CancerCORE

A New Innovation in Central Cancer Registry Software

Presented by:
Catherine Tapp, MPH & Chris Fisher
Arkansas Central Cancer Registry

NAACCR – Boston 2005
History

- Utilized several applications since 1996
- Needs became more specific
  - Prompted the creation of CancerCORE
  - Allowed for direct control and manipulation
  - June 2003 was the partial release (abstracting and DCO)
  - November 2003 – full version released
Specifications

- Developed in Visual Basic 6
- Data are stored on MS SQL Server 2000 databases
- Database design was purchased from Scotts Hill Assoc.
CancerCORE Components & Demonstration
Search Function

- Use Soundex and NYSIIS to pull alike results
- Profile database – clean
- Archive database – data repository/static
- Suspense database – contains the records that must be reviewed and cleaned before moved to profile.
- Temporary/outstanding database – holds imported and abstracted data
  - Accumulated throughout the day and synchronized at night (speeds things up)
Search Cont.

- Facility/Physician database – contact information is stored here
- Non-reportable database – stores cases prior to 1996 or non-reportables
- Death Clearance database – stores the mortality file and process results
- Non-hospital reporting database – stores pathology, hospice, nursing home data
- Facility auditing database
- Record linkage results database
Database Records (Profile)

- Patient
  - demographic and follow-up information
- Cancer
  - lists all cancers for the patient
  - Can view all cancers
- Report
  - List of reporting information for a cancer
  - Ex.) facilities that submitted and when
Record Form

- Many integrity checks reveal conflicts between fields – automatically updates.
- Overall data integrity check – uses edit checks.
- Highlights fields with invalid/suspicious values.
- Everything is selection box oriented and direct entry is allowed.
- Many fields are calculated on ‘the fly’ i.e.) NHIA, collaborative staging derived codes, beale codes, age of dx.
### Cancer Database Record: FISHER, RIDDICK (Profile: 652981)

#### System Information
- **Report ID:** 652981
- **Date Report Created:** 04/15/2005
- **Last Change Date:** 04/14/2005 11:04:00 PM

#### Facility Information/Referrals
- **Reporting Facility:** 710400 Saint Vincent Infirmary Medical C
- **Registry ID:** 000000710 Arkansas Central Cancer Registry
- **Registry Type:** 1 Central Registry (population-based)
- **Facility Referred From:** N/A
- **Facility Referred To:** N/A
- **Class of Case:** 0 Dx At Rep. Facility, All 1st Course Rx E
- **Primary Payer:** 99 Unknown

#### Data Transmission
- **NAACCR Record Version:** B 2006 Version 11
- **Transmit Date:** 04/23/2005
- **Record Type:** A Full Case Abstract Record Type
- **Vendor:** CCORE1.0
- **Date Case Received:** 04/23/2005
- **Date Case Loaded:** 04/23/2005

#### Case Information
- **Facility Patient ID:**
- **Medical Record Number:** UNK
- **Tumor Record Number:**
- **Date Tumor Rec Available:** / / 
- **Accession Number:**
- **Sequence Number:** N/A
- **Abstractor:** CJF
- **Report Source:** 1 Hospital Or Clinic
- **Date Case Completed:** 04/23/2005
Abstracting

- Data entry form with selection boxes and code searching functions
- On-line integrity check
  - Conflicts between fields
- Interactive fields i.e.) no surgery of primary site performed…application will update all other surgery related fields (surgical margins, reg/distant surg summ, lymph node surg summ)
Abstracting Cont.

