

EXECUTIVE SUMMARY

Measurement has been at the center of Transforming Care at the Bedside (TCAB), Phase II. It was to serve many functions: diagnostic, motivational and evaluative. Measurement has also been a challenge for the hospitals, the Institute for Healthcare Improvement (IHI) and the evaluation team. This report summarizes the findings from the evaluation team on the experience collecting and using the TCAB measures and focuses on the use of measures for diagnostic and motivational purposes.

Of the eight non-business case measures, only **falls** and **voluntary turnover** were reported to be collected as originally defined by most hospitals, easy to collect. Of these, two measures, only **falls** was used by the hospitals to identify or refine tests of change. **Team satisfaction, pressure ulcers, and patient satisfaction** were the other measures judged most useful for identifying or refining tests of change, with all three in the middle range of difficulty for data collection, and all except patient satisfaction collected as defined by almost all hospitals. Despite the difficulties in getting the **PDA measures** up and going, almost all hospitals reported using at least one measure to identify tests of change. **Deaths among surgical inpatients** was reported as being of middle difficulty to collect, but like **voluntary turnover**, was least useful in identifying or refining tests of change.

As we review the measures in the initial data set for coverage of the four domains, we find that there are two measures of **safety and reliability** (falls and pressure ulcers) that have been widely used. **Team vitality** has been assessed using two measures (voluntary turnover and the team satisfaction), and both have been widely used. There have been site-to-site variations in administration of the team survey that make comparability over time and across sites difficult. **Efficiency**, which in the TCAB context is better described as **shift to higher value work** to reflect the belief that staff hours should not be reduced but redirected, is measured by the PDAs. Although there has been considerable resistance to this measure, several interviewees expressed a desire to continue collecting it. **Patient centeredness** has morphed to some extent into patient satisfaction, and there are substantial variations in the administration of the short survey instrument across sites and many sites that used alternative measures. Patient satisfaction is believed to be among the most important TCAB measures as it was the most common response to our question: *“If you were asked to demonstrate to the outside world that TCAB had any value what measures would you collect?”*

Introduction

Measurement has been at the center of TCAB Phase II and was intended to serve three main functions:

- **Diagnostic:** To help units assess the need for improvement in a domain and provide the basis for further drill down and assessment of improvement opportunities.
- **Motivational:** To focus staff on the need for change, provide evidence that improvement was feasible, and maintain attention on activities that kept the numbers moving in the right direction.
- **Evaluative:** Assessing the extent to which there had been change in each of the four TCAB domains.

In an effort to understand the experiences with, and use of measures in the 13 hospitals participating in TCAB Phase II, the evaluation team conducted telephone interviews with nurse managers and QI staff. These interviews were conducted in March-April 2006, as Phase II was drawing to a close. At least one interview was conducted at each hospital. A total of twenty-seven interviews were conducted (fourteen QI staff or person responsible for data collection and thirteen Nurse Managers or proxy). At facilities with multiple TCAB units, we interviewed only one unit manager. We report responses at the hospital level (n=13).

Interviews examined which measures were collected, the fidelity to collecting measures as defined by the Institute for Health Care Improvement (IHI), the ease of collection, and how measures were used. This report focuses on the first two uses of measurement (diagnostic and motivational).

This report is organized as follows:

- 1: Measures collected and definitions
- 2: Ease of collecting each measure
- 3: How measures were used
- 4: Summary of findings by measure
- 5: Measures collected by PDA
- 6: Measures units identified as demonstrating TCAB value
- 7: Summary

Section 1: Measures collected and definitions

A core set of measures was to be collected and posted to the extranet at fixed periods. It was anticipated that units would identify and carry out additional measurement as part of planning, conducting and assessing tests of change (TOCs). Measurement has been a challenge for the hospitals, IHI and the evaluation team. After

an initial proposed measurement set was challenged by participating hospitals, the design team working with hospital representatives revised the measure set. The final measurement set consisted of fourteen measures (Table 1). The business case measures were to be examined separately from the unit-oriented measures.

