



This PDF version of the Criterion has a somewhat different appearance than the email version. Please locate articles by clicking on the table of contents or by scrolling up and down to the page number and article.

**Contents**

About this Issue ..... 1  
 Linking Central Cancer Registries and Institutional Biorepositories: Pilot Study ..... 2  
 The Importance of Path Numbers ..... 3  
 Hyphenated Surnames ..... 4  
 Business Rules Management Solutions - A CCR Project..... 4  
 Educational Opportunities ..... 6  
 In Memmori- am, Lori Lenberg, CTR..... 6

**About this Issue**

Katheryne Vance, BA, CTR

The cancer registry annual spring educational meeting events, National Cancer Registrars Association (NCRA) and North American Association of Central Cancer Registries (NAACCR), where we have an opportunity to enrich our knowledge, network with colleagues, see old friends, and make new friends have passed. It would be great to say the lazy crazy days of summer are here! If you are like me, summer is the time of year to catch up on the stack of files and papers on my desk in addition to my routine workload. There are just a few weeks of respite before gearing up again come fall.

In this issue, there are two articles about activities at the CCR. The first article by Rosemary Cress, Dr. P.H, and Margaret McCusker, M.D., M.S., is titled Linking Central Cancer Registries and Institutional Biorepositories to Facilitate Biospecimen-based Research - A Pilot Study. This received a second place award as a poster presentation in the category of Data Use at the recent NAACCR Annual Conference in Louisville, Kentucky. The second article on Business Rules Management Solutions - A CCR Project by Cheryl Moody, BA,CTR, describes an important CCR project that assures quality data by implementing a business model.

The regular column Chirps! is on vacation. In its place are two articles that will assist registrars in abstracting and will aid the CCR and the regional registries in case linkage. Mari Carlos, CTR, describes the importance of entering the exact and complete pathology report accession numbers in The Importance of Path Numbers. Did you ever think that a hyphen in a patient's surname could affect how the central data system, Eureka, links case reports? Crickett Dyke, CTR, describes the attributes of this symbol in Hyphenated Surnames and how to report a hyphenated surname.

The NAACCR Board of Directors has certified that the California Cancer Registry has attained its' Gold Standard for quality, completeness and timeliness again this year. This would not be possible without your continued dedication to submitting timely, high quality data! Your efforts contribute to NAACCR's mission of "Making Every Cancer Count!"

## Linking Central Cancer Registries and Institutional Biorepositories: Pilot Study

Rosemary Cress, Dr. P.H., Co-Principle Investigator Greater California SEER Registry  
Margaret McCusker, M.D. M.S., Medical Officer, California Department of Public Health

Biospecimens (blood and tissue specimens) are valuable tools for cancer research. Analysis of biospecimens is used to evaluate the molecular profiles of tumors; describe the molecular epidemiology of newly-identified oncogenes and their impact on recurrence and survival; and study the molecular epidemiology of rare tumors and tumors among specific population subgroups. A powerful research tool is created when existing biospecimen data are linked to cancer registries that provide high-quality, population-based data about cancer patients.

Last year, staff from the California Cancer Registry (CCR) conducted a pilot study to determine if biospecimen data from the UC Davis Cancer Center Biorepository (UCD CCB) could reliably be linked to records in the CCR database. Using standard CCR probabilistic data linkage procedures, UCD CCB records from 2005-2009 were linked to CCR cases diagnosed through 2009. Because the variable that is most often used for linkage (social security number) is not available in the UCD CCB records, the linkage was done using name, sex, date of birth, race/ethnicity, and pathology information. Because most patients have more than one biospecimen, only records that had a unique value for medical record number, tissue site and pathology specimen date were used.

Of 1,040 UCD CCB records with a unique medical record number, tissue site, and pathology data, 844 (81.2%) linked to CCR records. The Table 4 below shows the number of records in both databases by cancer site. Matches were highest for cancers of the cervix and testis/other male genital system, and lowest for cancers of the skin and bones/joints.

This pilot linkage demonstrated that existing biorepository data can be linked to cancer registry data to identify biospecimens for population-based research. Improvements to the variables critical for a successful linkage (first and last name, date of birth, facility medical record number, cancer site, and pathology report number) will improve the success of future linkages.

