Using Point of Service Clinical Documentation to Reduce Variability in Charge Capture

Sidney N. Thornton, Ph.D.[1], Hong Yu, M.D.[1], Reed M. Gardner, Ph.D.[1,2]
[1]Department of Medical Informatics, Intermountain Health Care, Salt Lake City, Utah
[2]Department of Medical Informatics, University of Utah, Salt Lake City, Utah

Data collected at bedside to document patient care can also be used to generate an itemized summary of charges including activity-based clinician charges. This approach becomes advantageous when the charge capture operation is transparent to the clinician who would otherwise have to review the care documentation, recall the appropriate charging rules, and exercise discretion in capturing charges. Documented procedures and supplies convert directly into patient charge rules. Documented patient care is more difficult to translate into activity-based charges because nursing care can vary in intensity and duration depending on the patient's needs. The problem can be overcome by embedding time or data-driven logic into the charging rules. Using this approach in the labor and delivery units of 7 IHC hospitals (114 beds), we generated consistent charge summaries. We improved the accuracy of patient charges from 65% to over 98% of our charge summaries having no missed charges.

INTRODUCTION

In most instances, hospital financial systems predate real-time clinical documentation systems, and because the data must be entered twice, the clinical and financial records often disagree. In this paper we demonstrate that the clinical documentation entered at the patient bedside (point of service) can also trigger algorithms that determine patient charges. The result of single data entry at the point of service is less duplicate effort and more accurate charge capture.

The Women and Newborns Clinical Program at Intermountain Health Care (IHC), representing hospitals throughout Utah and southern Idaho, has noted unacceptable error and variability in the itemized charges for patient care given during labor and delivery. The magnitude of the problem was described by an internal audit which found 35% of the itemized charge summaries for labor and delivery had at least one missed charge, 56% had at least one charge that was not found in the patient care documentation, and 95% had at least one discrepancy for charges related to the intensity of nursing activity. Patients receiving the same nursing care received different charges depending on how the nursing activity was interpreted for the charge summary. The variability was seen both within each hospital as well as between hospitals. More complex labor and deliver care procedures generated a wider range of nursing activity charges.

Analysis of the charge capture process for labor and delivery illuminated several key points that rely on clinician memory, judgment, and in many situations, arbitrary discretion. We noted that the clinicians would document patient care within the STORKBYTES[1] Clinical Information System (CIS) at the point of service, and would then duplicate the appropriate entries into the itemized charge summary according to their interpretation of what was chargeable. At a later point, the labor and delivery record would be reviewed once again by the billing clerk for additional itemized charges, indicating another point in the process requiring clinician discretion and judgment.

The approach we selected to reduce errors and to eliminate the variability arising from these discretion points was to embed logic within the STORKBYTES application that would capture charges systematically according to the documented patient care. The clinician could then focus on accurate and complete documentation of the care provided without worrying about the charge capture consequences. Patients who received equivalent care would receive the same charges as determined by the charging algorithm, regardless of who gave the care or in which hospital the care was given. This is consistent with best practice CIS design [2-4] and the

Institute of Medicine recommendation[5] for error prevention by collecting data once at the point of service to facilitate both clinical and business needs.

METHODS

Like most integrated health care delivery companies, IHC has combined its charging rules into a central Charge Master that is valid for all hospitals. The Charge Master for labor and deliver defines acuity charging rules for nursing activity based on the level of care provided to the patient.[6] Many CIS products[7-9] including IHC's HELP hospital-centric CIS[10-11] have procedure and supply cost capture functions but do not universally support acuity-based charge capture. The crux of our project lies in the successful conversion of these acuity rules into charging algorithms. These algorithms rely on the data elements charted by the nurses in STORKBYTES to document patient care and status. The intensity of nursing care is then computed by the algorithms rather than directly charted. Simply charting "intensive nursing care" requires the clinician to understand how the Charge Master defines intensive nursing care, and defeats the purpose of removing clinician discretion and judgment from charge capture.

For the charging algorithms to work, an agreement must be reached among all the clinical and administrative teams as to which charted data elements represent comparable intensity levels of nursing activity. In many cases agreement is easily reached, particularly where a clinical activity has a predictable effort and duration. For example, starting and maintaining a patient IV requires a certain predictable level of clinical activity that can be captured from a number of documented entries such as "IV started" or "IV hydration maintained."