Table 1. Measures Collected for TCAB, Phase II

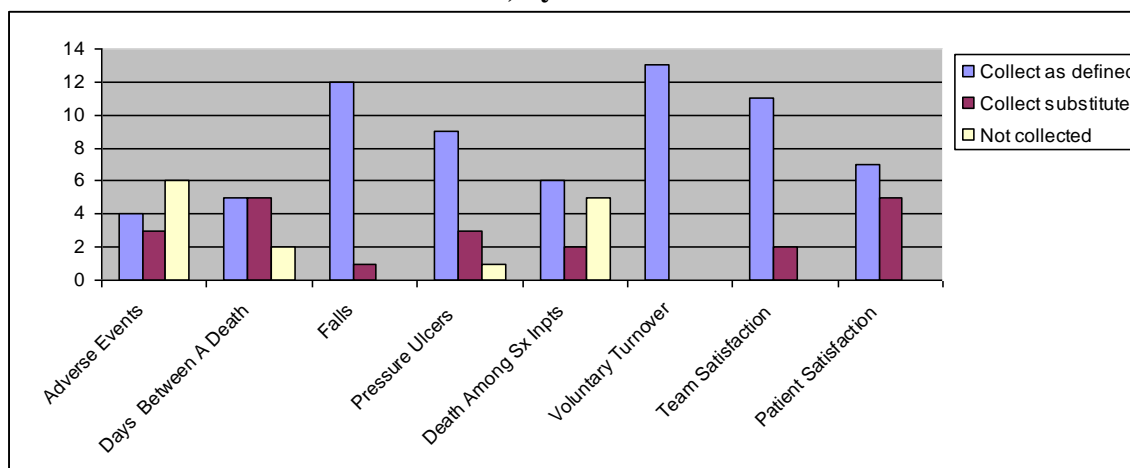
<p>Measures Collected by Chart Audit or Survey Or Incident Reporting Mechanisms</p> <ul style="list-style-type: none"> • Adverse events • Days between a death • Death among surgical inpatients • Falls (per 1000 pt days) • Patient satisfaction • Pressure ulcer (prevalence) • Voluntary turnover • Team satisfaction
<p>Measures Collected by PDA</p> <ul style="list-style-type: none"> • % time in documentation • % time in direct patient care • % time in value added work
<p>Business Case Measures (Collected from Administrative Data)</p> <ul style="list-style-type: none"> • Direct costs per adjusted patient day • Average length of stay • Nursing hours per adjusted patient day

The interview protocol included a series of detailed questions about each measure. First, we asked if each measure was collected, and, for each measure collected, we asked if it was collected as defined by IHI or if a substitute was collected. Finally, we asked for suggestions for alternative measures. A list of substitutes or modifications to existing measures is presented in Appendix A.

Of the eight measures collected by chart audit, survey, or incident reporting, three measures (falls, voluntary turnover, and team satisfaction) were collected by all thirteen sites (Table 2). These measures were most likely to be collected as defined by IHI.* Patient satisfaction and days between a death were the two measures for which a substitute was most frequently collected. Adverse events and death among surgical inpatients were the measures most frequently reported as not collected.

* Subsequent to this survey, it was determined that hospitals were uncertain about which staff departures were to be excluded under the exclusion rules for the voluntary turnover measure, and did not consistently report pre-TCAB turnover. Both issues are being addressed via revised reporting.

Table 2. Adherence to data collection, by measure*



* N= 13 for all measures except patient satisfaction and days between a death (n=12)

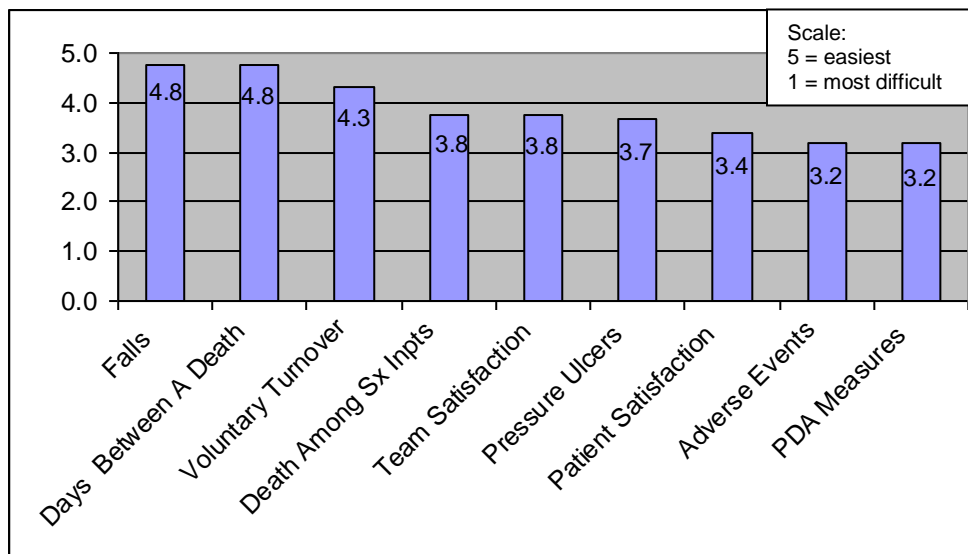
Section 2: Ease of collecting each measure

