How Bridging the Gap Between VB6 and .NET Provides Efficiency for Automated Tumor Linkage Procedures in Registry Plus Central Registry Database Software

Michelle Esterly
Northrop Grumman Under Contract to NPCR

Joseph D. Rogers, Team Lead
Applications, Statistics, and Informatics Support Team (ASIST)

NAACCR 2016 Annual Conference
St. Louis, Missouri
Acknowledgments

Jennifer E. Seiffert
CDC/NPCR Contractor

Sanjeev Baral
CDC/NPCR Contractor

John Jordan
CDC/NPCR Contractor
Project Overview

- Background
- Automated Tumor Linkage Process
- Initial Deployment
- Transition to .NET
- Future Plans
Background

- Limited Automation for Tumor Linkage
- Registry Automation Needs Increasing Rapidly
- Records Sent for Manual Review
Initial Deployment

- External Program programmed in .NET
- Not optimal; allowed for partial automation and testing
- Abstracts needed to be processed in VB6 for Consolidation
Automated Tumor Linkage Process

- Abstract links to a patient and runs through the automated tumor linkage process
- Based on SEER's 2007 “Multiple Primary and Histology Coding Rules”
- Utilized Existing Table Structure to Draft Logic
  - Added tables to define logic
Rule Sets

Breast
Colon
Lung
Melanoma
Prostate
Rectosigmoid
Rectum

*Remaining rule sets will be added in the future*
## Exceptions

<table>
<thead>
<tr>
<th>Name</th>
<th>Histology</th>
<th>Site</th>
<th>Behavior</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leukemia_Lymphoma</td>
<td>9590-9992</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kaposi Sarcoma</td>
<td>9140</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breast (Paget Disease)</td>
<td>8530, 8540, 8541, 8543</td>
<td>C500-C509</td>
<td></td>
</tr>
<tr>
<td>Unknown Primary</td>
<td></td>
<td>C809</td>
<td></td>
</tr>
<tr>
<td>Non Reportable Squamous Intraepithelial Neoplasia</td>
<td>8077</td>
<td>C210-211,C51,C529,C53</td>
<td>2</td>
</tr>
<tr>
<td>Non Melanoma Skin</td>
<td>8000-8005, 8010-8046, 8050-8084, 8090-8110</td>
<td>C44</td>
<td></td>
</tr>
<tr>
<td>Non Reportable Cervix</td>
<td>8010, 8070</td>
<td>C53</td>
<td>2</td>
</tr>
</tbody>
</table>
Site and Histology Pairs

- NAACCR Site and Histology Pair Tables
  - NAACCR Tumor Linkage Workgroup reviewed pair tables used in several Central Registries for consensus
  - Developed to provide guidance for tumor linkage
NAACCR Site Pair Table

- Agreement Code Definitions
  - Pairs that correspond to written SEER rules for same site
  - Pairs that are a reasonable extension of the SEER rules
  - Pairs that experience has shown are likely same but may be handled differently by registries
  - Pairs that require special handling
Registry Plus Site Pairs

- Utilizes NAACCR Site Pair Table
- Assigned Match Strength 4 to indicate pair of codes often representing same tumor

<table>
<thead>
<tr>
<th>ID</th>
<th>KEY</th>
<th>SITE_PAIR_LOW</th>
<th>SITE_PAIR_HIGH</th>
<th>InglRRID</th>
<th>MatchStrength</th>
</tr>
</thead>
<tbody>
<tr>
<td>889</td>
<td>C199C199</td>
<td>C199</td>
<td>C199</td>
<td>30</td>
<td>1</td>
</tr>
<tr>
<td>890</td>
<td>C199C209</td>
<td>C199</td>
<td>C209</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>891</td>
<td>C199C260</td>
<td>C199</td>
<td>C260</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>892</td>
<td>C199C268</td>
<td>C199</td>
<td>C268</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>893</td>
<td>C199C269</td>
<td>C199</td>
<td>C269</td>
<td></td>
<td>4</td>
</tr>
</tbody>
</table>
Programmed Rules

- **Comparing Sites with Linkage Rules**
  IA with a site that does not have programmed linkage rules is compared to a DBT having a site with programmed logic - automated decision

- **Sites Without Programmed Rules**
  If none of the sites being compared have programmed rules - manual review required
Rule Set Timing Tables

- Based on SEER Timing Interval Rules
- Developed Rule Sets Associated with DX Timing Rule and Behavior Timing Rule
- If Behavior differs, time interval between DxDates and sequence of occurrence reviewed
Timing/Interval Rule