- Three primary areas (patient, cancer, report)
- Full abstracts not required if pt. already exists; can abstract cancer portion only
- Performs data integrity checks – runs EDITs before saving…cannot be saved before cleaned!
<table>
<thead>
<tr>
<th>Field</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Last Name</td>
<td>MURPHY</td>
</tr>
<tr>
<td>First Name</td>
<td>DALE</td>
</tr>
<tr>
<td>Middle Name</td>
<td></td>
</tr>
<tr>
<td>Name Suffix</td>
<td></td>
</tr>
<tr>
<td>Maiden Name</td>
<td></td>
</tr>
<tr>
<td>Alias</td>
<td></td>
</tr>
<tr>
<td>Birth Date</td>
<td>01/10/1954</td>
</tr>
<tr>
<td>Social Security No.</td>
<td>999-88-7777</td>
</tr>
<tr>
<td>Sex</td>
<td>1 - Male</td>
</tr>
<tr>
<td>Race</td>
<td>01 - White</td>
</tr>
<tr>
<td>Spanish Origin</td>
<td>9 - Unknown</td>
</tr>
<tr>
<td>Telephone Number</td>
<td>(999) 999-9999</td>
</tr>
<tr>
<td>Family History</td>
<td>9 - Unknown</td>
</tr>
<tr>
<td>Alcohol History</td>
<td>9 - Unknown</td>
</tr>
<tr>
<td>Tobacco History</td>
<td>9 - Unknown</td>
</tr>
<tr>
<td>Occupation Text</td>
<td>UNKNOWN</td>
</tr>
<tr>
<td>Industry Text</td>
<td>UNKNOWN</td>
</tr>
<tr>
<td>State of Residence</td>
<td>AR - Arkansas</td>
</tr>
<tr>
<td>Street Address</td>
<td>35 MATTHEWS LN</td>
</tr>
<tr>
<td>_sequence_number</td>
<td>00 - One primary in the patient's lifetime</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>diagnosis_date</td>
<td>01/01/2004</td>
</tr>
<tr>
<td>primary_site</td>
<td>C000 - External upper lip</td>
</tr>
<tr>
<td>morphology_ICD-O_3</td>
<td>8270/3 Chromophobe carcinoma (C75.1)</td>
</tr>
<tr>
<td>laterality</td>
<td>0 - Not a Paired Site</td>
</tr>
<tr>
<td>grade</td>
<td>1 - Grade I; Well Differentiated</td>
</tr>
<tr>
<td>report_source</td>
<td>1 - Hospital or Clinic</td>
</tr>
<tr>
<td>physician_managing</td>
<td>999999 Unknown Physician</td>
</tr>
<tr>
<td>cancer_status</td>
<td>9 - Unknown</td>
</tr>
<tr>
<td>diagnostic_conf</td>
<td>9 - Unknown</td>
</tr>
<tr>
<td>admission_date</td>
<td>99/99/9999</td>
</tr>
<tr>
<td>date_of_1st_biopsy</td>
<td>99/99/9999</td>
</tr>
<tr>
<td>casefinding_source</td>
<td>10 - Reporting Hospital, NOS</td>
</tr>
<tr>
<td>state_at_diagnosis</td>
<td>AR - Arkansas</td>
</tr>
<tr>
<td>address_at_diagnosis</td>
<td>35 WITMORE AVE</td>
</tr>
<tr>
<td>supplemental_addr</td>
<td>UNKNOWN</td>
</tr>
<tr>
<td>city_at_diagnosis</td>
<td>CABOT - LONOKE COUNTY</td>
</tr>
<tr>
<td>county_at_diagnosis</td>
<td>085 - Lonoke County</td>
</tr>
<tr>
<td>zip_code_at_diagnosis</td>
<td>999999-9999</td>
</tr>
</tbody>
</table>
Quality Assurance (QA)
Quality Assurance

Primary QA Tasks

- Patient De-Duplication
  - No Outstanding Duplicates to Resolve
  - 0 User(s) Currently Processing

- Cancer De-Duplication
  - There are Outstanding Duplicates to Resolve
  - 0 User(s) Currently Processing

- Review Suspense Records
  - 121 Record(s) Need to be Reviewed
  - 0 User(s) Currently Processing

- City/County Reconciler
  - No Reconciliation is Taking Place

Other QA Tasks

- CancerCORE Data Integrity Checker
Merging

- Algorithms for patient and cancer de-duplication
- Calculates match scores and reports to user the records to resolve
- Merging provides a side-by-side view of the main and the matching records
- Data integrity checks
- Users can copy information from match field to main field (via arrow button)
Merging Cont.

