

# EDIT ERRORS ON FILE UPLOAD: USE OF REPORT ANALYSIS TO IMPROVE DATA QUALITY

Jenna Mazreku, CTR

Senior Systems Analyst, Production Automation & Quality Control  
California Cancer Reporting and Epidemiologic Surveillance (CalCARES) Program\*  
Institute for Population Health Improvement, UC Davis Health System



## BACKGROUND

In an effort to improve data quality, the Production Automation and Quality Control (PAQC) unit of the California Cancer Reporting and Epidemiologic Surveillance Program (CalCARES) began a focused analysis of edit errors identified when admission level abstracts are uploaded into Eureka, our integrated cancer database management system. The targeted months used in this analysis were August, September, and October of 2013. The *Edit Errors on File Upload Summary* report was utilized to capture these edit errors and provide documentation for how many initial edits are received on an admission.

## OBJECTIVE

The primary goal of this analysis was to find a strong business solution that would meet central and regional registry needs. The PAQC unit also looked for new opportunities for automation and target areas in need of education. Focus was also put on how the existing Edit Errors on File Upload report could be improved for ease of use.

## METHODOLOGY

The *Edit Errors on File Upload Summary* report was analyzed each month by the PAQC unit. The date parameters were set consistently and the report was used for analysis by being exported as a ".csv" file. For prospective analysis, the report was filtered to target the top edits overall.

The admissions associated with these edits were identified and utilized to determine if there were any additional edits associated with the admissions. This pool of data was then analyzed for trends between edits firing and their relation to admissions, facilities, vendors, and regions. A second analysis was performed retrospectively looking at edit errors that have been addressed in previous reports. They were monitored to ensure that proposed solutions were still accurate or to help determine if further action was needed.

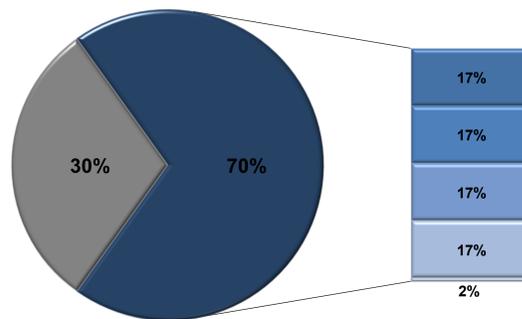
## OVERALL RESULTS

### AUGUST ANALYSIS

Focusing on the prospective analysis for the initial month, the *Edit Errors on File Upload Summary* report 7/31/13 – 8/31/13 showed that the top four edits were:

- ER1114: Addr at DX--Country
- ER1115: Addr Current--Country
- ER1116: Birthplace--State
- ER1130: Birthplace--Country

**Top Edit Errors: August**



These four edits accounted for 69% of the edits fired in the month of August and were firing on the same 140 admissions. There was also an additional edit that was firing on 10 of these cases, IF1046: Place of Death--State, Vital Status, which brought the percentage up to 70%. It was determined that all 140 cases were abstracted at the same Reporting Facility using the same software vendor. It was determined that this facility not only needed educational feedback on the allowable values for the fields, but also showed that they were able to complete cases with edit errors.

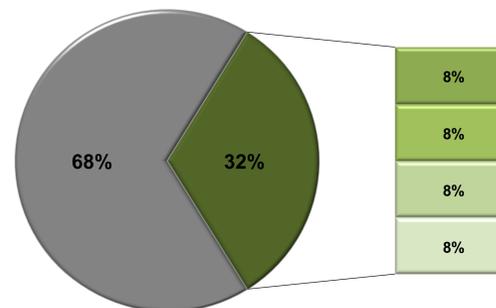
### SEPTEMBER ANALYSIS

Beginning with the prospective analysis, the *Edit Errors on File Upload Summary* 8/31/13 – 9/30/13 analysis showed that the top four edits were:

- IF740: Mult Tum Rpt As One Prim, Date of Diagnosis
- IF739: Ambiguous Terminology DX, Date of Diagnosis
- IF741: Multiplicity Counter, Date of Diagnosis
- IF744: Date of Multiple Tumors, Date of Diagnosis

**Top Edit Errors: September**

Remaining Edits  
IF739: Ambiguous Terminology DX, Date of Diagnosis  
IF744: Date of Multiple Tumors, Date of Diagnosis  
IF740: Mult Tum Rpt As One Prim, Date of Diagnosis  
IF741: Multiplicity Counter, Date of Diagnosis



Together these edits fired on the same 69 admissions and represent 32% of the total edit errors for the month. The distribution of these cases was investigated for any patterns. Six abstractors were responsible for reporting the admissions to the central registry utilizing the same vendor, Vendor M. 97% of the admissions were transmitted by the same Reporting Facility, Hospital A, while the remaining 3% were transmitted by Hospital B. Both Reporting Facilities were located within the same region. This distribution re-emphasized that abstractors are able to bypass edits to complete their cases.

**Prospective: Distribution of September Top Edit Errors**

| Abstractor   | Reporting Facility | Vendor   | Region   | Number of Cases |
|--------------|--------------------|----------|----------|-----------------|
| Abstractor 1 | Hospital A         | Vendor M | Region R | 32              |
| Abstractor 2 | Hospital A         | Vendor M | Region R | 2               |
| Abstractor 3 | Hospital A         | Vendor M | Region R | 20              |
| Abstractor 4 | Hospital A         | Vendor M | Region R | 8               |
| Abstractor 5 | Hospital A         | Vendor M | Region R | 5               |
| Abstractor 6 | Hospital B         | Vendor M | Region R | 2               |
| Total Cases  |                    |          |          | 69              |

There are five fields involved with these edits: Ambiguous Terminology DX, Date of Mult Tumors, Date of Mult Tumors Flag, Mult Tum Rpt As One Prim, and Multiplicity Counter. They all follow the same reportability date parameters, and it was found that these admissions were not abstracted following the rules, which caused the edits to fire on upload. These edits could have been cleared through following either the edit message or coding instructions.