Interviewees were asked to rate the difficulty/ease of collecting each measure on a scale of 1 to 5 with the following definitions:

- 1 = so difficult we did not collect
- 2 = an ongoing difficulty
- 3= somewhat difficult, but we figured it out
- 4= fairly easy, we had infrastructure but needed to tweak
- 5= easy to collect, we already had

Falls and days between a death ranked as the easiest (mean score 4.8) followed by voluntary turnover (mean score 4.3). It should be noted, however, that there was a high degree of substitution for days between a death. Five hospitals collected the days between a death measure as defined, while five others collected a substitute measure. Death among surgical inpatients, team satisfaction, and pressure ulcers were moderately easy to collect (mean scores ranged from 3.4 – 3.8). The measures ranked as most difficult to collect were adverse events and the PDA measures (mean scores 3.2) (Table 3).

Table 3. Ease of data collection



Note: For each measure, interviewees were asked: “How difficult or easy was [measure] to collect at the unit level?” Key: 1 = so difficult we did not collect; 2 = an ongoing difficulty; 3= somewhat difficult but we figured it out; 4= fairly easy, we had infrastructure but needed to tweak; 5= easy to collect, we already had. Scores represent means across interviewees.

Section 3: How measures were used

Measurement can be a useful tool not just as a means to compare hospitals’ performance, but as a means to inform both the identification of and utility of tests of change. Interviewees were asked if and how each measure was used, both in conjunction with testing changes (i.e., for identification of tests of change, for refining tests of change) and also whether they were used in the development of a business case (Table 4).

Identification of Tests of Change

All measures, with the exception of days between a death, were used to some extent to identify potential tests of change. The measures used most frequently used in this way were falls, team satisfaction, pressure ulcer prevalence, patient satisfaction, and adverse events.

Refining Tests of Change

Interviewees reported using measures for refining tests of change less frequently than they did for identification of tests of change. The two measures most commonly reported as used for refining tests of change were falls and team satisfaction.