- Facility and contact information is displayed for user
- Fields with discrepancies are highlighted for easy identification
- Optional function to show only discrepancies
- Keeps a log of all records that have been designated as not duplicates so they are not pulled for future reconciliation
<table>
<thead>
<tr>
<th>Field Name</th>
<th>Primary Record</th>
<th>Match Record</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cancer Case ID:</td>
<td>603663</td>
<td>648075</td>
</tr>
<tr>
<td>Cancer Login Date:</td>
<td>7/6/2004 3:56:00 PM</td>
<td>1/5/2005 8:53:00 AM</td>
</tr>
<tr>
<td>Last Change Date:</td>
<td>11/29/2004 7:46:00 AM</td>
<td>10/29/2004</td>
</tr>
<tr>
<td>Date of Diagnosis:</td>
<td>01/24/2003</td>
<td>01/24/2003</td>
</tr>
<tr>
<td>Primary Site:</td>
<td>C530 Endocervix</td>
<td>C559 Uterus, Nos</td>
</tr>
<tr>
<td>Histology ICD-O 2:</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Behavior ICD-O 2:</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Histology ICD-O 3:</td>
<td>8380 Endometrioid adenocarcinoma, NOS</td>
<td>8383 Endometrioid adenocarcinoma, ciliated cell variant</td>
</tr>
<tr>
<td>Behavior ICD-O 3:</td>
<td>3 Malignant</td>
<td>3 Malignant</td>
</tr>
<tr>
<td>Grade:</td>
<td>2 Grade II; Moderately Differentiated</td>
<td>1 Grade I; Well Differentiated</td>
</tr>
<tr>
<td>Laterality:</td>
<td>0 Not A Paired Site</td>
<td>0 Not A Paired Site</td>
</tr>
<tr>
<td>Sequence Number:</td>
<td>02 2 Of Two Or More Primaries Or Non-malignant Tumors</td>
<td>02 2 Of Two Or More Primaries Or Non-malignant Tumors</td>
</tr>
<tr>
<td>Reporting Source:</td>
<td>1 Hospital Or Clinic</td>
<td>1 Hospital Or Clinic</td>
</tr>
<tr>
<td>Casefinding Source:</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Date of First Contact:</td>
<td>01/24/2003</td>
<td>01/24/2003</td>
</tr>
<tr>
<td>Abstractor:</td>
<td>WD</td>
<td>RW</td>
</tr>
</tbody>
</table>

Reporting Facilities:

<table>
<thead>
<tr>
<th>Facility</th>
<th>Date Last Sent</th>
<th>Facility Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>710440</td>
<td>07/06/2004</td>
<td>University of Arkansas F</td>
</tr>
<tr>
<td>710563</td>
<td>08/24/2004</td>
<td>White County Medical C</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Facility</th>
<th>Date Last Sent</th>
<th>Facility Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>710563</td>
<td>01/05/2005</td>
<td>White County Medical C</td>
</tr>
</tbody>
</table>
## CancerCORE Data Integrity Checker

**Integrity Checks:**

<table>
<thead>
<tr>
<th>Run</th>
<th>Multi-Field</th>
<th>Source</th>
<th>Check Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>☐</td>
<td>No</td>
<td>NAACCR</td>
<td>Addr at DX--State</td>
</tr>
<tr>
<td>☐</td>
<td>No</td>
<td>CoC</td>
<td>Birthplace</td>
</tr>
<tr>
<td>☐</td>
<td>No</td>
<td>SEER</td>
<td>Race 2</td>
</tr>
<tr>
<td>☐</td>
<td>No</td>
<td>SEER</td>
<td>Race 3</td>
</tr>
<tr>
<td>☐</td>
<td>No</td>
<td>SEER</td>
<td>Race 4</td>
</tr>
<tr>
<td>☐</td>
<td>No</td>
<td>SEER</td>
<td>Race 5</td>
</tr>
<tr>
<td>☐</td>
<td>No</td>
<td>SEER</td>
<td>Sex</td>
</tr>
<tr>
<td>☐</td>
<td>No</td>
<td>SEER</td>
<td>Spanish/Hispanic Origin</td>
</tr>
<tr>
<td>☐</td>
<td>No</td>
<td>NAACCR</td>
<td>Summary Stage</td>
</tr>
<tr>
<td>☐</td>
<td>No</td>
<td>SEER</td>
<td>Behavior</td>
</tr>
</tbody>
</table>

