Professional Case Management
Vol. 14, No. 4, 185–191

Implementing an Automated Acuity Tool for Scoring Case Management Cases and Caseloads at Blue Cross Blue Shield of Massachusetts

Kathy Craig, MS, RN, CCM, and Anne Flaherty-Quemere, RN

ABSTRACT

Purpose/Objectives: Blue Cross Blue Shield of Massachusetts (BCBSMA) required a statistically validated tool to understand the acuity of its case management (CM) populations. The BCBSMA management sought reliable methods to quantify the weights of case managers' caseloads, allow associates and managers to have conversations about workloads based on objective measurable data, and analyze the data to ensure that the correct populations are being targeted for CM.

Primary Practice Setting(s): The tools, techniques, and strategies described in the Acuity and eQuity Workflow Solutions are suitable for all healthcare practice settings. Acuity scoring metrics are designed especially for CM settings such as independent stand-alone companies, inpatient and outpatient practices, and multistrata health plans and hospital consortia.

Findings/Conclusions: The customized BluCuity assessment tool, tested in the interrater reliability (IRR) study arm of the project, was validated to a high degree of concordance in BluCuity scoring among CM raters with an overall agreement of 82% (20 novel raters, 20 cases, 190 rater pairs, 3,800 pairwise case reviews; $p = .03$). The BluCuity tool and workflow strategies were implemented into the BCBSMA information technology (IT) system for oncology and other CM teams in 2008. Report data demonstrated the acuities of case managers' cases, the relative case and caseload weights among a team of case managers, and the differences in cases and caseloads across CM product offerings.

Implications for Case Management Practice: Acuity-based tools such as BluCuity to access the severity, intensity, and complexity of CM cases operationalize quality assessments into quantifiable data. Acuity score reports can join other accounting and operational reports to determine appropriate caseloads, case assignments, and staffing. To assess consistency of judgment among case managers, CM organizations can implement a formal IRR testing methodology to evaluate the level of clinical assessment reliability and establish training protocols. The automated IT acuity assessment system and data outputs can be used to assess the overall weight of a caseload based on diagnosis code groupings such as oncology, pediatric, transplants, and stroke, and illustrate the intensity of services required by CM clients at different times or in different programs. Organizations that sell CM services can use the ability to score the acuity of CM cases to scale and justify the pricing of their services.

Key words: *acuity, caseload, interrater reliability*

Not only was Tuesday the day case managers in the oncology team at Blue Cross Blue Shield of Massachusetts (BCBSMA) received new cases, the first Tuesday of July 2008 was the scheduled rollout day for the new BluCuity tool. Buzz about BluCuity had been building for 2 months. BluCuity constituted the cornerstone of the acuity-based information technology (IT) system, a project designed to score case management (CM) cases for complexity using a customized Acuity Tool that was statistically validated for reliability. The achievement

Address correspondence to Kathy Craig, MS, RN, CCM, Craig Research Continuum, LLC, 313 Victoria St, Kingston, Ontario, Canada K7L 3Z2 (kdcraigrn@earthlink.net).

This integration project was undertaken by Blue Cross Blue Shield of Massachusetts (BCBSMA) and piloted by the case managers of the Oncology Team, BCBSMA, Quincy, MA. The acuity-customization project was commissioned by Anne Flaherty-Quemere, the BCBSMA Case Management Senior Business Lead. The IT project was accomplished by Janelle Miller, BCBSMA Systems and Reports expert and part of Business Integration. Their enthusiastic support for this project is appreciatively acknowledged.

…although the active caseload logs were known to range from 40 to 60 cases, case managers were discovered to be carrying loads of almost equal sizes in the amounts of cases in their case intake queues (from 30 to 50 cases). When these true caseloads were combined, the case numbers swelled from 70 to 110 cases; these were their real caseload sizes.

heralded a milestone that frontline and business personnel alike could mark as a breakthrough toward the twin goals of better data-driven identification of cases and caseloads.

