Sousa-Poza, 2010). Respondent burden can be significant, however, particularly when free-form responses are required. For this reason, TDs typically are limited to 1–2 days per participant, which potentially limits the potential for pattern recognition and generalizability of time-use estimates (Lin, 2012). Moreover, error may be introduced through the data coding process, and significant resources may be required for coding and data entry of the free-form responses.

Similar to the work sampling approach, ESM involves collection of data at multiple randomly selected times over a predefined period (i.e., day, week, month). Respondents are provided with a programmable device that is activated (e.g., to beep, vibrate, or buzz) randomly throughout the data collection period. In response to the alarm, respondents record information about what they are experiencing in that moment. In contrast to work sampling, the information recorded in ESM can be rich in detail and include multiple aspects of the

time experience (e.g., cognitive, behavioral, and affective) (Juster, Ono, & Stafford, 2003). Self-report forms for ESM typically contain a set of core items, which may use a variety of response options: free-form text, fill in the blank, semantic differential scales, visual analog scales, and checklists (Ver Ploeg et al., 2000). As in work sampling, time-use estimates represent proportions of time spent rather than actual duration of time spent. Because detailed information is recorded in ESM, respondent burden and resources for data coding and entry can be significant.

SRRs of time spent require respondents to recall and estimate how much time they “normally” or “typically” spend on a list of predefined activities within a given time frame (e.g., day, week, month). Response options can be crafted to assess relative and/or absolute time spent. Response options to assess absolute time spent may be open-ended, allowing respondents to fill in a specific amount of time, or they may include ordinal scales with ranges of time duration (Manson, Levine, & Brannick, 2000; Otterbach & Sousa-Poza, 2010; Ver Ploeg et al., 2000). Stylized items may measure relative time spent by asking respondents to use Likert-type response options to rate the amount of time spent performing an individual task relative to all other tasks being considered (Manson et al., 2000). SRRs can be completed using a variety of formats, including interviews (by phone or in person), paper and pencil mail surveys, and online surveys. Although respondent burden and data collection costs are lower with SRRs than for the other self-report methods, the potential for recall bias and aggregation error is greater.

All direct observation and self-report methods have the potential for biased estimates that favor time spent on socially desirable activities (Donaldson & Grant-Vallone, 2002). Greater overall bias in time-use estimates has been consistently demonstrated in SRRs compared to TDs and direct observation (Bratt et al., 1999; Collopy, 1996; Juster et al., 2003; Lin, 2012; Otterbach & Sousa-Poza, 2010). Discordance between measures can be significant. For example, Bratt et al. (1999) reported mean absolute differences of 59–60 minutes in daily time-use estimates, and Collopy (1996) reported median absolute differences of 32–47%.

Reported concordance among time-use estimates across methods is highly variable. Hunting et al. (2010) reported moderate agreement (60%) between time-use estimates across nine task categories among construction workers, but differences in concordance were noted based on relative task proportion and job role. Intraclass correlation coefficients (ICCs) for major tasks (i.e., those performed > 1 hour/day) ranged from 0.52 to 0.85 and were considered good to excellent. In contrast, ICCs for minor tasks (e.g., those performed for < 1 hour/day) were primarily poor (ICCs 0.39–0.54). There was a trend toward higher self-report time-use estimates for major tasks and lower self-report estimates for minor tasks. Agreement among time-use estimates was higher among workers in specialized job roles with consistent work patterns (few tasks routinely performed in a controlled environment) compared to workers in roles with greater variability in work patterns (many tasks performed in response to variable situational contexts). In job roles with variable work patterns, over half of the time-use estimates differed by more than 1 hour. This is particularly relevant as work patterns among bedside nursing also are highly variable and context-dependent. Despite variable ICCs, there was good agreement (79%) between methods regarding the rank order of time-use estimates.

Burke et al. (2000) compared self-report time-use estimates from bedside nurses to time-use estimates obtained through direct observation and work sampling. No method effect was noted for estimates of the percentage of time spent across four main categories of care, suggesting good concordance between work sampling and self-report. However, significant discordance was noted between estimates of time duration for specific activities. Self-report estimates were 2–3 times longer than direct observation estimates.

