



STRATEGIES FOR DATA QUALITY BEYOND EDITS

PART 2 - WEBINAR

Fernanda Silva Michels, MSc, PhD, CTR
Program Manager of Data Quality and Integration - NAACCR


1

NAACCR REGISTRY OPERATIONS WORKSHOP: DATA QUALITY BEYOND EDITS MARCH 16, 2022

Interactive workshop for central registry directors and managers to share their experiences and ideas for non-standard data quality checks

111 attendees were divided by subgroups to discuss topics like:

| | | | | | |
|--|---|---|--|---|--|
| What unique QA reports or procedures do you routinely run? | How often do you perform this activity? | What software do you use to perform the activity? | Is there something you're checking all the time that could be an edit? | How much staff time does it take to perform the activity? | How do you identify cases that have already been reviewed? |
|--|---|---|--|---|--|



2

CREATION OF THE TOOLS



Based on the participants' expertise and examples



NAACCR developed a tool called: "Strategies for Data Quality Beyond Edits"



Available on [Registry Operations Guidelines - NAACCR](#)



3

3

Registry Operations Guidelines

CENTRAL REGISTRY STANDARDS

- CENTRAL REGISTRY STANDARDS
- DATA EXCHANGE STANDARDS & RECORD DESCRIPTIONS
- DATA STANDARDS & DATA DICTIONARY
- STANDARDS FOR COMPLETENESS, QUALITY, ANALYSIS, AND MANAGEMENT OF DATA
- STANDARDS FOR CANCER REGISTRIES, STANDARD DATA EDITS
- PATHOLOGY LABORATORY ELECTRONIC REPORTING
- XML DATA EXCHANGE STANDARD
- IMPLEMENTATION GUIDELINES
- REGISTRY OPERATIONS GUIDELINES**
- INTER-REGISTRY DATA EXCHANGE AGREEMENT
- MEANINGFUL USE

DATA QUALITY -

In March 2022, NAACCR held an interactive workshop titled "Data Quality Beyond Edits" for central registry directors and managers to share their experiences and ideas for non-standard data quality checks. Based on their expertise and examples, NAACCR developed these tools:

- pdf version (checklist) [Strategies for Data Quality Beyond Edits](#)
- excel version [Strategies for Data Quality Beyond Edits](#)

NAACCR TIP SHEETS +

TELECOMMUTING BEST PRACTICES +

DATA ITEM CONSOLIDATION MANUAL +

DEATH CLEARANCE MANUAL +

DISCHARGE AND CLAIMS DATA BEST PRACTICES GUIDE +

INTERSTATE DATA EXCHANGE GUIDELINES +



4

4

PDF VERSION

Print out and use as a check list

| Data Quality Beyond Edits | | |
|---------------------------|---|-------|
| Groups | Description | Check |
| Dx Date | Review diagnosis date at the turn of the year (ex. Dec/Jan with unknown day). | |
| Age | Review patients with current age >100 with vital status alive. More than likely patient is deceased, check death files or obituary search. | |
| Sex | Run NAACCR SEX Edits. Review/verify transgender cases. Check text to confirm sex code. Review male breast cases. Review text to confirm sex. | |
| Race | Review multi race patients against race priority rules. Use Appendix D of SEER Reporting Manual. Unknown race/ethnicity - review text, follow-up to reporting source. | |
| SSN | Review invalid SSN: Per SSA, no SSN has the first three digits "000," "666," or "900-999" or with the second group of two digits "00" or with the third group of four digits "0000".* | |
| Address | Look for blocks of addresses that might not be accurate (i.e. prisons). | |
| Tips | Linkages to improve all demographics: — Death files — Birth files — Motor vehicle files — Claims data | |

* 1. RM 10201.035 Invalid Social Security Numbers (SSNs). Social Security Administration.



5

EXCEL VERSION

Workshop: Data Quality Beyond Edits. It is not meant to lay out a specific methodology, but rather to serve as a guide for registries who wish to conduct more extensive quality control activities not validated by standardized edits. For more tools and resources visit www.naacccr.org.

Sources for Quality Control Task Resolution

Cases with discrepancies identified during quality control tasks can be resolved using a variety of sources. Time available for resolution may depend on registry workload and personnel resources.