Case management companies require a validated instrument and process methodology to indicate, with demonstrated reliability, the overall complexity of the cases their case managers handle. A statistically validated tool is needed to help understand the acuity of the CM population, understand the weight of the case manager's caseload, allow associates and managers to have conversations about workloads based on objective measurable data, and analyze the data to ensure that the correct population is being targeted for CM. To approach these goals, the acuity-based scoring system called Acuity Workflow Solutions (and its electronic counterpart eQuity Workflow Solutions) was customized and incorporated into the internal IT system at BCBSMA in the first quarter of 2008.

Background

The operational premise of the acuity workflow system is that different cases have different acuity weights. Acuity weights are metrics derived from three critical domains of CM assessment: (1) client need-severity; (2) primary and backup carer (caregiver) need-severity; and (3) CM intervention-intensity (Craig & Huber, 2007; Huber & Craig, 2007a, 2007b). Scores are obtained from the selections case managers make from factors called drivers and subdrivers that worsen or improve by degrees within the tool's three domains. Case managers use the scoring methodology to convert their quality assessments into quantifiable data. The data-coded measurements categorize and communicate both the complex problems patients and their caregivers are encountering and the complex interventions case managers need to institute and coordinate to address the need severities. In the paper-based acuity architecture, the four driver/subdriver combinations are seen to occur in cells or boxes within the three domains visible as columns. From each domain, case managers select the most representative driver/subdriver combination from the cells of corresponding severity, intensity, and complexity components identified as Degree 1 through Degree 5. [A thorough description of driver/subdriver combinations is available in the studies by Craig & Huber, 2007.]

Project Expectations

The BCBSMA project supervisor identified eight expected outcomes of the acuity workflow project (Table 1). Immediate objectives of the assessment process were to identify the acuity of each case as a means of documentation and prioritization and to track caseloads based on a defined principle of case acuity that has been tested for concordance (agreement) among the case managers who use the tool. A longer term purpose was to build a reporting architecture to enable the capture of outcomes in individual cases and in aggregated cases related to the work that case managers enact on behalf of their clients (members or patients). Outcomes measures based on CM case acuities would be joined with other business measures to confirm case managers' positive contributions and to signify areas of improvement for ensuring the positive business impact and value-added effect of CM.

The eQuity Workflow Solutions project included three phases that aimed to evaluate whether the assessment tool used by case managers at BCBSMA was a well-designed and validated IT product that met the company's goals. The project's three main phases were (1) development, (2) testing, and (3) implementation. The project team drew from instrumental personnel in the CM oncology team and the systems and reports department. The five-membered team was managed internally by the CM business lead. The external consultant directed the overall project's activities and carried responsibilities for each phase's deliverables.

Development Phase

Development included a case manager survey, a *Reflective Practice Analysis*, and the Acuity Tool customization. In March 2008, case managers involved in the initial implementation were surveyed to establish a set of preimplementation responses against which to compare responses at postimplementation. Survey questions asked about the status

TABLE 1
BluCuity Tool's Eight Expected Achievements

1. *Train all case managers involved in the initial implementation on how to use the BluCuity tool.*
   - Before end of the fourth quarter of 2008
2. *Formalize the development of weekly, automated BluCuity reports for case management (CM) individuals and teams.*
   - In the first quarter of 2009
3. *Implement data quality reporting.*
   - To identify coaching and teaching opportunities for case managers such as
     - Case managers who require instruction to improve the use of the BluCuity tool
     - Case managers who have delays in assessing BluCuity in their cases
   - To establish policies for process and performance metrics based on acuity workflows and BluCuity scores
4. *Discover weights of current caseloads.*
   - For individual case managers
   - Within each CM team
   - Between care coordination teams
5. *Incorporate BluCuity-based scores with other reports and metrics.*
   - To evaluate and refine current metric standards for caseloads per staff
   - To improve case handling and satisfaction
6. *Combine insights from BluCuity data derived from individual case manager's cases with content from operational sources.*
   - To evaluate weights of case manager's caseloads for improving efficiency
   - To determine case manager's differing capacities for assignment of new cases
7. *Assess weights of cases according to aggregated data.*
   - By diagnosis codes and groupings
   - By CM populations, such as oncology, pediatrics, behavioral health, and adult high risk
   - By value-laden designations, such as geographic or provider distinctions
8. *Formalize the interrater reliability (IRR) process at Blue Cross Blue Shield of Massachusetts.*
   - To ensure strong IRR agreement in similar test cases and thereby demonstrate that case managers use the BluCuity tool to assess and score complexity with high degrees of accuracy, consistency, and dependability
   - To reveal documentation and case profiles that comprise problematic cases in which case managers generate weak IRR agreement for quality improvement study or rectification

