

Growing Pains: Lessons Learned from the Implementation of the NAACCR v12 Record Layout



David O'Brien, PhD, GISP
Data Analyst

Alaska Cancer Registry

Section of Chronic Disease Prevention & Health Promotion
Alaska Department of Health and Social Services



Alaska Division of Public Health
Prevention Promotion Protection

Alaska Cancer Registry



What is the NAACCR Record Layout?

- File format used to exchange (most) cancer registry data in the U.S. and Canada
 - Most path lab data being exchanged in HL7 and pipe-delimited formats
- “Flat” text file, one patient record per line
- Fixed-field width (fields are not delimited)
- Managed by NAACCR’s IT & Uniform Data Standards (UDS) Committees



What's New With v12?

- All date fields changed
 - Now in YYYYMMDD format
 - Only valid years (no 9s, 8s, or 0s)
 - Blanks indicate unknown day, month, year
 - New date flags indicate why date is blank
 - Only affects how dates are exchanged as part of Record Layout, not how stored in the registry database



What's New With v12?

- 126 new data items (incl. 26 date flags)
- 14 newly retired data items
- Total of 489 active data items, & 15 “reserved”
- Most text fields expanded – 16 fields now have 1000 characters (mostly Dx & Tx text)
- As a result of these changes, no. of characters per line more than tripled to 22,824 !!!





Challenges encountered during implementation

Remember...

Conference theme is “Keeping pace w/...technology”



Alaska Division of Public Health
Prevention Promotion Protection

Alaska Cancer Registry



New Date Format

- All computer operations involving dates (calculations, queries, etc.) using the NAACCR record layout had to be reviewed and most were revised
 - Old operations expecting MMDDYYYY, not YYYYMMDD
 - Year queries had to be changed from Right(...,4) to Left(...,4).



Longer Line Length

- Many text editors cannot edit and save 22,824 characters per line
- Popular editors that can't properly edit v12 files:
 - Programmer's File Editor (freeware) truncates each line at 16,385 characters
 - UltraEdit (\$60) wraps line at 4096 characters -- each line is displayed as six separate lines, so editing is difficult



Text Editors for v12 Files



■ Notepad++

– Download from

<http://notepad-plus.sourceforge.net/uk/site.htm>

– Freeware

■ TextPad

– Download from <http://www.textpad.com>

– Available for purchase for about \$25

– Free trial version available for download



Importing a v12 File into MS Access

- File>Get External Data>Import>Text Files
- Click on “Advanced” command button
- Click on “Specs” command button
- Choose the NAACCR v12 import specification



Naaccrin_2010v12 : Database (Access 2000 file format)

Open Design New

- Objects
- Tables
- Queries
- Forms
- Reports
- Pages
- Macros
- Modules
- Groups
- Favorites

Create table in Design view PrvMar09v12_2008

Import Text Wizard

Your data seems to be in a 'Fixed Width' format. If it isn't, choose the format that more correctly describes your data.

Delimited
 Fixed Width

Sample data from

```
1 A31
2 A31
```

Advanced...

PrvMar09v12_2008Test Import Specification

File Format: Delimited Fixed Width

Field Delimiter: []
 Text Qualifier: (none)

Language: English
 Code Page: OEM United States

Dates, Times, and Numbers

Date Order: MDY Four Digit Years
 Date Delimiter: / Leading Zeros in Dates
 Time Delimiter: : Decimal Symbol: .

OK
 Cancel
 Save As...
 Specs...

Field Information:

Field Name	Data Type	Start	Width	Indexed	Skip
Field1	Text	1			
Field2	Long Integer	17			
Field3	Text	95			
Field4	Text	145			
Field5	Long Integer	156			
Field6	Long Integer	166			
Field7	Long Integer	175			

Import/Export Specifications

Import/Export Specifications:

- Naaccr 10
- Naaccr 12**

Open
 Delete
 Cancel



What, you don't have NAACCR v12 import specs?

**DON'T
PANIC!**



Alaska Division of Public Health
Prevention Promotion Protection

Alaska Cancer Registry



Creating v12 Import Specs

- 489 data items, but 255 columns max, so need to combine data items
- Refer to NAACCR Standards Vol 2, Data Dictionary, Ch 7: Record Layout Table
- Start entering information into the fixed width Specs table
- Make large text fields (>255 char) into memo fields
- And after several hours...



Microsoft Access

File Edit View Insert Format Records Tools Window Help Adobe PDF

Type a question for help

Naaccrin_2010v12

Open Design

Objects

- Tables
- Queries
- Forms
- Reports
- Pages
- Macros
- Modules

Groups

- Favorites

1NAACCR12 Specs : Table

Field Name	Data Type	Start	Width	Indexed	Skipped
RecordType	Text	1	1	No	0
RegistryType	Text	2	1	No	0
FINCodingSystem	Text	3	1	No	0
Reserved00	Text	4	13	No	0
NAACCRRecordVersion	Text	17	3	No	0
NPIRegistryID	Text	20	10	No	0
RegistryID	Text	30	10	No	0
TumorRecordNumber	Text	40	2	No	0
PatientID	Text	42	8	No	0
PatientIDHosp	Text	50	45	No	0
DxCity	Text	95	50	No	0
DxState	Text	145	2	No	0
DxZip	Text	147	9	No	0
DxCounty	Text	156	3	No	0
CensusInfo	Text	159	17	No	0
MaritalStatus	Text	176	1	No	0
Race1	Text	177	2	No	0
Race2	Text	179	2	No	0
Race3	Text	181	2	No	0
Race4	Text	183	2	No	0
Race5	Text	185	2	No	0
RaceCoding	Text	187	2	No	0
Spanish	Text	189	1	No	0
ComputedEthnicity	Text	190	2	No	0
Sex	Text	192	1	No	0
AgeatDiagnosis	Text	193	3	No	0
DOB	Text	196	10	No	0
Birthplace	Text	206	3	No	0
OccupationInfo	Text	209	209	No	0
NHIA	Text	418	1	No	0
RaceNAPIIA	Text	419	2	No	0

Record: 1 of 252

Datasheet View

NUM

NAACCR
v12 specs!