In contrast, the agreement process is not as straightforward for fetal distress and other complications of labor, which vary greatly from patient to patient. Our approach to this dilemma has been to evoke time windows that will open and close based on the charted data. Using the fetal distress example, a 30-minute time window will open as soon as the first sign of fetal distress is charted. If in that 30-minute window the fetal distress is documented to be resolved, the time-window will close and no charge will be generated. On the other hand, if within the 30-minute time window there is neither resolution nor additional charted criteria for fetal distress, then the charge will be generated and another 30-minute window initiated. For other complicated clinical activities we have created data-driven windows that actively search for corroborating data to justify the charge generation. Determining the specific, computable criteria for opening and closing time or data-driven windows has been difficult and has required some compromise, particularly when building consensus among several hospitals. We estimate that 2/3 of the total pre-pilot development effort was spent on having nurses reach valid and testable agreements for nursing acuity.

Currently we have implemented the software as a pilot project in 7 of the 10 IHC hospitals using STORKBYTES. These 7 hospitals participate in more than 20,000 deliveries annually. All ten hospitals are scheduled to complete the pilot and evaluation phases of the project during 2002. This paper reflects a mid-pilot assessment of our approach. We examined the effectiveness of our approach in reducing errors and variability in the itemized charge summary. A separate study is being conducted to evaluate the cost consequences of this implementation and to address whether the patient care documentation accurately reflects the care provided including lost charges due to insufficient documentation.

This study assesses the agreement between charge capture and the clinical documentation. We conducted chart reviews using IHC's Charge Master as the standard. The coded data entries in STORKBYTES had been previously tested and validated to trigger the appropriate charges and are continually reviewed by an inter-hospital standards committee. In this study we examined the free-text data entries, looking for clinical findings that should have been entered as coded data. Free-text entries circumvent the charge capture algorithms and result in missed charges. For this study we distinguish free-text entry as data typed into the "other" or "comment" fields, and differentiate these entries from data values that are entered as coded.

For a given chart, the free-text entries were separated from the coded entries. The free-text entries were manually reviewed by two experts and sorted by whether they impacted a Charge Master rule. A charge-related free-text entry for which there was an equivalent coded entry was not counted as a missing charge. A charge-related free-text entry that did not have a redundant coded entry to trigger the appropriate charge was counted as a missing charge.

For this study 10 cases were selected for chart review from each of the 7 hospitals currently pilot-testing the software (n = 70). We selected the 10 most recent cases at each facility for which high-intensity nursing activity was provided, and we disregarded cases with only medium or low intensity nursing activity. We selected high-intensity nursing activity because it included all but the most extreme complications of labor, and required the greatest amount of interpretation and agreement on the charging rules. The average and standard deviation was computed for both the individual hospitals and for the all of the hospitals combined.

We estimated nurses perceived confidence of the accuracy of patient charges using the "manual review" flag in STORKBYTES. If the itemized charge summary seemed suspicious to the nurse, or the merits of the case were highly unusual, the nurse could select the manual review flag so that the nurse manager could review the case before sending it for billing. The percentage of cases for which the flag was not set was calculated at each hospital for the 30 days prior to this study.

Charging issues that have surfaced since the beginning of the pilot have been documented. Many of these issues have required a consensus among the hospitals. The back and forth deliberations among the nurses has generated an e-mail trail. We used the number of discrete e-mail communications as a surrogate for issue complexity.

RESULTS AND DISCUSSION

Table 1 summarizes the chart reviews for the labor and delivery cases requiring high-intensity nursing care that were conducted for each of the 7 pilot hospitals. The most dramatic finding is the single missed charge found in the 70 charts we reviewed. This is substantially lower than the pre-pilot finding in which 35% of the cases reviewed had at least one missing charge. The single missed charge was for a piece of equipment that accounted for less than 1% of the total labor and delivery charge.

On average, one in 5 of the reviewed cases had a charge-related free-text entry. Except for the one charge-related free-text entry that resulted in the missed charge, all other free-text entries were redundant with coded charting entries that triggered the appropriate charge. The fact that these redundant free-text entries do not add value or clarity to the chart is an educational issue for the labor and delivery staff. The high standard deviation value for total

Table 1. Summary of the chart review giving the missed charges and the perceived confidence measure for each of the 7 hospitals pilot testing the labor and delivery charge capture software.