Note. Eight expected achievements related to the BluCuity implementation project describe the primary goals that the Blue Cross Blue Shield of Massachusetts management team identified as reasons to undertake the customization of the Acuity Tool and to incorporate it into the IT system used by the oncology team's case managers.

of, and satisfaction with, current acuity assessment strategies. Replies were yes/no responses and answered on a 7-point Likert-type scale. Two of the 10 case managers scored current assessment capabilities as "adequate." Eight classified them as "less than adequate," "poor," "very poor," or "not present." In a free-form comment section, case managers expressed the need for a quick, accurate tool that added objectivity to case dispersal for improved efficiency and caseload equitability.

Reflective practice analysis was carried out by the consultant to discover the unique terminology used by case managers at BCBSMA and to uncover discrepancies between management's process expectations and case managers' actual process steps and work-arounds. By sitting with staff, the consultant observed and queried case managers about their work steps. The analysis succeeded in identifying specific terminology that was channeled into customizing the iteratively unique BluCuity tool. Case managers were encouraged to provide feedback about the BluCuity drafts to bolster the investment in ownership. Feedback furthered the tool's ability to reflect the clinical decision-making process as an organic part of their actual practices rather than an artificial and awkward overlay onto their practice steps.

The practice analysis uncovered one specific process bottleneck in the form of large sizes of cases in the case managers' intake queues in relation to their active case logs. For example, although the active caseload logs were known to range from 40 to 60 cases, case managers were discovered to be carrying loads of almost equal sizes in the amounts of cases in their case intake queues (from 30 to 50 cases). When these true caseloads were combined, the case numbers swelled from 70 to 110; these were their real caseload sizes. Information was garnered about how these queues affected caseloads and burdened case managers in ways that management had undervalued or not appreciated fully before. Accounting for the true numbers of throughput cases aided the design of workflow strategies to acknowledge the backlogs better and to develop measurements that appreciated the caseloads more accurately.

During the development phase, members of the project team considered how to incorporate BluCuity into the case managers' clinical information system. Initially, the tool was planned for use in a hard-copy paper format. Entry of the acuity scores and driver/subdriver selections in the IT system required manual efforts by each case manager. However, the BCBSMA care management IT system and data-extraction expert served as the team's internal IT system lead and spearheaded the introduction of the automated BluCuity methodology into

Caseload weights are derived by adding the weighted subtotals of case numbers multiplied by the acuity score. Here are two examples.

Case manager Sid has 65 cases that weigh 135 craigs.[a] His caseload consists of 15 cases with acuity 1, 31 cases with acuity 2, 18 cases with acuity 3, 1 case with acuity 4, and no cases with acuity 5. A caseload distributed weight is obtained by adding 15 (15 cases x 1 acuity) plus 62 (31 cases x 2 acuity) plus 54 (18 cases x 3 acuity) plus 4 (1 case x 4 acuity), which total 135 craigs.

15 cases x acuity 1 = 15; 31 cases x acuity 2 = 62; 18 cases x acuity 3 = 54; 1 case x acuity 4 = 4; 0 cases x acuity 5 = 0
15 + 62 + 54+ 1 + 0 = 135 craigs

Case manager Kathleen's 65 cases weigh 225 craigs. She has 1 case with acuity 1, 5 cases with acuity 2, 28 cases with acuity 3, 25 cases with acuity 4, and 6 cases with acuity 5. Although the case numbers for Sid and Kathleen are both 65, the distributed weights of 135 and 225 represent very different caseloads.

1 x 1 = 1; 5 x 2= 10; 28 x 3 = 84; 25 x 4 = 100; 6 x 5 = 30
1 + 10 + 84 + 100 + 30 = 225 craigs

[a]A craig is the unit of caseload weight derived from the acuity scoring metrics described in the Acuity and eQuity Workflow Solutions™ methodologies.