**Table 4: Records Matching by Cancer Site in the UCD CCB and CCR Databases, 2005-2009  
(n=1,040)**

Description	# Cases Used	# Matches	% Matches
Cervix	5	5	100.0%
Testis/other Male Genital System	7	7	100.0%
Corpus and Uterus, NOS	37	36	97.3%
Respiratory System	114	106	93.0%

Breast	48	44	91.7%
Kidney	108	97	89.8%
Bladder	49	44	89.8%
Colorectal	38	34	89.5%
Other Urinary System	9	8	88.9%
Lymphoma	18	16	88.9%
Stomach	8	7	87.5%
Pancreas	40	34	85.0%
Endocrine System	12	10	83.3%
Brain and Other Nervous System	23	19	82.6%
Other Digestive System	10	8	80.0%
Liver	10	8	80.0%
Ovary	33	25	75.8%
Prostate	376	274	72.9%
Oral Cavity and Pharynx	18	13	72.2%
Soft Tissue Including Heart	48	34	70.8%
Miscellaneous	18	12	66.7%
Bones and Joints	6	2	33.3%
Skin	5	1	20.0%
Prepared by the California Cancer Registry, California Department of Public Health, Cancer Surveillance Section.			

## The Importance of Path Numbers

Mari Carlos, CTR  
Cancer Surveillance Program - Los Angeles

Historically registrars came up with a number of innovative ways to make path report numbers fit into the too small space allotted by most cancer reporting software. A favorite way was to simply eliminate any leading zeroes. For instance path report number S07-00291 might have been entered in the path number field as S07-291. Less frequently this number might be abbreviated as 00291, or S0700291. While in a purely manual system these would usually result in a correct match, the anticipated automation of path/abstract matching will be in jeopardy if the entire path report number is not recorded accurately.

One of the data set changes of 2008 was the expansion of DxRx Report Number fields from two to five [now called Path Report Numbers (1-5)], and the addition of 12 more character spaces to each of these fields. With 20 characters per field, it should never be necessary to abbreviate a path report number.

The stated reason for the expansion from 8 characters to 20 characters was documented in the Changes and Clarifications to Volume I. It stated: "In order for the CCR's central data base system (Eureka) to integrate pathology report processing with new case abstract processing, the system needs a way to easily match abstracts to path reports." This means that, among other things, path report numbers should be recorded exactly as they are written, including all zeroes, hyphens and spaces. If the reporting facility has generated pathology reports, the Path Report Numbers field(s) must include the complete number for every report.

## Hyphenated Surnames

Crickett Dyke, CTR  
Desert Sierra Cancer Surveillance Program

Apparently, there are some limitations in the software for hospital facesheets that do not allow the hyphen character ('-') in the last name field and, therefore, a hyphenated surname is actually reflected as a single surname, all squished together. So...*Teresa Jones-Packard* is recorded as *Teresa Jonespackard* on the hospital facesheet. The cancer registrar records the name as a non-hyphenated surname, *Jonespackard*, because we loyally record what's on the facesheet. This has had a troublesome direct effect at the regional registry level when linking new case reports to the same patients already in the database. Eureka searches and links differently for *Jonespackard* than it does for *Jones-Packard*. Eureka takes a hyphenated surname and automatically fills each half of the name into the Alias and/or Maiden name fields and then searches on each half of the surname. When it's a 'run-on' name, Eureka will many times miss linking the appropriate patient, possibly causing incorrect patient counts.

We would like to encourage registrars to double check the patient admission signature form when they suspect a surname should be a hyphenated surname. The patient will usually sign their name as two separated names or as a hyphenated name and that can be verified. If, in fact, it should be a hyphenated surname, please record the surname in your cancer reporting software *with* the hyphen included, even though it is not visible on your hospital facesheet. This would help us to ensure better patient and tumor linkages at the regional registry level and keep our cancer counts correct.

We would also like to encourage registrars to make use of the Alias name fields when there is a variation in the patient's name. For example, one admission may have the patient's first name as *TERESA*, and an outpatient biopsy report may have her as *TERRY*. Please add *TERRY* as an Alias-First Name. The same would be true for William and Bill, or James and Jim. Also, when a patient has a change in surname (marriage, divorce, etc), please record the new current name in the Surname field and record the old name in the Alias-Last Name field.