Staffing Indices From Administrative Databases

The fourth approach to quantification of time use common among health services researchers involves staffing indices derived from data collected for administrative purposes in conjunction with normal daily operations. The most common staffing indices include nursing hours per patient day (NHPPD), registered nurse full time equivalents (RN FTEs), and nurse patient ratios (NPR), all of which involve aggregated measures of worked hours (from payroll data) and patient census (from billing data; Buerhaus & Needleman, 2000; Currie, Harvey, West, McKenna, & Keeney, 2005; Heinz, 2004; Kane et al., 2007; Numata et al., 2006; Thungjaroenkul et al., 2007). Payroll data about worked hours reflect time spent in crudely defined work roles, such as direct versus non-direct care. Time spent on specific activities within either role is not captured. Patient census data are primarily recorded for billing purposes and reflect the number of inpatients present in a facility at a single point once every 24 hours (most often at midnight when a new billing cycle begins). Actual time spent in the care of specific providers is not captured.

Although staffing indices may be more efficient than other time-use methods and immune from bias associated with perceived social desirability, other limitations of this approach have been reported. Standardized definitions for key variables used in the computation of staffing indices (e.g., direct care and FTE) across organizational databases have not been established. Consequently, significant discordance among staffing indices has been reported (Spetz, Donaldson, Aydin, & Brown, 2008). Moreover, staffing indices are biased estimates of time use because they are known to overestimate time spent in the direct care role (Upenieks et al., 2007). Finally, although staffing indices may be useful to study trends over time, they do not provide sufficient precision to accurately quantify the effect of time spent in nursing care on patient outcomes or uncover the mechanism of action through which nurse staffing might affect patient outcomes (Buerhaus & Needleman, 2000; Clarke, 2007). Staffing indices are crude surrogates for time spent with patients, and it is not known whether increases in nurse staffing actually result in increased time spent with patients.

Health services researchers and hospital administrators are eager for time-use estimates for bedside nursing staff that are accurate, precise, unbiased, reliable, and cost-effective. A single method superior in each of these characteristics has yet to be identified, and exploration of new methods is warranted.

Real Time Location System (RTLS) for Time-Use Estimation

Electronic capture of time and motion through real time location systems (RTLS) is an innovative and promising approach to time-use measurement that is now available for application in the healthcare setting. Radiofrequency identification (RFID) technology is the

primary underlying mechanism for RTLS, and the terms are often used interchangeably. The original application of RTLS in healthcare was for asset tracking and supply chain management. RTLS application subsequently expanded to include tracking patient throughput to document dwell times and identify bottlenecks in patient flow. More recently, experimentation with RTLS as a method to quantify time spent at the bedside by nursing staff has been reported (Hendrich, Chow, Skierczynski, & Lu, 2008). Furthermore, a detailed description of the RFID mechanism for automated wireless time-use data collection has been reported (Jones, 2013).

In RTLS, location tracking is achieved using microchips embedded in tags worn by subjects of interest. Each microchip intermittently transmits a unique electronic signal to uniquely identified sensing devices (readers) placed in locations of interest. Upon recognition by a reader, a second wireless signal carrying information regarding the identification number of the tag and the location of the reader is transmitted. This transmission is received and recorded by another sensing device (interrogator), which adds an electronic timestamp reflecting the time the signal was received. A real time movement history for each tag is generated, enabling computation of time spent in each location.

Vulnerability to artifact is a limitation of the RTLS methodology. Artifact can result in invalid, misread, and/or missed RTLS signal entries, all of which represent measurement error and adversely affect the accuracy of time-spent estimates (Jones, 2012). Consequently, a filtering process designed to identify and correct entries due to artifact is a necessary adjunct for the RTLS methodology.

The application of RTLS to obtain time-use data for bedside nurses is in its infancy, and minimal empirical evidence regarding efficacy and feasibility has been reported (Fahey, Lopez, Storfjell, & Keenan, 2013). Therefore, the purpose of this pilot study was threefold: (1) to assess the efficacy of RTLS time-use estimates for bedside nursing staff compared to the gold standard of direct continuous observation; (2) to assess inter-rater reliability of manually filtered RTLS time-use estimates; and (3) to identify the monetary resources required to support time-use estimation among bedside nursing staff using RTLS technology.

Methods

A prospective validation design in a convenience sample of three medical–surgical nursing units in a large academic medical center was used to meet the study aims. After approval from the facility Institutional Review Board and completion of consent procedures, RTLS equipment (tags, readers, and interrogators) was installed and calibrated on each of the three nursing units in accordance with vendor specifications. The equipment was positioned to allow identification of specific patient rooms and common public areas (e.g., nurses' station, medication room, staff break rooms, supply rooms, and conference rooms). Equipment was not positioned to allow identification of staff restrooms or corridor space between rooms (hallways).