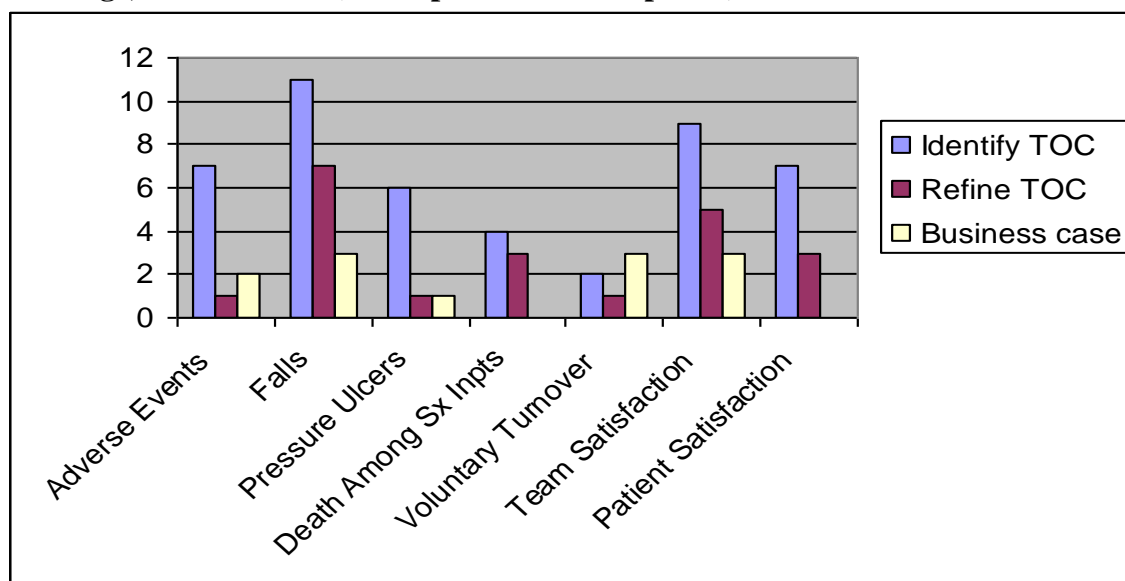
Business Case

Six of the hospitals reported using (or planning to use) at least one measure to develop a business case. Measures for which a business case was reported to be either planned or completed were falls, voluntary turnover, team satisfaction, adverse events and pressure ulcers. One site stated they were considering using the measure of nursing hours per adjusted patient day to create a business case. One site planned to do a business case on surgical site infections, a measure they were collecting in addition to TCAB-required measures.

Measures not used

Only one measure, days between a death, was not used for identifying tests of change, refining tests of change or for developing a business case.

Table 4. Usefulness of measures for: identification of tests of change, refining tests of change, business case (either planned or completed)



Note: Not all places collected all measures. Days between a death (not shown) was not used for any of these purposes

Section 4: Summary of findings by measure

Falls

Frequency of reporting: Monthly

Falls were measured as the number of inpatient falls (x 1000)/total number of inpatient days. Nearly all sites (n=12) reported collecting this measure as defined by IHI staff; one site reported collecting incidence (not prevalence). The falls measure was rated as very easy to collect 4.8/5.0. Most facilities reported that collecting a unit-level

measure of falls pre-dated TCAB. Five facilities stated they had a falls initiative separate from TCAB. This measure was reported to be overall most useful as nearly all (n=11) reported using this to identify tests of change and more than half (n=7) reported using the measure to refine tests of change. This measure was also used as a motivator for staff. For example, “days between falls” was posted on some units to create awareness of falls. Two hospitals reported using falls as the topic for a completed or planned business case.

Team Satisfaction

Frequency of reporting: Twice

Staff satisfaction was collected using an 11-item tool which all unit staff (e.g., RN, MD, dietary personnel, unit clerk) were asked to complete. All sites reported collecting this as defined by IHI. Most sites (9/13) reported using this measure to identify tests of change. This was rated as the second most useful measure for this purpose. Only two sites reported using it to refine tests of change. This was rated as fairly easy to collect (4.3/5.0).

One concern about this measure was the type of staff asked to complete the tool. Most hospitals focused on nursing staff for completion of this, however, at least one interviewee stated that physicians were asked to complete the tool and noted that they do not identify with the unit as their home base and were answering this based on their office environment.

This measure was spread in at least one hospital.

Voluntary Turnover

Frequency of reporting: Monthly

Voluntary turnover was measured using the National Quality Forum (NQF) protocol which measures uncontrolled separations for RNs and advanced practice nurses. All thirteen sites collected this measure as defined by IHI although about half (6) said this was a new measure for them. This measure was relatively easy to collect, although as noted above, it was subsequently determined that sites were not consistently applying the voluntary turnover exclusion rules. Sites rated it 4.3/5.0. This measure was used to a limited extent in identifying tests of change; only two sites reported using the measure in this way, while one site reported using the measure to refine tests of change. This was useful in creating a business case. Three sites reported using it in their business case.

Pressure Ulcer Prevalence

Frequency of reporting: Three times

Pressure ulcer prevalence was collected in all but one facility. Most (n=9) reported collecting this as defined by IHI; and three reported collecting a substitute. About half the sites (5) reported that this was an existing measure while (n=4) reported that was a new measure. This was useful in identifying areas for improvement with over half (n=6) reporting that they used this measure to identify areas for improvement. It was

rarely used to refine tests of change (n=1) and used in one site as the basis of their business case.

Patient Satisfaction

Frequency of reporting: Monthly

Patient satisfaction was collected using a 4-item tool. Facilities were almost split on collecting this as defined (7/12) versus reporting a substitute that was already collected (5/12) and felt to be an equivalent measure. Among the seven sites that collected and reported the 4-question tool, this was rated as moderately difficult to collect (3.4/5.0).