### Outstanding Integrity Checker Results

- **Run Date:** 5/3/2005 11:15:00 PM
- **Passed:** 1,022,845
- **Failed:** 383
- **Skipped:** 1,488
- **Resolved:** 2
- **Remaining:** 381

The integrity checker will run a number of interfield checks to validate field values. Those fields that fail are marked to be resolved by the user. The user has the ability to select which checks to run using the integrity check list.
Failed CancerCORE Data Integrity Checks to Resolve

Record Information

<table>
<thead>
<tr>
<th>Patient ID:</th>
<th>63465</th>
<th>Item Number:</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient Name:</td>
<td>[redacted]</td>
<td>Total Items:</td>
<td>235</td>
</tr>
<tr>
<td>SSN:</td>
<td>[redacted]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Checks Failed:</td>
<td>2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Case ID: 200584
Check Name: Histologic Type (COC)
Fields: Histology ICD-O 2 (2600)
Description: Not a valid value for histology ICD-O 2.

Case ID: 200584
Check Name: Histologic Type ICD03 (COC)
Fields: Histology ICD-O 3 (2600)
Description: Not a valid value for histology ICD-O 3.
Facility Auditing Tool

- Provides users a way of organizing, scheduling and performing QA (reabstraction) and case-finding audits
- Work flow chart to better understand how the audit process works for a particular audit.
- Calendar available for scheduling audits
Case-finding Audit

- System will take records to cross match with the database
  - Determines what was/was not reported
  - User receives report for follow-up
- Reports are generated with the final results
  - These are provided to audited facility
QA (Reabstracting) Audit

- System automatically generates random list of records to reabstract
- Option of blinded or non-blinded audit
- System highlights discrepancies (minor or major) as data are entered
- Upon reconciliation, user can select reasoning for the discrepancy (i.e. abstractor error, auditor error)
- Detailed summary report is generated
  - # of discrepancies, type, overall accuracy %
Audit Schedule
Schedule and view upcoming or previous casefinding and reabstracting audits.

Casefinding Audit
Population-based central cancer registries are mandated to collect data on cancer and other reportable tumors for residents of the geographic area. The central cancer registry must verify that all reportable tumors diagnosed in their residents are included in the central registry database. One process used to monitor case completeness for the registry and for individual reporting facilities is a casefinding audit. A casefinding audit examines casefinding procedures and evaluates completeness of case ascertainment. The casefinding audit may be performed to evaluate completeness of case ascertainment for an individual reporting facility, or it may be performed to evaluate completeness of case ascertainment for the central cancer registry.*

Reabstracting (QA) Audit
Population-based central cancer registries are mandated to collect data on cancer and other reportable tumors for residents of their geographic area. For the data to be useful in research, the information must be accurate and consistent. Reabstracting audits are performed to evaluate the accuracy and validity of submitted data.*

Exit Facility Auditing Tool

Casefinding Audit

Scheduling and notification of the facility to be audited is necessary to ensure a smooth process with as little inconvenience to the facility as possible.

Data entry should be conducted at the facility from original sources. Records will have to be synchronized from the laptops to continue the audit process.

1) Scheduling and Notification
2) Review Sources and Data Entry
3) Synchronize (if necessary)
4) De-Duplication of Audit Records
5) Record Linkage with Central Registry Database
6) Preliminary Reports and Resolution
7) Final Results

Reconciliation can only be performed at the central cancer registry. All audit records are needed as well as access to most of the CancerCORE databases.

Automatic e-mail notifications are sent to the facility contact for preliminary and final audit results.