RTLS tags were adhered to commercially available lanyards with metal clips to facilitate attachment to participants' uniforms. The lanyards used to attach the tags to uniforms in this

study allowed some tag movement. One end of the lanyard was clipped to the uniform and was relatively immobile, while the other end remained unattached and free to move. Consequently, it was possible for the lanyard to become twisted, so that the side with the tag faced the body instead of outward. In this position, the body could potentially block signal transmission between tag and reader. Tags were typically clipped to the front of uniforms in the area of the collar or lapel. In this position it was possible for the tags to slide underneath outer garments (e.g., sweater or scrub jacket) and become temporarily covered. Moreover, carrying linen or other supplies close to the chest could also result in tags being temporarily covered.

RTLS equipment was sequentially activated on each participating unit for a period of one week beginning Monday at 0700 and continuing through the following Sunday at 0700. At the beginning of each shift all bedside nursing service personnel (registered nurses, licensed vocational nurses, and nurse assistants) were provided RTLS tags and instructed to attach them to their uniform for the remainder of the shift. Research staff was present during each change of shift to ensure that each staff member was equipped with a tag in vivo; however, direct observation of time use among staff was not performed. At the end of the in vivo data collection period, the raw RTLS data were downloaded for manual filtering.

Manual Filtering Process

RTLS artifact often presents as tag transmissions not linked to a specified location of interest. Locations of interest are those areas equipped to monitor time use and are identified by the presence of a uniquely identified RTLS reader. All locations without readers are therefore considered non-monitored areas. Unlinked tag transmissions appear on the movement history (chronological sequence of location changes) as from a “non-monitored area,” which is the generic term applied to all areas outside the sensing range of a reader. Two main scenarios result in unlinked tag transmissions: (1) when the location of the tag is truly outside the sensing range of an activated reader (true signal), and (2) when the location of the tag signal is inside the sensing range but is undetected by the reader (noise/artifact). Therefore, RTLS time-spent estimates for the location defined as non-monitored area are based on a combination of legitimate (true signal) and erroneous (noise/artifact) timestamp entries that should have been linked to a specific location of interest. This can create biased time-spent estimates, that is, the underestimation of time spent in the defined locations of interest.

Effective filtering of RTLS data requires a working knowledge of how RTLS reader placement, RTLS tag position, and unit geography influence time-spent estimates. Unit hallways were not initially defined as locations of interest for this study and were therefore not equipped with readers. Thus, tag transmissions from a hallway as nursing personnel moved from one room to another were outside the range of an RTLS reader and legitimately identified as originating from a non-monitored area. Undetected tag transmissions originating within sensing range of an activated reader are often related to tag position. Tag signals can be missed by readers when tags are positioned behind a solid object or covered up. Transmissions from covered tags that are missed by readers are erroneously recorded by the interrogator as originating from a non-monitored area rather than the true location.

A manual filtering process was developed based on knowledge of unit geography and the identification of discernible patterns in the sequence of non-monitored area timestamp entries on the movement histories (Table 1). Logic algorithms were developed to help distinguish signal (valid entries) from noise (spurious entries) and to guide corrective action for the noise entries. For example, in a sequence involving a timestamp entry for a non-monitored area chronologically positioned between entries for the same specific location, the entry could be logically attributed to within-room artifact related to a blocked signal due to tag position (Table 1, first 5 entries, Room 1 sequence). The logic algorithm stipulated that these timestamp entries be integrated to form a single time-space location. When two adjacent locations were separated by significant geographical distance, the non-monitored area entry occurring between them could be logically attributed to a true signal (Table 1, entries between Room 1 to Room 5 and Room 5 to Room 10). The non-monitored area was thought to most likely reflect time spent in the hallway as the subject traveled from one location to another. The logic algorithm stipulated that these timestamp entries be decoded to the hallway location.

When two locations are in very close proximity, it is plausible for a staff member to move from one location to the other during the interval between tag transmissions (typically set at 3–6 seconds). Therefore, short transit times are not always reflected in the movement history by a time-stamp entry for a non-monitored area (Table 1, Room 10 to Room 9 and Room 9 to Room 8). No corrective action was stipulated in the logic algorithm for this kind of sequence in the movement history.

Transit time between adjacent locations can be extended for a variety of reasons, however, as in the case of interruptions and engagement in other activities in the hallway. Therefore, a timestamp entry for a non-monitored area between entries for locations in very close proximity could arguably reflect either a true signal (from the hallway) or within-room artifact (from either adjacent location). Designation of these entries as signal or noise is associated with comparatively more uncertainty (Table 3, entry between Room 8 and Room 7) and requires consideration of additional contextual information, such as previous movement history patterns around the same geographic location. Therefore, the logic algorithm stipulated that the non-monitored area entry be decoded to one of the specific adjacent locations or to the hallway location based on judgment.