FIGURE 1
Description of how to obtain caseload weights.

the care management system. The IT design used a behind-the-scenes solution to help address the intake queue dilemma that added a pseudoweight for cases that had not been scored for acuity. The pseudoacuities could be separated in the reports from cases with scores that represented true acuity assessments. The true case acuity and pseudoacuity scores combined to quantify the quality-based clinical assessments as case weights that underpinned the caseload weights. See Figure 1 for description of how to obtain caseload weights.

Also, during the development phase, policies to incorporate BluCuity assessment into the CM workflows were started. Performance expectations were drafted, such as identifying a turnaround time of five business days for case managers to complete acuity assessments and produce acuity scores after engaging members (patients) in the CM program.

BluCuity Testing Phase

The BluCuity IT workflow was intended to construct a reliable reporting foundation upon which to build a robust data-informed, two-way communications conduit for case acuity, caseload acuities, and acuity-based decision making. A crucial lynchpin of the communications and reporting architecture was the confirmation that case managers could use the BluCuity tool to produce data that have a high degree of reliability. If it could not be demonstrated by an overwhelming margin that case managers would assess and assign similar BluCuity scores in similar cases, then the entire workflow process was only an exercise in wishful thinking.

To re-create the process executed to validate the original Acuity Tool, an interrater reliability (IRR) study arm of the project was undertaken during the second quarter of 2008. The purpose of the IRR study was to demonstrate to what extent the BluCuity assessments could be shown as a reliable and replicable source of data. The goal of the IRR study was to test and quantify the capacity of the BluCuity tool to be used by case managers to produce acuity scores that were the same-or-similar in same-or-similar cases. This ability was called *agreement* or *concordance*. The IRR study used a version of the BluCuity tool thoroughly vetted by the project team and case managers. Two members of the project team scrutinized 50 oncology cases for completeness and acceptability. Through a repeatable randomization technique, 20 cases were designated for use in the IRR test. Case managers involved in the upcoming implementation participated in the IRR study. Raters were *blinded* to (intentionally kept unaware of) knowing they were reviewing the same 20 test cases. In addition, they did not know which case fellow raters were reviewing at any particular time. The IRR, conducted over 2 half-day intervals at BCBSMA, was proctored to preserve the IRR blinding integrity.

By using the BluCuity criteria to assess clients' severity and CM intervention intensity, supervisors were empowered to demonstrate that some case managers were undervaluing their work by undervaluing a case's complexity. This data-driven demonstration led to a plan for teaching, training, and communicating about accuracy in assessing the complexity and acuity in oncology cases.

The IRR study demonstrated that concordance among case managers reached an overall agreement of 82% (from 20 raters; 190 rater pairs, 3,800 pairwise case reviews; $p = .03$; Craig, 2008). The 82% agreement in CM assessments represented an overwhelming margin that far exceeded average, and it closely approached the 85% benchmark target for agreement in healthcare disciplines such as microscopy. When viewing pathology specimens, achieving diagnostic concordance in at least 85% of tested matches is critical. Although CM clinical judgments carry no diagnostic criticality, it nevertheless remains important to exceed 80% and to approach the 85% benchmark whenever possible. The acuity scores in the test cases done by the raters[1] were analyzed using computational cross comparisons in more than 6,500 pairwise simulations. The 82% agreement among novel users in more than 6,500 pairwise matches showed statistical rigor and verified that the BluCuity tool's customized components were sound foundations to access acuity in oncology cases. Most important, the IRR methodology reasserted the BluCuity tool's validity. Thereafter, frontline personnel, CM supervisors, and front-office business managers at BCBSMA could know that highly reliable data stemmed from the BluCuity assessments. Because preliminary statistical findings had been reported previously to BCBSMA supervisors in June 2008, the IT acuity project had proceeded toward implementation.