Some measure of patient satisfaction was used by seven hospitals to identify tests of change; and by three to refine tests of change. None reported using it for a business case.

Six of the sites reported that they customized the 4-item tool. Examples of customization included: adding specific questions pertaining to innovations they were testing at the time and adding questions such as “willingness to recommend”.

One hospital reported spreading the 4 item tool to other units.

Days Between a Death

Frequency of reporting: Monthly

Days between a death was collected by plotting the days between a death every time a death occurs. Sites were asked to plot this monthly, even if a death did not occur. This measure was collected as defined by IHI in 7 sites; a substitute was collected in 5 sites. This measure was ranked as very easy to collect (4.8/5.0). However, it was not used for identifying tests of change, refining tests of change, nor for a business case.

Adverse Events /1,000 Patient Days

Frequency of reporting: Monthly

Sites were asked to collect this measure using a trigger tool for a more global measure of adverse events or a tool that specifically assesses adverse drug events. Only seven hospitals reported collecting this measure, of these, it was a new measure for all but one and three reported collecting a substitute. This measure was rated among the most difficult of all the measures to collect (3.2/5.0). However, it appears to have been useful for identifying tests of change: all facilities collecting this data reported using this it to identify change activities.

Death Among Surgical Inpatients

Frequency of reporting: Three times

This measure was also known as “unanticipated deaths” or “failure to rescue”. Sites were asked to report the NQF protocol for this measure. This measure was rated as

easy to collect (4.8/5.0). However, it was not used for identification of tests of change, refining tests of change, or for a business case.

Section 5: Measures Collected By PDA

Frequency of reporting: Monthly

Three measures were derived from the PDA data: (1) percent time in documentation, (2) percent time in direct patient care, and (3) percent time in value added work. This measure was rated among the most difficult to collect (3.2/5.0). The challenges primarily fell into one of four areas (Table 5).

Table 5. PDA Data Collection Challenges

Challenge	Count
Difficulties with software & data	8
Nurse participation in data collection	8
Mechanical problems or battery charge	5
Confusion about coding	5

Note: N=13

Challenges with software and data were noted by at least one interviewee in the majority of hospitals. Examples included difficulties getting the software installed onto the hospital's computer system due to incompatibility with network interface or lack of administrator rights on the unit computer and difficulties downloading data. Another issue, specific to hospitals with more than one TCAB unit, involved not having multiple licenses to install the software. One interviewee noted there is no ownership of the PDA data collection—meaning that there was no one to turn to for technical assistance if things broke or were not working.

Another commonly mentioned challenge was achieving nurse participation in collecting the PDA data. This was reported by at least one interviewee in the majority of hospitals. Many interviewees reported that considerable effort is required to gain staff compliance in carrying the devices. The barriers to achieving staff nurse participation in PDA data collection included: frustration with the equipment, nurse concerns that the PDAs were extra work, the distraction from patient care, and that the devices were cumbersome to carry.

Mechanical difficulties such as the PDAs not holding a charge was another commonly cited challenge.

Finally, confusion about coding, or concerns about the accuracy of the data were also mentioned. One interviewee reported that the nurses said they were not careful about how they coded their time, although she felt that their results seemed accurate. At least one interviewee said that they did not use these measures to identify or refine tests of change because they did not think they could trust the data.

We asked if the PDA data collection was “currently functioning as planned” and 10/12 hospitals reported that it was functioning (or is mostly functioning) as planned. Additional detail on this, such as how many different nurses are carrying the PDA, might

be useful to assess the validity of the data. Interviewees at five hospitals noted that these were important measures and that they wished to continue collecting them. One hospital has bought additional PDAs and has built PDA data into their spread work.

In eleven of the thirteen hospitals, at least one of the PDA measures was used to identify tests of change. The measures of time in direct patient care and percent time in documentation were used to identify tests of change (6/12 hospitals reported using these measures in this manner); the measure of value added care was used less frequently to identify tests of change (3/13 hospitals).