Efficacy of RTLS Time-Use Estimates

Efficacy of RTLS time-use estimates was assessed in conjunction with the RTLS installation and calibration process and involved simultaneous collection of time-use data in a simulated setting with two methods: RTLS and direct continuous observation. On each nursing unit, following completion of the calibration process by vendor staff, an RTLS tag was attached to the uniform of a single participant, and the readers and interrogators were activated. The participant of interest, who was not officially on duty as a bedside nurse, was instructed to simulate movement throughout the nursing unit as he or she would when engaged in patient care. The principal investigator (PI), equipped with a stopwatch, shadowed the participant of interest on each nursing unit and recorded the movement history and duration of time spent in each location. At the end of each simulated session, the raw RTLS movement history data

for the tags attached to the participant's uniform were downloaded and time-spent measures were computed for each location. The raw simulation data were manually filtered using logic algorithms to remove artifact and isolate valid signal entries. Time-spent estimates were computed for each location based on the manually filtered RTLS movement histories. Time spent in a location was computed as the number of seconds that elapsed between the first timestamp recorded for a tag in a location and the first timestamp recorded for that tag in a new location.

Efficacy was assessed by the following indicators of method agreement using SPSS version 19: intra-class correlations (ICC), mean difference ($\bar{d}$), mean absolute difference (mean $|d|$), and 95% limits of agreement (95% LoA) (Bland & Altman, 1999; Shrout & Fleiss, 1979). Method agreement was quantified using the differences between time estimates obtained through direct observation and through RTLS (based on raw and filtered data). Differences between time-spent estimates were obtained by subtracting each RTLS estimate from the paired direct observation estimate. The absolute values of these differences were subsequently computed. Method agreement (concordance) is inversely related to the magnitude of the mean absolute difference between time-spent estimates. The direction and magnitude of method bias is reflected in the mean difference between the paired time-spent estimates: a positive mean difference would reflect under estimation of time spent by the RTLS method compared to direct observation, and a negative mean difference would reflect overestimation of time spent by the RTLS method. The Bland-Altman 95% LoA reflects the degree of variation in the differences between methods and defines the range within which 95% of the differences will lie. The 95% LoA is computed as the mean difference $\pm$ (1.96 $\times$ standard deviation).

Inter-Rater Reliability of RTLS Time-Use Estimates

Two analysts were trained in the application of the logic algorithm for manual filtering. An inter-rater reliability check was completed post training for quality control. Standardized guidelines regarding inter-rater reliability checks are not available, and no previous reports of inter-rater reliability for manually filtered RTLS time-use estimates were available to support a sampling plan. However, a community standard of 5% has been reported for studies involving chart abstraction and a minimal reliability threshold set at 0.75 (Liddy, Wiens, & Hogg, 2011; Reisch et al., 2003). An inter-rater reliability below this threshold triggers the following actions: repeating data collection, retraining data collectors, and more frequent reliability checks.

Because the logic algorithm used in this study had not previously been evaluated, a 10% sample of data filtered immediately following training was selected for reliability checking. Ten percent of the raw RTLS data generated by the nursing service personnel from the first 12-hour shift on the first unit was selected at random and treated to manual filtering by both analysts. Time-spent measures were computed based on RTLS data filtered by each analyst. Inter-rater reliability was assessed based on measures of rater (analyst) agreement for time-spent estimates using the same procedures used to assess method agreement. Analyst agreement was quantified using the differences between time estimates obtained after manual filtering by analysts A and B. Differences between time-spent estimates were

obtained by subtracting each RTLS time-spent estimate derived from the data filtered by analyst B from the paired RTLS time-spent estimate derived from the data filtered by analyst A. The absolute values of these differences were subsequently computed.

Analyst agreement (inter-rater agreement) is inversely related to the magnitude of the mean absolute difference between time-spent estimates. The direction and magnitude of analyst bias is related to the mean difference between the paired time-spent estimates: a positive mean difference would reflect underestimation of time spent by the analyst B compared to analyst A, and a negative mean difference would reflect overestimation of time spent by analyst B. The Bland-Altman 95% LoA reflects the degree of variation in the differences between analysts and defines the range within which 95% of the differences will lie.

Resource Requirements for RTLS Time-Use Estimates

Resource requirements to obtain RTLS time-use estimates include the following: RTLS equipment (hardware and software), human resources for installation and calibration, and human resources for manual data filtering and computation of time-use estimates. Data for resources related to RTLS equipment, installation, and calibration were obtained through interviews with an anonymous vendor. Data for resources required for manual filtering and computation of time-use estimates were based on analyst self-report.