Hospital	Total charted items per case		Total free-text entries per case		Total charge-related free-text		Missed charges	Perceived confidence
	Avg	*St.dev*	*Avg*	*St.dev*	*Avg*	*St.dev*	*Count*	*%*
1	396	100	2.8	3.4	0.10	0.32	0	85
2	671	101	3.2	2.1	0.30	0.48	1	94
3	720	195	4.1	2.1	0.20	0.42	0	96
4	665	170	2.3	1.9	0.30	0.48	0	98
5	693	249	5.1	5.2	0.30	0.48	0	98
6	503	218	5.0	6.2	0.10	0.32	0	93
7	431	108	1.9	3.6	0.00	0.00	0	96
All	**583**	**208**	**3.5**	**3.9**	**0.19**	**0.39**	**1**	**95**

free-text entries substantiates the widely held view that some clinicians prefer to use free-text. Again, this becomes an educational issue for the labor and delivery staff. Nonetheless we were pleased with finding an average of only 3.5 free-text entries per case.

The average measure of perceived confidence among all hospitals was 95% with a range from 85 to 98%. The one case having a missing charge among the 70 charts reviewed cases gives an overall indicator of accuracy of 98.5%, which agrees with the high end of the perceived confidence range. Reasons for low perceived confidence (85%) may include clinician caution during the pilot and lack of feedback regarding the accuracy of the charging algorithms. Perceived confidence should rise over time as accuracy feedback is received and as more and more of the issues flagged for review are resolved.

Figure 1 shows a total of 103 charging algorithm issues that surfaced from the beginning of the pilot resulting in 927 discrete e-mail communications among the nursing and administrative staff trying to agree on the rules. The majority of these charging issues (53%) were resolved with a single response. An example would be when a nurse needed to know if a particular charted element should trigger a charge. A smaller cluster of charging issues (22%) was resolved with 8 or less communications, where a quick consensus was found among the participating hospitals. An example would be arriving at a consensus of how to chart duplicate supplies. The remaining cluster (26%), resolved with more than 8 communications, required reconciliation of differing charging practices by hospital. These cases dealt with defining acuity triggers within the documentation. The issue generating the largest volume to date was patient blood loss, which accounts for 19% of the total communication volume. Although the charging rules about patient bleeding have been articulated in the Charge Master for several years, each hospital had treated the issue independently. The large volume of communication ensued from our effort to determine how the quantity and rate of bleeding corresponds to equivalent care.

Figure 2 shows the charging issues and their volume by date they first became an issue for discussion. The two clusters correspond to the software implementation for the two groups of pilot hospitals. The cluster for the first three pilot hospitals is much lower than the cluster for the second set of four hospitals. This result suggests that the complexity of issue resolution is proportional to the number of hospitals having to build a consensus. The issues raised during the second wave of pilot hospitals not only had to be acceptable among themselves but also with the first set of hospitals. This observation may also reflect the bias of the pre-pilot work, where the clinicians from the first set of pilots had a more immediate interest in the rule building process. The band of low-volume issues, those resolved in three or less communications, follows the expected trend where the frequency of issues decreases after the initial onset of each pilot.

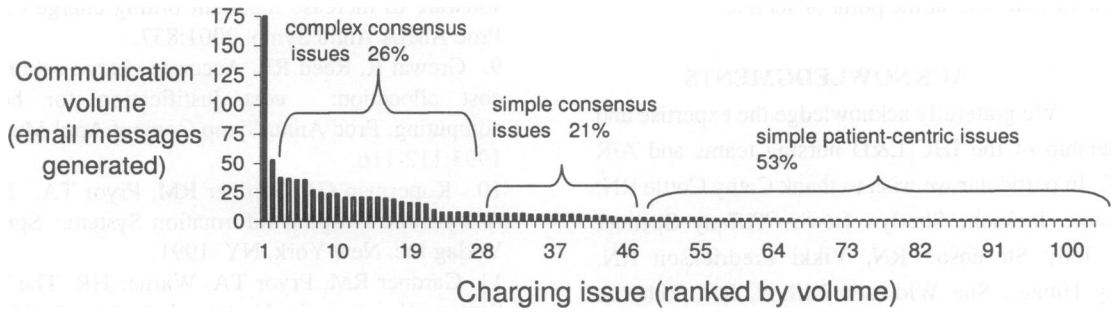

Figure 1. Charging issues ranked by volume of email messages sent among nursing management and administration while resolving the issue.