1. **Check source record text** – coded items should match source record text and all applicable coding rules should be applied per the appropriate coding manual.
2. **Check with reporting facilities** – discrepancies that can not be resolved within the record may be followed back to the reporting facility for further investigation.
3. **Check with third-party sources** – discrepancies that can not be resolved within the record may be checked with other sources including data aggregators (i.e., Lexis Nexis), vital records data, linkage with other state agency data (i.e., hospital discharge data, drivers license records, voter registration) or claims data.

- Download the file
- Check the tabs
- Edit according to your needs



6

EXCEL VERSION

Data Quality Beyond Edits NAACCR

Site Histology

| Groups | Description | # of cases to review | Staff Assigned | Date started | Date Completed | Comments |
|---------------|---|----------------------|----------------|--------------|----------------|----------|
| Head and neck | Review site C00 with histology 8090 - 8110; verify primary should not be skin. | | | | | |
| Hemato. | Review ALL - Acute Lymphoblastic Leukemia (histology 9811-98xx) cases with age >25. | | | | | |
| C50 | Review site C50 (2018+) with histology 8010 (to see if should be recoded to 8500 per Solid Tumor Rules - STR). Review site C50 with histology 8543/2 (confirm if in situ). | | | | | |
| C53-C57 | Review site C559 for more specific code. | | | | | |
| C70-C72 | Review site C70-71 with Surgery Code = 40-55 to verify correct code assigned. Codes 40 and 55 are partial or total resections of lobe of brain. Not done very often. | | | | | |
| C80 | Review unknown primary sites (C809) - if your % of C809 is higher than expected. | | | | | |
| STR | Special studies when STR change; use changes from 2007 Multiple Primary and Histology rules (MPH) section to ID cases to review. | | | | | |

NAACCR 7

7

DATA QUALITY BEYOND EDITS NAACCR

Workshop: Data Quality Beyond Edits. It is not meant to lay out a specific methodology, but rather to serve as a guide for registries who wish to conduct more extensive quality control activities not validated by standardized edits. For more tools and resources visit www.naacrr.org.


Sources for Quality Control Task Resolution


Cases with discrepancies identified during quality control tasks can be resolved using a variety of sources. Time available for resolution may depend on registry workload and personnel resources.


1. **Check source record text** – coded items should match source record text and all applicable coding rules should be applied per the appropriate coding manual.
2. **Check with reporting facilities** – discrepancies that can not be resolved within the record may be followed back to the reporting facility for further investigation.
3. **Check with third-party sources** – discrepancies that can not be resolved within the record may be checked with other sources including data aggregators (i.e., Lexis Nexis), vital records data, linkage with other state agency data (i.e., hospital discharge data, drivers license records, voter registration), or claims data.

NAACCR 8

8





Data Quality Beyond Edits



Demographics


| Groups | Description | Check |
|----------------|---|--------------------------|
| Dx Date | Review diagnosis date at the turn of the year (ex. Dec/Jan with unknown day). | <input type="checkbox"/> |
| Age | Review patients with current age >100 with vital status alive. More than likely patient is deceased, check death files or obituary search. | <input type="checkbox"/> |
| Sex | Run NAACCR SEX Edits. Review/verify transgender cases. Check text to confirm sex code. Review male breast cases. Review text to confirm sex. | <input type="checkbox"/> |
| Race | Review multi race patients against race priority rules. Use Appendix D of SEER Reporting Manual. Unknown race/ethnicity - review text, follow-up to reporting source. | <input type="checkbox"/> |
| SSN | Review invalid SSN: Per SSA, no SSN has the first three digits "000," "666," or "900-999" or with the second group of two digits "00" or with the third group of four digits "0000".* | <input type="checkbox"/> |
| Address | Look for blocks of addresses that might not be accurate (i.e- prisons). | <input type="checkbox"/> |
| Tips | Linkages to improve all demographics: — Death files — Birth files — Motor vehicle files — Claims data | <input type="checkbox"/> |


* 1. RM 10201.035 Invalid Social Security Numbers (SSNs). Social Security Administration. <https://secure.ssa.gov/apps10/poms.nsf/lnx/0110201035>.