## Business Rules Management Solutions - A CCR Project

Cheryl Moody, BA, CTR  
CCR Data Standards and Quality Control Unit

### Background

California Cancer Registry (CCR) reduced visual editing of incoming Admissions from 100% to 40%. This decision precipitated a new approach to verifying data quality in the state database. A project team was developed to create a business model for writing, programming, and implementing automated business rules as a quality control tool.

### The Business Rules Management Solutions Project

Project team members consisted of three Certified Tumor Registrars functioning as business rule writers and one Programmer. Initial efforts focused on verifying Admission level data as analytic or non-analytic. The data field "Class of Case" was selected as the key data field for this determination. Class of Case was prioritized and the decision made to begin with Class 49 (Death Certificate Only) cases.

### Rule Writing

The Rule Writers decided which data fields were typically reviewed during visual editing to confirm Class 49. Those data fields became the foundation for the automated effort. Below are the primary data fields Rule Writers felt were integral to the Class 49 rule set:

- Date of Diagnosis
- Date of Last Patient Contact
- Type of Report Source
- Vital Status
- Hospital Number
- Diagnostic Confirmation

In addition to auto-coding and correcting Class 49 admissions, the rule set is also designed to identify existing Class 49 edits that are applicable. Rule Writers will analyze existing edits and request additional edits as appropriate.

### **Programming**

Rule Writers and Programmer collaborated to transition rules from a written rule set into a programmed code applied to the state's database at every level of processing. The code is designed to determine if the admission meets all requirements of the rule set or fails at one or more of the required data items. Incorrectly coded cases are auto-corrected to the appropriate code. In the rare instance where two or more correct codes are feasible, the case is auto-sent to a visual editor for review.

### **Queries**

Once rules are written and programmed, query results are obtained to determine the following:

- Total number of Class 49 cases in the database
- Total number of errors for each data field

Query totals are compared to programmer's test results. Rules and/or code are amended as appropriate until totals are identical.

There were 35,742 Class 49 admissions in the database at the beginning of this project with errors totaling 1,645 (5%). Identified errors were corrected either through a global fix or manually prior to implementation of the automated Class 49 Rule Set.

### **Testing**

The programmed rules are then applied to the database in a test environment. Rigorous testing is performed at every feasible level of processing to determine Rule Set functionality. Rule Writers create test cases and analyze test results at the following levels:

- File Upload
- Visual Editing
- Corrections
- Follow-up
- Database Inquiry

Following the comprehensive quality assurance testing, subject matter experts are invited to test the rules. Feedback from both the QA and Peer review tests are analyzed and rules and/or code amended as needed. Approval to implement is obtained from the Data Standards and Quality Control Unit of the CCR. The Class 49 rule set was successfully implemented in July 16, 2009. The original rule set was modified to expand the auto-change options to include the following data fields: Class of Case,

Hospital Number, Date of Diagnosis, Date of First Admission, Date of Last Patient Follow-up, Type of Report Source, and Diagnostic Confirmation. The revised rule set was implemented in July 2010. A third revision is in process to add Treatment Date Flags as a result of the 2010 conversion. Implementation is planned for June, 2011.

### **Summary**

In addition to Class 49 rules, the project team has completed writing and programming Class 38 (autopsy rules (Autopsy Only) which were successfully implemented in July 2010. Work is currently underway to develop Class 43 (Pathology or Lab Specimen Only) rules. Additionally, Volume III has been reviewed for potential auto-change rules which are also in progress.

Post-implementation analysis of the Class 49 admissions in the database revealed 36,912 cases with a 0% error rate.

### **Educational Opportunities**

104B July 27, 2011, Regional Meeting, CPIC, Fremont. Contact: Shawky Matta at [Shawky.Matta@cpic.org](mailto:Shawky.Matta@cpic.org)

105B November 2-4 , 2011, 38th CCRA Annual Meeting, Culver City. Contact: Maia Bohm at [Bohm.Maia@hcahealthcare.com](mailto:Bohm.Maia@hcahealthcare.com)

### **In Memmorial, Lori Lenberg, CTR**

Lori Lenberg, CTR, longtime cancer registrar at Kaiser's San Diego facility and the Kaiser regional office in Pasadena, died June 26 after a long battle with cervical cancer. Per her family wishes, contributions can be made in her name to the American Cancer Society.