Taking a retrospective look at the top four edits analyzed in the previous month of August: ER1114, ER1115, ER1116, and ER1130. It was noted that all four edit errors carried over into this month's report. The edits followed the same pattern as previously reported with the four edits firing on the same 10 admissions. Together these edits accounted for 5% of the total edit errors for the month. The admissions were reported by different abstractors and Reporting Facilities, which shows that the problem was not isolated to a single abstractor, reporting facility, vendor, or region.

**Retrospective: Distribution of August Top Edit Errors**

| Abstractor    | Reporting Facility | Vendor   | Region   | Number of Cases |
|---------------|--------------------|----------|----------|-----------------|
| Abstractor 7  | Hospital C         | Vendor N | Region S | 1               |
| Abstractor 8  | Hospital D         | Vendor N | Region R | 4               |
| Abstractor 9  | Hospital D         | Vendor N | Region R | 2               |
| Abstractor 10 | Hospital E         | Vendor N | Region S | 3               |
| Total Cases   |                    |          |          | 10              |

### OCTOBER ANALYSIS

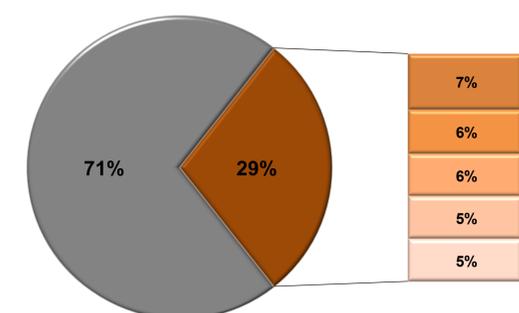
Beginning with the prospective analysis, the *Edit Errors on File Upload Summary* 9/30/13 – 10/31/13 analysis showed that the top four edits were:

- IF992: Class of Case, RX at Hosp
- IF885: CS Mets at DX--Lung, CS Mets at DX
- IF850: CS Mets at DX--Bone, CS Mets at DX
- IF852: CS Mets at DX--Brain, CS Mets at DX

It was decided to include IF875: CS Mets at DX--Liver, CS Mets at DX because of the edits relationship to the other CS Mets at DX edits.

**Top Edit Errors: October**

Remaining Edits  
IF885: CS Mets at DX--Lung, CS Mets at DX  
IF852: CS Mets at DX--Brain, CS Mets at DX  
IF992: Class of Case, RX at Hosp  
IF850: CS Mets at DX--Bone, CS Mets at DX  
IF875: CS Mets at DX--Liver, CS Mets at DX



It was found after analysis that together these edits spread across seven regions, fifteen reporting facilities, and seven vendor versions. This distribution was larger than the previous month's analysis, which just increased the fact that abstractors being able to bypass edits is not an isolated problem in any of the venues analyzed. Because of the coding relationship between the CS Mets at DX--Lung, CS Mets at DX--Bone, CS Mets at DX--Brain, and CS Mets at DX--Liver; it was decided that this was a great opportunity for using automation programming to correct the coding of these fields. Auto-Change Rules based on each edit were drafted and implemented in Eureka, which prevented these edit errors from increasing manual workloads through a programming effort.

Finishing the analysis with a retrospective look at of the top edit errors identified in the previous months, it was identified that these top edit errors remained active issues on the report. Despite the fact that these edits were firing less individually, together they still accounted for 20% of the total edit errors for the month.

**Retrospective: Distribution of August & September Top Edit Errors**

| Edit Number: Name                                  | Total for Edit Number | Percentage |
|--|-----------------------|------------|
| ER1114: Addr at DX--Country                        | 12                    | 1%         |
| ER1115: Addr Current--Country                      | 12                    | 1%         |
| ER1116: Birthplace--State                          | 12                    | 1%         |
| ER1130: Birthplace--Country                        | 12                    | 1%         |
| IF740: Mult Tum Rpt As One Prim, Date of Diagnosis | 38                    | 4%         |
| IF739: Ambiguous Terminology DX, Date of Diagnosis | 37                    | 4%         |
| IF741: Multiplicity Counter, Date of Diagnosis     | 38                    | 4%         |
| IF744: Date of Multiple Tumors, Date of Diagnosis  | 38                    | 4%         |
| Total Edits  | 199                   | 20%        |

## CONCLUSION

Analysis of initial errors has become an overall great resource for the CCR. The initial summary reports have been used to understand the relationship between edit errors and their associated admissions, vendors, facilities, and regions. This understanding has resulted in business discussions on the best way to resolve the issue:

- Resolution at the source through education
- Resolution at the back-end with software programming

From the perspective of the central registry, impacts on regional and central workload must be taken into account. In this situation, a collaboration of education and software programming is the best option at this point in time. The proposed solution starts with hospital transmits being rejected if they contain edit errors. The ability to tell when cases need to be rejected would require back-end work through programming, however it would be met halfway with Regional registries discussing with their hospital registries regarding the failures and provide education on the proper coding procedures when needed. This prevention of incorrect coding would improve data quality and decrease manual work efforts over time.