9

9





Data Quality Beyond Edits



Site Histology

| Groups | Description | Check |
|----------------------|---|--------------------------|
| Head and neck | Review site C00 with histology 8090-8110. It should be skin - make it not reportable. | <input type="checkbox"/> |
| Hemato. | Review ALL - Acute Lymphoblastic Leukemia (histology 9811-98xx) cases with age >25. | <input type="checkbox"/> |
| C50 | Review site C50 (2018+) with histology 8010 (to see if should be recoded to 8500 per Solid Tumor Rules - STR). Review site C50 with histology 8543/2 (confirm if in situ). | <input type="checkbox"/> |
| C53-C57 | Review site C559 for more specific code. | <input type="checkbox"/> |
| C70-C72 | Review site C70-71 with Surgery Code = 40-55 to verify correct code assigned. Codes 40 and 55 are partial or total resections of lobe of brain. Not done very often. | <input type="checkbox"/> |
| C80 | Review unknown primary sites (C809) - if your % of C809 is higher than expected. | <input type="checkbox"/> |
| STR | Special studies when STR change; use changes from 2007 Multiple Primary and Histology rules (MPH) section to ID cases to review. | <input type="checkbox"/> |



10

10

Data Quality Beyond Edits

Staging

| Groups | Description | Check |
|----------------------------|--|--------------------------|
| Stage | Review cases with unknown stage (TNM, EOD or Summary Stage). | <input type="checkbox"/> |
| Grade/ behavior | Review rare in situ sites (C25.0, C25.1, C56.9, C57.0, C62.9, C34) with behavior /2. | <input type="checkbox"/> |

11

11

Data Quality Beyond Edits

Deduplication and Multiple Primaries

| Description | Check |
|--|--------------------------|
| Review bilateral primaries to make sure not 2 primaries. | <input type="checkbox"/> |
| Resolution of duplicate patients and tumors. | <input type="checkbox"/> |
| Review patients with sequence number >08. | <input type="checkbox"/> |
| Run NAACCR de-duplication protocol using Match*Pro. | <input type="checkbox"/> |

12

12

»» Data Quality Beyond Edits

Additional visual review

| Description | Check |
|--|--------------------------|
| Review diagnostic confirmation coded to 4-9 to make sure a more specific or microscopic code could not be used. | <input type="checkbox"/> |
| Review laterality coded to 9 to see if a specific laterality can be assigned. | <input type="checkbox"/> |
| Review cases if difference between date of diagnosis and date of treatment is long to verify correct date entered - based on NCCN (National Comprehensive Cancer Network) guidelines | <input type="checkbox"/> |
| Review text for accuracy of coding for site C44 - Breslow depth - compare to T value | <input type="checkbox"/> |
| Review text for accuracy of lab value coding for site C619 - PSA lab value (and value closest & prior to bx date) | <input type="checkbox"/> |

13

13

⊘ Data Quality Beyond Edits

Do **NOT** do any of the following:

- Do **NOT** code vital status based on current age and/or prognosis.
- Do **NOT** code race based on address.
- Do **NOT** change a data item just to pass edits.
- Do **NOT** change data items without clear evidence (don't make assumptions).
- Do **NOT** assume that the most recently received record is the most complete/most correct one.
- Do **NOT** use override option just to pass the edit.



14

14


✓ Data Quality Beyond Edits NAACCR

Always do the following:

Review your standard setting organizations and/or funding agencies data quality report for the most recent diagnosis year.
Run frequencies by unknown primary site, race/ethnicity, histology, sequence number, sex, diagnosis date.
Review your overrides.
Work with your software vendor to flag reviewed cases.


15



QUESTIONS?

NEW NAACCR FORUM:
STRATEGIES FOR DATA
QUALITY FORUM

16



THANK YOU!

To all Cancer Registries directors and managers that shared their experiences and ideas with us back in March.

To the NAACCR Workshop Group:

- Betsy Kohler
- Ann Marie Hill
- Stephanie Hill
- Lori Havener
- Lori Koch
- Wendy Aldinger

FMICHELS@NAACCR.ORG