Implementation Phase

With the BluCuity tool statistically validated in June 2008, implementation to the oncology team successfully occurred on July 1, 2008. All new cases were assessed for acuity using the BluCuity tool when the cases were opened. Cases that had been active in oncology before the BluCuity implementation rollout date were assessed for acuity at the next phone contact with the patients (members). Although the initial rollout used a paper-based version, the automated BluCuity assessment tool was launched in August 2008 in the care management system. In the fourth quarter of 2008, case managers provided feedback to the acuity project consultant. After discussing the associates' feedback with the project team, BluCuity was refined minimally to avoid disrupting the tool's validity. Enhancements included improved capture of end-of-life palliative and hospice care as well as increased visibility of care transitions. BluCuity, Version 2.4, was synchronized with the electronic IT BluCuity and released in the late fourth quarter of 2008. In addition, a hard copy "job aide" was developed to help associates better understand the application of the driver/subdriver selections. The job aide provided key term definitions, intended use descriptions, and narrative documentation samples for each domain in large type font for easier reference on a daily basis and for subsequent training and coaching opportunities.

Challenges and Achievements

Several challenges were encountered. During feedback interviews, case managers new to the CM field reported that they found it difficult to understand some of the condensed, one-line acuity descriptions, which had to be brief or abbreviated because of computer field limitations. Efforts to improve the understanding of the acuity components were achieved through the release of a revised BluCuity version and the production of a "job aide," a noncomputerized helper sheet with definitions of abbreviations and examples of acuity driver/subdriver combination selections.

Two-way channels of communication were established and encouraged to increase the feelings of ownership by frontline personnel. Case managers engaged in comanaging cases with behavioral health and pediatric CM found gaps in the BluCuity tool customized for oncology. Although these challenges persisted, it was recommended that discussions should begin regarding customization and validation of BluCuity for these CM practice specialties.

Because initial reports contained "dirty data," cleanup efforts were begun during the postimplementation phase. Continuous quality improvement (CQI) processes were not in place during the implementation phase in the third and fourth quarters of 2008. However, the CQI and data examination

[1]These raters are BCBSMA case managers who had not used BluCuity in practice settings before the IRR testing.

processes were started in the first quarter of 2009. A formal work group was established, which included senior managers, managers, systems and reporting experts, and the external consultant. The CQI and reporting work group began to meet biweekly in January 2009 with an agenda to further enhance the BluCuity tool and its use, automate the reporting architecture, and scrutinize the reports for data content utility and shortfalls. Although a delay in folding the BluCuity data into the other operational reports occurred, it was decided that data arising from BluCuity scores in cases and the averages of caseload acuities eventually will be incorporated with the operational reports to obtain a more holistic picture. In addition to agreeing to commit resources to develop a high-level report that blends the BluCuity and other operational data, decision making about the frequency of how often to produce the individual reports and how to use the blended report for consistency was begun.

A large part of the success story was that data from the case managers about case acuities were flowing and could be tabulated, considered, and analyzed. Previously, case reviews had been able to identify cases into subjective low, medium, and high groups without the ability to run reports categorizing the caseloads. After the BluCuity workflow implementation, the Data Trifecta tap for process, performance, and outcomes data had been turned on. As an example, Table 2 shows a sample of the IRR test output data for 20 raters' BluCuity scores in five test cases with diagnosis codes and scoring totals.

By using the BluCuity criteria to assess clients' severity and CM intervention intensity, supervisors were empowered to demonstrate that some case managers were undervaluing their work by undervaluing a case's complexity. This data-driven demonstration led to a plan for teaching, training, and communicating about accuracy in assessing the complexity and acuity in oncology cases by using the BluCuity tool as the fulcrum of measurement. Also, establishing an IRR process in which periodic IRR tests are conducted among groups of case managers was seen as a teaching aid, measurement tool, and CQI method