- Ranging from 60 Days to Five Years
- Intervals Expressed as Days

<table>
<thead>
<tr>
<th>Timing Rule ID</th>
<th>Days</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>60</td>
<td>60 days</td>
</tr>
<tr>
<td>2</td>
<td>182</td>
<td>6 months</td>
</tr>
<tr>
<td>3</td>
<td>365</td>
<td>1 year</td>
</tr>
<tr>
<td>4</td>
<td>1095</td>
<td>3 years</td>
</tr>
<tr>
<td>5</td>
<td>1825</td>
<td>5 years</td>
</tr>
</tbody>
</table>

- Ex. SEER Colon MP/H Timing Rule
  - **Rule M5** Tumors diagnosed *more than one (1) year* apart are multiple primaries.
General Description

- Linkage Rule Tables by Rule Set
- Values of Selected Data Items Compared
  - Primary Site
  - Laterality
  - Histology
  - Behavior
  - Diagnosis Date
  - Reporting Facility
## General Description

- **Code Designations**

<table>
<thead>
<tr>
<th>COMPARISON CODE</th>
<th>DEFINITION</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Same</td>
</tr>
<tr>
<td>2</td>
<td>Different</td>
</tr>
<tr>
<td>3</td>
<td>Maybe Same</td>
</tr>
<tr>
<td>0</td>
<td>Irrelevant</td>
</tr>
</tbody>
</table>

- **Results of Comparison Accumulated**

- **Accumulated Values**
  - Identify a specific row in table of linkage rules
  - Identified row contains Disposition
Comparison and Disposition

- Accumulated Comparison
  - Site Pair = C187
  - Histology = 8140
  - Laterality = 0
  - Behavior = 3
  - DxDate = 20151201
  - Reported by different facilities
Tumor Linkage Result

- **Automated Decisions**
  - Link to Existing Tumors
  - Create New Tumors

- **Manual Review**
  - Diagnosis Date Prior to 2007 or Unknown
  - Tumor Linkage not determined
# Codes for Failed Tumor Linkage

<table>
<thead>
<tr>
<th>CODE</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Failed tumor linkage – Diagnosed before 2007</td>
</tr>
<tr>
<td>2</td>
<td>Failed tumor linkage – Non-reportable benign tumor</td>
</tr>
<tr>
<td>3</td>
<td>Failed tumor linkage – Linkage not applicable/Exception</td>
</tr>
<tr>
<td>4</td>
<td>Failed tumor linkage – Rule set not implemented</td>
</tr>
<tr>
<td>5</td>
<td>Failed tumor linkage – Possible tumor link (Match Strength 4)</td>
</tr>
<tr>
<td>6</td>
<td>Failed tumor linkage – Linkage rule for combination of codes requires manual review</td>
</tr>
<tr>
<td>7</td>
<td>Failed tumor linkage – Links to more than one tumor</td>
</tr>
</tbody>
</table>
Calculating Sequence Number Central

- General Assumptions
  - Sequence Number Central of existing tumors correct and meets rules of registry
  - Existing tumors pass inter-record edits

- Identical Diagnosis Date
  - If date of IA is identical to DBT, Summary Stage reviewed
Calculating Sequence Number Central

- Sequence Number Hospital Groupings
  - Between 00 and 59
  - Between 60 and 88

- Assigned Number for Incoming Tumor Compared to Sequence Number Hospital
Calculating Sequence Number Central

- **Auto Calculation**
  - Based on chronological order (Diagnosis Dates)

- **Re-sequencing of Remaining Tumors**
## Testing Results – Defined Rule Sets

<table>
<thead>
<tr>
<th>Site</th>
<th>Total Cases Run Through Automation by Site</th>
<th>% Automated Disposition Determined</th>
<th>% Manual Review</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prostate</td>
<td>222</td>
<td>97</td>
<td>3</td>
</tr>
<tr>
<td>Lung</td>
<td>883</td>
<td>86</td>
<td>14</td>
</tr>
<tr>
<td>Colon</td>
<td>399</td>
<td>83</td>
<td>17</td>
</tr>
<tr>
<td>Rectum/Rectosigmoid</td>
<td>160</td>
<td>77</td>
<td>23</td>
</tr>
<tr>
<td>Skin-Melanoma</td>
<td>359</td>
<td>69</td>
<td>31</td>
</tr>
<tr>
<td>Breast</td>
<td>1027</td>
<td>56</td>
<td>44</td>
</tr>
<tr>
<td>Total</td>
<td>3050</td>
<td>74</td>
<td>26</td>
</tr>
</tbody>
</table>
Transitioning to .NET

- CRS Plus re-written in .NET
  - Releasing June 2016

- Full impact of Tumor Linkage automation will be available

- Records that are linked, consolidated, and edited successfully sent to database
Future Plans

- Add Rules Sets
- Re-evaluate Exceptions
- Increase Automation
- Review/Modify as SEER MP/H Rules Updated
Questions

Thank You!

Joseph D. Rogers, Team Lead
JRogers@cdc.gov

Michelle L. Esterly
MEsterly@cdc.gov

Jennifer E. Seiffert
JSeiffert@cdc.gov