Process Planning: "yellow"
Compiling: "green"
Reconcile: "blue"
Complete: "red"

All Casefinding Audits:

<table>
<thead>
<tr>
<th>Facility Name</th>
<th>Facility No.</th>
<th>Audit Date</th>
<th>Audit Type</th>
<th>Status</th>
</tr>
</thead>
</table>

List All Audits
List Upcoming Audits
Open Audit
Death Clearance

- Interactive flow chart for processing
- Once a process is completed it is marked and the next process becomes available
- DC can be performed on more than one death certificate year at a time
- Can search and view all death certificates
- Record linkage performed between profile database and death certificates
  - List of records to be reconciled as matches or non-matches is generated
Death Clearance Cont.

- Once reconciliation is complete, all true matches are automatically updated with vital status information and those fields that are unknown in the database with those on the death certificate i.e.) ss #, occupation text.
- Linkages between false matches that are cancer deaths are performed with the non-reportable, non-hospital and archive databases to lower burden of follow-back.
- Remaining records are linked with discharge summaries to provide additional information to improve follow-back.
Death Clearance Cont.

- Follow-back processing provides a system to flag records as contacted
- User can mark records as DCO, MDO or discovered true match
- System automatically abstracts those DCO’s based on the death certificate.
Death certificates must be received from Vital Stats. The death certificates can then be imported into the CancerCORE death certificate database. This database contains all certificates that have been imported in the past and can be searched but not modified.

Once the death certificates for the process year have been imported the record linkage with the profile database can be performed.

The record linkage will perform a probabilistic match to determine if the deceased exist in the profile database. Possible matches must be reviewed to determine if they are false or true matches. True matches and patient-only matches can then be used to update the vital status information within the patient’s record in the profile database.
Discharge summaries are received from health stats. The discharge summaries contain all discharge information from facilities in Arkansas.

False matches and patient only matches are then linked with the archive, non-reportable, and pathology databases. This will determine if the record is relevant and should continue to the follow-back process.

You can mark summaries as irrelevant.

Follow-up contact remaining. Not matching determines those marked as not matching.
CancerCORE II 1.1 - Arkansas Central Cancer Registry Database Management System

Death Clearance (PORTED AND STRIPPED TEMPORARY VERSION)

Select Process Year: 2001

Search Death Certificate DB

Irrelevant Records

Relevant Records (Ready)

Discharge Summary (Done)

Record Linkage (Done)

Follow-Back

Non-Reportable/ Irrelevant

Submitted/ Abstracted

MDO (Physician Only)

DCO

Finished/Summary

When follow-back and DCO auto-abstracting is complete then the process can be officially flagged as complete for the process year.

Follow-back is a very time consuming process of contacting physicians and facilities about the remaining cancer-related death certificates that did not match within the profile database. The goal is to determine how to process the remaining records. Those records that cannot be determined are marked as DCO (death certificate only) and are auto-abstracted to the profile database.

Discharge summaries are received from health facilities. The discharge summaries contain all discharge information from facilities in Arkansas.

You can go back up to 3 years of discharge summaries to link with the death certificates remaining for follow-back.

Note: Matches are then linked, and pathology is determined. If the record is relevant, the record is relevant to the follow-back process.
Record Linkage Utility

- Links external data with any of the databases
- Linkage wizard provides choices of databases to link with
- Utility accepts many types of file formats i.e.) ASCII, Access, Excel
Analysis

- Query on Profile database or Static file
- Numerous selection criteria
- User defined queries
- Canned reports i.e.) case counts by facility/year, errata report for facilities, accession report
- Results are displayed in detailed summary reports that can be sorted, exported, and mapped
(PATIENT.LASTNAME =N'fisher')
Future Plans

- Redeveloping application in VB Net
- Allow users in field access to the database from any location
- Annual report generating module
- Cluster tracking system
- Enhanced GIS integrated utility
- Additional management reports
Acknowledgements

- Ken Hill of Scotts Hill Associates
- Eric Durbin of KCR
- NAACCR for the death clearance and facility auditing procedures (Registry Fundamental CD’s)
- CDC/NPCR
Thank You