Section 6: Measures units identified as demonstrating TCAB value

Interviewees were asked: *“If you were asked to demonstrate to the outside world that TCAB had any value what measures would you collect?”* Responses are summarized in Table 6. The two measures that interviewees believed to be most important indicators of TCAB success were patient satisfaction and staff satisfaction.

Table 6. Measures that are the most important indicators of TCAB success

MEASURE	# mentions (n=27)
Patient Satisfaction	8
Staff Satisfaction	7
Measures of specific prototypes and measures of nursing processes	4
Deaths	3
Patient safety/medication errors	3
Cost	3
Falls	3
Adverse events	2
Pressure ulcers	2
Nursing hours	2
PDA Data	2
Average length of stay	2

Note: Other measures mentioned by one interviewee were: nurse autonomy, nurse empowerment, failure to rescue, coordination of care, unit-specific measures which reflect the unit’s patient population, readmission rates, core measures set, nurse turnover, medication errors, the number of innovations spread, unplanned transfers to the ICU, and pneumonia.

Section 7: Summary

Of the eight non-business case measures, only **falls** and **voluntary turnover** were reported to be collected as originally defined by most hospitals, easy to collect, and of these, **only falls** was used by the hospitals to identify or refine tests of change. **Team satisfaction, pressure ulcers, and patient satisfaction** were other measures judged most useful for identifying or refining tests of change, with all three in the middle range of difficulty for data collection, and all except patient satisfaction collected as defined by almost all hospitals. Despite the challenges with **PDA use and measures**, almost all hospitals reported using at least one measure to identify tests of change. **Deaths among surgical inpatients** was of middle difficulty to collect, but like **voluntary turnover**, was least useful in identifying or refining tests of change.

As we review the measures in the initial data set for coverage of the four domains, we find that there are two measures of **safety and reliability** (falls and pressure ulcers) that have been widely used. **Team vitality** has been assessed using two measures—voluntary turnover and the team satisfaction. While these two measures were used equally to create business cases, the voluntary turnover measure was used far less frequently than the team satisfaction measure for identifying and refining tests of change. As we move to examining the use of the team satisfaction data as the basis for comparing across facilities, variations in the composition of the sample across hospitals and variations in who completes the survey on each unit over time make comparisons between hospitals and within units problematic. **Efficiency**, which in the TCAB context is better described as **shift to higher value work** to reflect the belief that staff hours should not be reduced but redirected, is measured by the PDAs. Although there has been considerable push back on this measurement, several interviewees expressed a desire to continue collecting it. **Patient centeredness** has morphed to some extent into patient satisfaction, and there are substantial variations in the administration of the short survey instrument across sites and many sites used an alternative measure. Patient satisfaction is believed to be among the most important TCAB measures as it was the most common response to our question: *“If you were asked to demonstrate to the outside world that TCAB had any value what measures would you collect?”*.

APPENDIX A

SUGGESTED SUBSTITUTES/MODIFICATIONS TO EXISTING MEASURES

MEASURE	SUBSTITUTE OR MODIFICATION
Adverse Events or Adverse Drug Events	<ul style="list-style-type: none"> • Prefer to use our own (customized) Adverse Events tool • Report close calls which we monitor with incident reports • Reporting NDNQI because going for Magnet • Real-time auditing of indicators would be better but time consuming
Days between a death	<ul style="list-style-type: none"> • Transfers to higher level of care • Number of deaths per month easier and more useful than days between (said twice) • Prefer to measure readmission rates: 3 days for surgical pt discharge and 30 day readmission rates • Potentially preventable process • Raw mortality data
Falls	<ul style="list-style-type: none"> • Falls rate (per 1000 pts) not raw # of falls • Fall incidence
Voluntary Turnover	<ul style="list-style-type: none"> • Change exclusion rules: Relocation and child care important reasons people leave • Collect quarterly not monthly
Team satisfaction	<ul style="list-style-type: none"> • Move to a quarterly survey • Too long. Would prefer 5 questions, 6 at most • Have used 2 other surveys for vitality: Magnet survey every other year. If continue to use, would like survey every year
Pressure Ulcers	<ul style="list-style-type: none"> • Pressure ulcer incidence
Patient Satisfaction	<p>Picker (used in at least 2 places) Press Ganey (used in 3 places)</p>
Death Among Surgical Inpatients	<p>No suggestions</p>