Results

Efficacy of RTLS Time-Use Estimates

A sample of 68 distinct time-space events was generated during simulation. Time-spent estimates were obtained by continuous direct observation and RTLS using raw (RTLS-R) and filtered (RTLS-F) movement histories. Events in patient rooms ($n = 49$) and non-patient rooms ($n = 19$) were represented in the sample. RTLS-F time-spent estimates demonstrated superior agreement with direct observation estimates based on $\bar{d}$ (-3 to -22 seconds) and mean $|d|$ (13–65 seconds) compared to RTLS-R estimates, which underestimated time spent ($\bar{d} = 35$ –47 seconds; mean $|d| = 36$ –84 seconds) (Table 2; Fig. 1). Method agreement varied based on location category and RTLS filter status, with the highest agreement demonstrated by RTLS-F estimates for time spent in patient rooms ($\bar{d} = -3$; 95% LoA -37 to 31). Moreover, 75% of the estimates for time spent in patient rooms obtained by RTLS-F were within 21 seconds of those obtained through direct observation. RTLS-F estimates also demonstrated superior reliability (ICC = 0.69–0.99) compared to RTLS-R (ICC = 0.34–0.83), with the highest reliability demonstrated by RTLS-F estimates for time spent in patient rooms (ICC = 0.99).

Inter-Rater Reliability of RTLS Time-Use Estimates

A total of 271 distinct time-space events were identified following the filtering process by two analysts from the 10% random sample of RTLS-R movement history data generated in vivo. Events in patient rooms (65) and non-patient rooms (206) were represented. Perfect agreement was demonstrated in 23% of the estimates, and half of the estimates demonstrated agreement within a 19-second range (Table 3). Although the $\bar{d}$ across all locations was small (2 seconds; Fig. 2), there is evidence of some bias based on location because estimates by

analyst A were higher than those of analyst B for patient rooms ($\bar{d}$ = 58 seconds) and lower for non-patient rooms ($\bar{d}$ = -16 seconds). Consistency of agreement for time-spent estimates between analysts was high across all locations (ICC = 0.81–0.95).

Resource Requirements for RTLS Time-Use Estimates

The time requirement for manual data filtering for the in vivo sample was 30–90 minutes per staff member per 12-hour shift and varied based on movement patterns. Vendor-reported costs for RTLS equipment sufficient to support time-use monitoring on a single nursing unit was \$21,012, with an additional cost of \$4,350 for assistance with hardware installation (personal communication, July 31, 2012). The equipment can be reused across different nursing units, but vendor-supported installation costs recur for each separate unit.

Many vendors offer software programs designed to filter RTLS data as it enters the interrogator, so that entries likely attributed to artifact are decoded prior to being recorded and stored. Because filtering occurs before data storage, the raw data are not retained and therefore not available for subsequent analyses. The algorithms used for automated filtering are proprietary, and the software requires customized configuration for each nursing unit to incorporate information about unit geography. Vendor-reported costs (personal communication, July 31, 2012) for automated filtering include a renewable unit-based license (\$7,500 to filter one unit at a time) and configuration fees (\$12,000/unit). RTLS start-up costs to equip a single nursing unit for time-use estimation with automated filtering (\$44,862) are significantly greater than with manual filtering (\$25,362 plus labor).

Discussion

This pilot study provided evidence to support the efficacy of RTLS technology for time-use estimation among bedside nursing staff. RTLS time-use estimates demonstrated good concordance and reliability with those obtained through the gold standard of continuous direct observation. Moreover, the level of agreement documented in this study between direct observation and RTLS was superior to what has been previously reported for other methods (i.e., work sampling, self-report, and staffing indices).

These findings should be considered preliminary rather than confirmatory, because the sample size was small (68 time-space events), and the method comparison was conducted under simulated rather than live conditions. The differences noted between raw and filtered time-use estimates highlight the detrimental effects of artifact on the accuracy of time-use estimates derived from raw RTLS data. Therefore, optimization of RTLS as a method for obtaining time-use estimates requires effective strategies to reduce and filter artifact.

The manual filtering process applied in this study significantly reduced the bias demonstrated in RTLS-R time estimates and improved concordance and reliability with estimates obtained through continuous observation in the simulated setting. Moreover, when applied in vivo, agreement and consistency between analysts was high for the resulting RTLS-F time-use estimates. Some bias was noted between analysts, however. Much of the discordance was related to types of movement history sequences associated with greater uncertainty (e.g., Table 1, Room 8 to Room 7 sequence). Such discordance may be reduced

by further refinement of the algorithm to address highly uncertain sequences and by additional analyst training.