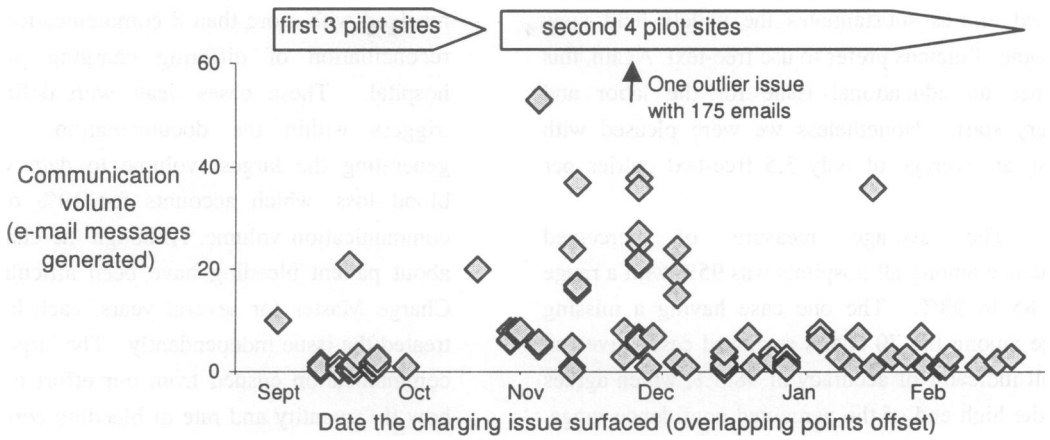

Figure 2. Charging issues by date and communication volume

CONCLUSIONS

The IHC pilot in labor and delivery for deriving charges from the patient care documentation collected at the point of service showed a substantial improvement in reducing the number of missed charges while generating uniform charge summaries for equivalent care. The software implementation reduces the clinician's burden of data interpretation for the purpose of charge capture thereby freeing the clinician for more patient-centric work. This pilot implementation demonstrates the ability to adequately generate activity-based charges from patient care documentation. However, the effort in defining the specific fields and the logic that will link them to the charging rules increases proportionately with the number of facilities trying to reach a consensus. The effort required to reach these agreements is compensated by the remarkable improvement in the quality of the charge capture process that stems from having both clinical and financial needs served by data collected once at the point of service.

ACKNOWLEDGMENTS

We gratefully acknowledge the expertise and leadership of the IHC L&D nursing teams and A/R staff. In particular we wish to thank Cathy Cottle RN, DeLaine Jackson, Sherlyn Lewis, Tiffany Spencer RN, Jody Stevenson RN, Vikki Fredrickson RN, Holly Hunter, Sue Widmann MSPH, RN, Rebecca Allen, RN, Kris Hansen RN, and Diane Jensen. We also acknowledge computer and programming assistance from Robert Reese and Andy Barraco.

REFERENCES

1. Twede M, Gardner RM, Hebertson RM. A PC-based system for intrapartum monitoring. Contemporary OB/GYN, "Special Issue -- Technology 1985" Oct 1984; 24:13-17.
2. Reason J. Human Error. Cambridge University Press. New York, NY 1990.
3. Norman DA. The design of everyday things. Doubleday/Currency. New York, NY 1988.
4. Leape LL, Kabcenell A, Berwick DM, et al. Reducing adverse drug events. Institute for Healthcare Improvement. Boston, MA 1988.
5. Institute of Medicine. To err is human: building a safer health system. Kohn LT, Corrigan JM and Donaldson MS, editors. National Academy Press, Washington D.C. 2000.
6. Baker JJ. Activity-based costing and activity-based management for health care. Aspen Publisher, Inc., Gaithersburg, MD. 1998.
7. Schmitt KF, Wofford DA. Financial Analysis projects clear returns from electronic medical records. Healthcare Financial Management 2002 Jan:52-57.
8. Strayer SM, Study evaluating handheld computer software to increase inpatient billing charge capture. Proc AMIA Annu Symp. 2001:837.
9. Grewal R, Reed RL. Accurate charge capture and cost allocation: cost justification for bedside computing, Proc Annu Symp Comput Appl Med Care 1993:112-116.
10. Kuperman GJ, Gardner RM, Pryor TA. HELP: A Dynamic Hospital Information System. Springer-Verlag Inc, New York, NY 1991.
11. Gardner RM, Pryor TA, Warner HR. The HELP System: Functionality Update. Intl J Med Informatics 1999 Jun;54:(3):169-182.