OR... Just call me and I'll send you my import specs!



Alaska Division of Public Health
Prevention Promotion Protection

Alaska Cancer Registry



Software Upgrades: Central Registry Database

■ Vendor:

- Added the new data items to the data dictionary
- Added new import/export file specification for Type A (abstract), C (confidential), and I (incident) files
- Upgraded Edits Metafile and CS
- Converted codes in existing database records



Software Upgrades: Central Registry Database

■ User:

- Added 126 new data items to data entry grid in logical places
- Removed 14 retired data items from data entry grid
- Added state-specific fields to new v12 file specs
- Created descriptive name for each new data item; assigned default values if appropriate
- In Web Plus, developed “help prompts” for each new data item



Software Upgrades: Central Registry Database

■ User:

- Very time consuming process!
- Central level software: Most programs have at least 3 display types (source record, consolidated tumor, consolidated demographics)
- Hospital level software: Programs can have many display types (AK had 9 in Web Plus!) depending on type and size of the hospital or physician office



Software Upgrades: Central Registry Database

■ User:

- For AK, ODBC connection to the data dictionary in SQL Server through MS Access greatly reduced time in updating new data items by doing a lot of copy & paste



Software Upgrades: Central Registry Database

- Vendors QC their software, but users still find bugs
 - Users enter real-world data vs. vendor's test data

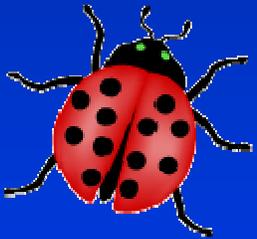


- » Example: Path Date Spec Collection
- » 14-digit date, but vendors used to 8-digit date
- » NAACCR Vol II data dict had YYYYMMDD
- » Uploading a 14-digit date into an 8-digit date field crashed the program!



Software Upgrades: Central Registry Database

- Vendors QC their software, but users still find bugs!
 - User's database slightly different configuration than vendor expected due to previously skipped minor upgrades
 - User might use the software differently than was done in the vendor's test environment
 - A vendor modification inadvertently breaks something else that was previously working just fine



Software Upgrades: Link Plus

- Link Plus currently cannot process more than 10,000 characters per line (new Beta version can)
- Need to be able to export a “Type C” file from your central registry database software
- Or... import a “Type A” file into MS Access, delete the Dx & Tx text fields to make a “Type C” file
- Or... create a custom text file with just the fields you need
- Or... use the Beta version



Software Upgrades: Link Plus

- Common practice to reuse Link Plus configuration files
- All reused config files need to have the date properties modified from MMDDYYYY to YYYYMMDD format
- Unknown date elements need to be modified from 9s to blank



Software Upgrades: eMaRC Plus

- eMaRC Plus translates HL7 and pipe-delimited files path files into NAACCR record layout
- User needs to add new data items to data entry grid & modify defaults to v12 standards
- In v11, big problem with not enough characters (about 200) to accommodate the descriptive text in the path reports
- In v12, descriptive text fields now 1000 characters, so much better for conversion of path files



Software Upgrades: SEER*Prep

- SEER*Prep processes your registry data w/state population data so it can be accessible in SEER*Stat
- Newest version has DD file for v12
- Some NAACCR file editing is needed prior to processing:
 - Add combined state-county FIPs
 - Add bladder behavior recode (2-->3)
 - Add state-specific fields
 - Update Date of Last Contact to death clearance date for survival analysis



Software Upgrades: SEER*Prep

- Run edited file through NAACCR's NHAPIIA SAS program to add derived Race/Ethnicity & create "Type I" file for import into SEER*Prep



Other Upgraded Software

- Needed to install new versions of:
 - EditWriter
 - GenEDITS
 - Inter-Record Edits Utility



Dates and Date Flags

- All Dates are exchanged in new standard, but can be stored in old standard in registry database
- If date and date flag are not edited together, then you can get a date that does not agree with the flag
 - Example: Known date & non-blank flag



Dates and Date Flags

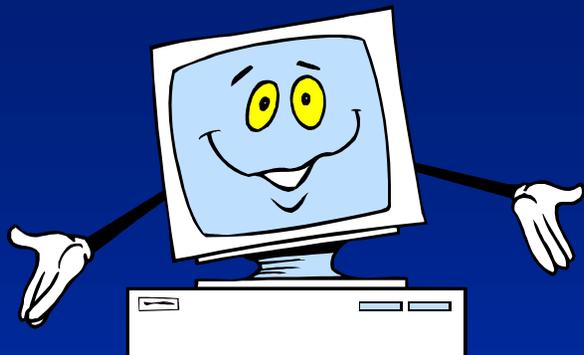
- When there is a known date & non-blank flag, software will blank out the date and report only the flag in exchange file!
- AK developed a routine that checks database for inconsistencies between date and flag



What Did We Learn?

- Still a lot of work for the user to upgrade registry software to v12 standards after vendor's upgrade
 - Added new data items to data entry grid, added descriptive names, default values, help prompts
 - Checked for bugs & worked with vendor to correct
 - Modified date-related SQL queries (left side, not right)
- Needed new tools to edit v12 files
- Needed to reinstall new utility programs compatible with v12 files & standards





Thanks very much!



Alaska Division of Public Health
Prevention Promotion Protection

Alaska Cancer Registry