Although effective, the manual filtering process was labor-intensive; therefore, human resources must be considered when evaluating the cost-effectiveness of RTLS as a time-use methodology. Human resources can be reduced through enhanced efficiency of the manual filtering process or replacement with an automated filtering process. Enhanced efficiency for manual filtering can be achieved by reducing the volume of entries on a movement history requiring manual review for potential artifact (i.e., non-monitored area entries) and reducing the time required to distinguish between a true signal and noise or artifact entry. Two strategies may be pursued to achieve these goals: (1) equip unit hallways to detect tag transmissions; (2) develop alternative methods to affix tags to uniforms that will limit tag movement and prevent incidental blockage. Both of these strategies should reduce the volume of non-monitored area entries. Moreover, with hallway transmissions uniquely identified in the raw data, it should be easier and less time-consuming for analysts to distinguish between within-room artifact and true transit time. For studies involving repeated measures on the same nursing units over time, automated filtering would be more cost-effective.

Perhaps the most significant limitation of RTLS is the fact that the timestamps recorded in RTLS are linked to locations rather than activities. Although RTLS may yield precise (in seconds) and accurate estimates of time spent in specific locations such as patient rooms, this method does not directly capture time spent engaged in specific activities. However, in situations where location and activity are related, this limitation is less relevant. In the context of an inpatient nursing unit, patients are primarily confined to specific locations, that is, their assigned patient rooms. Thus, nurse-patient interaction is likewise confined. Therefore, a compelling argument to support the RTLS estimate of time spent in patient rooms as a surrogate measure of time use for the broad activity category of nurse-patient interaction can be made. Moreover, in this study, RTLS time-use estimates were most efficacious for patient room locations.

Conclusions and Implications

The results of this pilot study support proof of concept for continued experimentation with the application of RTLS technology to obtain time-use estimates for bedside nurses and particularly to quantify nurse-patient interaction time. RTLS time-use estimates demonstrate good concordance and reliability with direct observation. In comparison with conventional methods, the RTLS methodology offers: (1) increased efficiency through automated data collection and data entry; (2) less intrusiveness to nurse-patient interaction and workflow; (3) decreased vulnerability to response bias; (4) absence of respondent burden; and (5) opportunity for a precise unit of measure (i.e., seconds). The major limitations of RTLS include recurring data filtering costs and the linkage of time-spent estimates to locations rather than activities.

Acknowledgments

This research was supported by Grant Number UL1RR024982, titled “North and Central Texas Clinical and Translational Science Initiative” (Milton Packer, MD, PI) from the National Center for Research Resources (NCRR), and its contents are solely the responsibility of the author and do not necessarily represent the official view of the NCRR or NIH. Information on Re-engineering the Clinical Research Enterprise can be obtained from <http://nihroadmap.nih.gov/clinicalresearch/overview-translational.asp>. The authors thank Linda H. Yoder for assistance in the preparation of this manuscript.

References

- Buerhaus PI, Needleman J. Policy implication of research on nurse staffing and quality patient care. *Policy, Politics & Nursing Practice*. 2000; 1(1):5–15.
- Bland JM, Altman DG. Measuring agreement in method comparison studies. *Statistical Methods in Medical Research*. 1999; 8(2):135–160. [PubMed: 10501650]
- Bratt JH, Foreit J, Chen P, West C, Janowitz B, De Vargus T. A comparison of four approaches for measuring clinician time use. *Health Policy and Planning*. 1999; 14:374–381. [PubMed: 10787653]
- Burke TA, McKee JR, Wilson HC, Donahue RM, Batenhorst AS, Pathak DS. A comparison of time-and-motion and self-reporting methods of work measurement. *Journal of Nursing Administration*. 2000; 30(3):118–125. [PubMed: 10725940]
- Clarke SP. Registered nurse staffing and patient outcomes in acute care. Looking back, pushing forward. *Medical Care*. 2007; 45:1126–1128. [PubMed: 18007161]
- Collopy F. Biases in retrospective self-reports of time use: An empirical study of computer users. *Management Science*. 1996; 42:758–767.
- Currie V, Harvey G, West E, McKenna H, Keeney S. Relationship between quality of care, staffing levels, skill mix and nurse autonomy: Literature review. *Journal of Advanced Nursing*. 2005; 51:73–82. [PubMed: 15941463]
- Donaldson SI, Grant-Vallone EJ. Understanding self-report bias in organizational behavior research. *Journal of Business and Psychology*. 2002; 17:245–260.
- Fahey L, Lopez KD, Storfjell J, Keenan G. Expanding potential of radiofrequency nurse call systems to measure nursing time in patient rooms. *Journal of Nursing Administration*. 2013; 43:302–307. [PubMed: 23615373]
- Finkler SA, Knickman JR, Hendrickson G, Lipkin M, Thompson WG. A comparison of work-sampling and time-and-motion techniques for studies in health services research. *Health Services Research*. 1993; 28:577–597. [PubMed: 8270422]
- Heinz D. Hospital nurse staffing and patient outcomes: A review of current literature. *Dimensions of Critical Care Nursing*. 2004; 23:44–50. [PubMed: 14734900]
- Hendrich A, Chow MP, Skierczynski BA, Lu Z. A 36-hospital time and motion study: How do medical-surgical nurses spend their times? *Permanente Journal*. 2008; 12(3):25–34.
- Hunting KL, Haile E, Nessel L, Welch LS. Validity assessment of self-reported construction tasks. *Journal of Occupational and Environmental Hygiene*. 2010; 7:307–314. [PubMed: 20349389]
- Jones T. Validation of the perceived implicit rationing of nursing care (PIRNCA) instrument. *Nursing forum*. 2013
- Jones TL. Radiofrequency identification. Exploiting an old technology for measuring nurse time and motion. *Computers, Informatics, Nursing*. 2012; 30:463–472.
- Juster FT, Ono H, Stafford FP. An assessment of alternative measures of time use. *Sociological Methodology*. 2003; 33:19–54.
- Kalisch BJ, Landstrom G, Williams RA. Missed nursing care: Errors of omission. *Nursing Outlook*. 2009; 57:3–9. doi:10.1016/j.outlook.2008.05.007. [PubMed: 19150261]
- Kane RL, Shamliyan TA, Mueller C, Duval S, Wilt TJ. The association of registered nurse staffing levels and patient outcomes: Systematic review and meta-analysis. *Medical Care*. 2007; 45:1195–1204. [PubMed: 18007170]
- Larson EL, Aiello AE, Cimiotti JP. Assessing nurses hand hygiene practices by direct observation or self report. *Journal of Nursing Measurement*. 2004; 12:77–85. [PubMed: 15916321]