TABLE 2
Case Acuity Sample

	Acuity																								
	Case 1 (*ICD-9* 710)					Case 2 (*ICD-9* 998.5)					Case 3 (*ICD-9* 890.1)					Case 4 (*ICD-9* 518.8)					Case 5 (*ICD-9* 518.81)				
Rater	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
R1					X					X					X					X					X
R2					X				X						X					X			X		
R3					X		X								X					X					X
R4					X			X						X					X						X
R5					X				X						X				X						X
R6					X					X					X					X					X
R7					X				X						X				X						X
R8					X				X						X					X					X
R9					X					X					X					X					X
R10					X					X					X				X					X	
R11					X					X					X					X			X		
R12					X					X					X					X				X	
R13					X					X					X					X				X	
R14					X			X							X					X					X
R15					X					X					X					X					X
R16					X					X				X					X						X
R17					X				X						X					X					X
R18					X					X					X					X					X
R19					X					X					X					X				X	
R20					X					X					X					X					X
Total	0	0	0	0	20	0	1	2	5	12	0	0	0	2	18	0	0	0	5	15	0	0	2	4	14

Note. Case acuity sample provides insight into data that can be obtained by using the interrater reliability testing methodology within the Acuity and eQuity Workflow Solutions.

important for the development of consistency and accuracy in assessments on an ongoing basis not only within the oncology team but also across other teams.

Summary

The Acuity Tool was customized into the BluCuity assessment tool. It was tested in the IRR study arm and validated to a high degree of concordance in BluCuity scoring among CM raters. The hard-copy BluCuity tool was launched successfully into the BCBSMA IT system. The implementation of BluCuity was rolled out to the oncology team in the second quarter of 2008 and to the rest of CM in the fourth quarter of 2008. Case managers and management staff have used BluCuity to help frame how they think about and systematically quantify members' issues. Both frontline and supervisory personnel have reported that the BluCuity workflow process is influencing and improving documentation in the clinical IT system. Data reports have begun to appear that show distributions of acuities. In addition, case and team managers have begun to turn to these reports with eagerness to discover the acuities of cases the case managers carry, the weights of the caseloads they shoulder, the pitfalls the BluCuity workflow process encounters, as well as the struggles the case managers encounter, and the successes case managers are achieving.

Relevance to Broader CM Field

Case management organizations can use tools such as BluCuity to access the complexity of CM cases. The Acuity and eQuity Workflow Solutions can be used to demonstrate the weights of cases that one case manager is carrying, the relative case and caseload weights among a team of case managers, or the differences in cases and caseloads for various diagnosis groups or CM product offerings. Acuity can be incorporated into operational reports to determine appropriate caseloads, case assignments, and staffing. Furthermore, to ensure consistency of judgment among case managers, CM organizations can implement a formal IRR process to establish and promote a high level of clinical assessment reliability. The automated IT acuity assessment system and data can be used to assess the overall weight of a caseload based on diagnosis code groupings such as oncology, pediatric, transplants, and stroke. The acuity-derived data can be incorporated into accounting reports to illustrate the intensity of services required by CM clients at different times or in different programs. In a broader application of the ability to score the acuity of CM cases, organizations that sell CM services can use acuity to scale and justify the pricing of their services.

References

Craig, K. D. (2008, August 6). *Inter-rater reliability studies for the initial draft version of BluCuity: An Acuity Workflow Solutions™ Project, Part B—Inferential statistical analyses* (p. 2). Internal report to BCBSMA from Craig Research Continuum, LLC. Kingston, Ontario, Canada: Craig Research Continuum, LLC.

Craig, K., & Huber, D. (2007). Acuity and case management: A healthy dose of outcomes, Part II. *Professional Case Management, 12*(4), 199–210.

Huber, D., & Craig, K. (2007a). Acuity and case management: A healthy dose of outcomes, Part I. *Professional Case Management, 12*(3), 132–144.

Huber, D., & Craig, K. (2007b). Acuity and case management: A healthy dose of outcomes, Part III. *Professional Case Management, 12*(5), 254–269.

Kathy Craig, MS, RN, CCM, is the President and Founder of Craig Research Continuum, LLC. She is an applied researcher, analyst, and tool inventor with a 30-year practice heritage in nursing and in domestic and international case management.

Anne Flaherty-Quemere, RN, serves as is the Senior Manager of BCBSMA Case Management Business Unit Leader. She focuses on developing quality-based data metrics to enhance patient service delivery. In her 25-year healthcare career, more than 12 years are in leadership and supervisory roles.