- Liddy C, Wiens M, Hogg W. Methods to achieve high inter-rater reliability in data collection from primary care medical records. *Annals of Family Medicine*. 2011; 9:57–62. [PubMed: 21242562]
- Lin K. Revisiting the gap between stylized and diary estimates of market work time. *Social Science Research*. 2012; 41:380–391. [PubMed: 23017759]
- Manson TM, Levine EL, Brannick MT. The construct validity of task inventory ratings: A multitrait-multimethod analysis. *Human Performance*. 2000; 13:1–22.
- Norrish BR, Rundall TG. Hospital restructuring and the work of registered nurses. *Milbank Quarterly*. 2001; 79:55–79. [PubMed: 11286096]
- Numata Y, Schulzer M, van der Wal R, Globberman J, Semeniuk P, Balka E, Fitzgerald JM. Nurse staffing levels and hospital mortality in critical care settings: Literature review and meta-analysis. *Journal of Advanced Nursing*. 2006; 55:435–448. [PubMed: 16866839]
- Ogunfeditimi F, Takes L, Paige VJ, Wyman JF, Marlow E. Assessing the productivity of advanced practice providers using a time and motion study. *Journal of Healthcare Management*. 2013; 58:173–185. [PubMed: 23821897]
- Otterbach S, Sousa-Poza A. How accurate are German work-time data? A comparison of time-diary reports and stylized estimates. *Social Indicators Research*. 2010; 97:325–339.
- Pape ES. A model for analysis of work sampling statistics. *IEE Transactions*. 1992; 24(5):134–138. doi: org/10.1080/07408179208964253.
- Pelletier D, Duffield C. Work sampling: Valuable methodology to define nursing practice patterns. *Nursing and Health Statistics*. 2003; 5:31–38.
- Robinson MA. Work sampling: Methodological advances and new applications. *Human Factors and Ergonomics in Manufacturing & Service Industries*. 2010; 20(1):42–60.
- Reisch LM, Fosse JS, Beverly K, Yu O, Barlow WE, Harris EL, Elmore JG. Training, quality assurance, and assessment of medical record abstraction in a multisite study. *American Journal of Epidemiology*. 2003; 157:546–551. [PubMed: 12631545]
- Rutherford, P., Lee, B., Greiner, A. Transforming care at the bedside. IHI Innovation Series white paper. Boston, MA: Institute for Healthcare Improvement; 2004. Retrieved from <http://www.IHI.org>
- Schubert M, Glass TR, Clarke SP, Aiken LH, Schaffert-Witvilliet B, Sloane DM, De Geese S. Rationing of nursing care and its relationship to patient outcomes: the Swiss extension of the International Hospital Outcomes Study. *International Journal for Quality in Health Care*. 2008; 20:227–237. [PubMed: 18436556]
- Shrout PE, Fleiss JL. Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*. 1979; 86:420–428. [PubMed: 18839484]
- Spetz J, Donaldson N, Aydin G, Brown DS. How many nurses per patient? Measurements of nurse staffing in health services research. *Health Services Research*. 2008; 43:1674–1692. [PubMed: 18459953]
- Thungjaroenkul P, Cummings GC, Embleton A. The impact of nurse staffing on hospital cost and patient length of stay: A systematic review. *Nursing Economics*. 2007; 25:255–265. [PubMed: 18080621]
- Upenieks VV, Akhavan J, Kotlerman J, Esser E, Ngo MJ. Value-added care. A new way of assessing nursing staffing ratios and workload variability. *Journal of Nursing Administration*. 2007; 37:243–252. [PubMed: 17479044]
- Ver Ploeg, M, Altonji, J, Bradburn, N, Da Vanzo, J, Nordhaus, W., Samaniego, F., editors. Time-use measurement and research. Report of a workshop. Washington, DC: National Academy Press; 2000.
- Weigl M, Müller A, Zuoanc A, Angerer P. Participant observation of time allocation, direct contact and simultaneous activities in hospital physicians. *BMC Health Services Research*. 2009; 9:1–11. [PubMed: 19118494]

FIGURE 1.

Time spent: Difference (observation–filtered RTLS) versus mean values with 95% limits of agreement. RTLS, real time location system; *SD*, standard deviation.

FIGURE 2.

Time spent: Difference (analyst A–analyst B) versus mean values with 95% limits of agreement. RTLS, real time location system; *SD*, standard deviation.

Table 1

Steps of Real Time Location System (RTLS) Manual Filtering Process

Step I: Identify Sequence Patterns in RTLS-R Movement History	Step II: Identification of Probable Artifact	Step III: Recode Probable Artifact Entries to Create RTLS-F Movement History
Room 1	Room 1	Room 1
Non-monitored area	Probable within room artifact	[Entry filtered out]
Room 1	Room 1	[Entry filtered out]
Non-monitored area	Probable within room artifact	[Entry filtered out]
Room 1	Room 1	[Entry filtered out]
Non-monitored area	Probable true signal	Non-monitored area—hallway
Room 5	Room 5	Room 5
Non-monitored area	Probable true signal	Non-monitored area—Hallway
Room 10	Room 10	Room 10
Room 9	Room 9	Room 9
Room 8	Room 8	Room 8
Non-monitored area	Within room artifact or true signal	Room 8, hallway, or room 7
Room 7	Room 7	Room 7

Note. RTLS-R, raw RTLS data; RTLS-F, manually filtered data.

Table 2

Summary of Method Agreement Indicators

	Observation vs. Raw RTLS			Observation vs. Filtered RTLS		
	All Locations	Patient Rooms	Non-Patient Rooms	All Locations	Patient Rooms	Non-Patient Rooms
Events (<i>n</i>)	68	49	19	68	49	19
Time spent (seconds)						
$\bar{d}$ (<i>SD</i>)	38 (85.24)	35 (45.38)	47 (146.44)	-8 (65.60)	-3 (17.44)	-22 (122.13)
95% LoA	-129 to 205	-54 to 124	-240 to 334	-137 to 121	-37 to 31	-261 to 217
Mean $ d $	50	36	84	28	13	65
Quartiles (seconds)						
25%	11	5	22	5	5	11
50%	26	23	30	13	10	22
75%	51	51	68	23	21	64
ICC	0.63	0.83	0.34	0.85	0.98	0.69

Note. RTLS, real time location system; *SD*, standard deviation; *d*, mean difference; 95% LoA, 95% limits of agreement; *Mean |d|*, mean absolute difference; ICC, intra-class correlation coefficient.

Table 3

Summary of Inter-Rater (Inter-Analyst) Agreement Indicators

	All Locations	Patient Rooms	Non-Patient Rooms
Events (<i>n</i>)	271	65	206
Time spent (seconds)			
$\bar{d}$ (<i>SD</i>)	2 (111.90)	58 (104.66)	−16 (108.48)
95% LoA	−217 to 221	−147 to 263	−229 to 197
<i>Mean</i> <i>d</i>	56	58	55
Quartiles (seconds)			
25%	1	0	5
50%	19	15	20
75%	51	73	51
Absolute agreement (%)	23	28	22
ICC	0.94	0.81	0.95

Note. *SD*, standard deviation; 95% LoA, 95% limits of agreement; $\bar{d}$, mean difference; *Mean* |*d*|, mean absolute difference; ICC, intra-class correlation coefficient.