

LOOKING FROM THE INSIDE OUT: EVALUATING THE 2010 DATA CHANGES IMPLEMENTATION

B. Gordon and A. Houser - C/NET Solutions,
California Cancer Registry

Summary

- Where I'm coming from
- Signs of 2010 distress and success
- Why were the changes so complex?
- Vendor/registry issues
- Our goals and perspective
- User interface challenges
- Complexity in our future
- How we managed stress

Morphing this talk

- Continuing specs changes meant no time for reflection.
- Most states still in flux, too dynamic to track.
- Refocus on “Why was this all so hard?”

What we bring to the table

- We care about this as implementers of CS in national tools and hospital software. At the front end, we needed solid software out early in 2010.
- I also had the chance to participate in the informatics dialog and decision making on the CSV2 Informatics workgroup and NAACCR UDS.
- Alan Houser was a key player in the development of 2010 edits tools and helped with the NAACCR and CSV2 Implementation Guides.

Signs of 2010 Distress

- Seriously slipped deadlines
- Very late changes to specs

Slippage

Milestones	Months Slipped	Completed	Notes
2010 Data Items approved by UDS	2	Aug-08	Hema added Aug 2009
Project Requirements Doc finished	2	Jan-09	finalized 1/9/2009
AJCC-7 Chapters Released	4	Oct-09	
Final Schema Specs	6	Jan-10	some changes after
initial vendor release	3	Jan-10	
complete system testing	8	Feb-10	dll kept changing
2010 Conversion Specs	9	Mar-10	class conv changed 6/8/10
production release	7	Apr-10	next revision fall 2010
CDC Conversion Program	1	Jun-10	goal: 1 month after specs done
CSV2 Edits	8	Jun-10	schemas kept changing
Full implementation of CSV2	5	Jun-10	
CSV2 Implementation Guide	12	Jun-10	CSV2 IG will be released 6/16+
XML Schemas Proofed	12	Jul-10	started Jan 2010 with no budget

Very late changes to specs

- Class of Case conversion changed
 - Jerri Linn's email of June 8
- CSV2 Implementation Guide
 - June 16? target for update release
 - 'Clarifying' CS version fields
 - Updating SEER requirements section
 - Adding info on schema discriminators
 - Etc.

Late Changes, continued

- Surprise introduction by SEER in August of hematopoietic codes and rules impacted all registry implementations and training.
- Schema numbers were changed again in January 2010
- Staging for some schemas changed in Feb 2010

Late Deliveries

- AJCC-7 Chapters – Oct 2009
- CS software to states/vendors Feb 2010
- CDC Conversion program – June 2010
- CSV2 Implementation Guide – updates still coming

Signs of Success

- Good NCI-CDC partnering on CSV2 software development/testing
- Excellent regression testing by IMS
- Complex dll ready for vendor evaluation January 2010, in spite of late specs.
- Programming teams responded quickly to every change in specs, including CDC, IMS, Elaine C, Susan C, and Kathleen B

More signs of success

- Testomatic came out early and worked well
- The edits engine <-> csv2 dll interface was completed early so Susan could write edits.
- Many smart, dedicated people worked overtime to get things done. Thanks, especially, to Elaine and Jennifer.

Why was this year so hard?

Complexity issues

- Raw numbers – schemas (152) and tables (5000+)
- Couldn't know the inputs until mapping was complete
- NAACCR held back changes for several years, then the flood
- Late AJCC signoff added pressure to all teams.

Mapping & Implementation Complexities

- More schemas and SSF's and many more tables
- More 'extra' tables, adding complex table relationships
- More data inputs to CS calc (e.g. age, year dx, lymph invasion)
- New schema discriminator SSF₂₅ which must be coded first.
- Need to generate more stage outputs to meet the backwards compatibility requirements.
- New table attributes (usage, currency, role)
- Five different types of obsolete codes (flavors!)

How tables affected the timeline

- Took time to map new schemas and outputs from AJCC manual pages
- Took time to triple the number of tables (5000+)
- Extra time spent when mappers found issues:
 - Requiring new indirect tables or
 - Requiring a new schema split/new discriminator

For example: Five tables needed to derive Prostate T AJCC 7

- CS Extension-Clinical Extension
- CS TS/Ext Eval
- SSF 3 Extension - Pathologic Extension
- Special Calculation Table for TNM 7 Non-Invasive Pathologic Extension
- Special Calculation Table for TNM 7 Invasive/Unknown Pathologic Extension Eval

Vendor/registry issues

- Vendor issues with blank vs. 988 codes.
- How to handle 2010 cases entered with CSV1?
- Complexities when CSV1 SSF's were marked 'obsolete'.
- Conversion program and Genedit's not available until late.
- Pick lists must show longer explanations.
- Software delays threaten case on-time production.
- Registries must do pre-conversion data preparation and cleanup.
- Late decisions from standard setters on which SSF's are required .

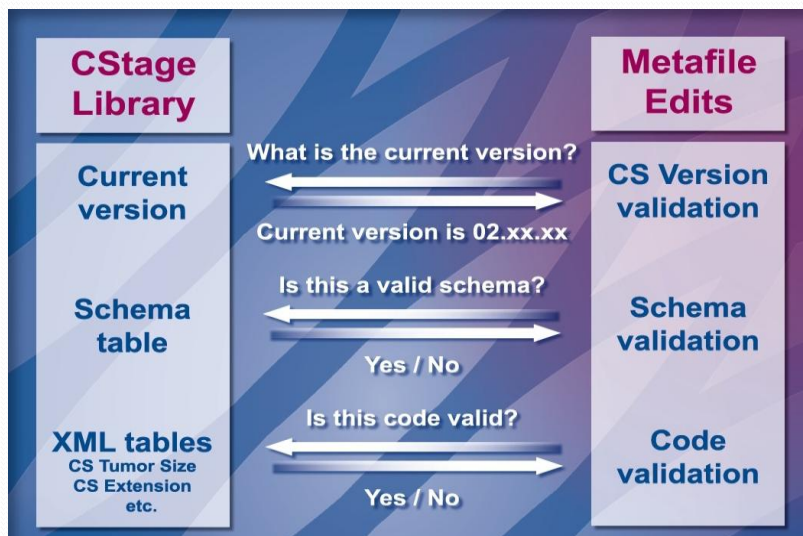
What C/NET Solutions and other vendors needed from CS Developers

1. Support pick lists from dll (no duplicated code lists)
2. Allow edits to call the CS dll (no duplication of rules)
3. Make sure extensive dll regression testing has been done (since rules are complex)
4. Let us and other vendors get an early look at dll interface to check for viability
5. Get early copy of conversion s/w to create test files.
6. Get a Genedit's version that can read NAACCR 12
7. Halt changes in specs in January 2010

Scorecard on goals

- ✓ CDC dll fully supports pick lists
- ✓ Edits <-> CSV2 interface completed by C/NET and CDC teams
- ✓ SEER-IMS regression testing was impressive
- ✓ Early look at dll worked
- ❖ We had to create our own conversion pgm
- ❖ We had to write our own edits batch 'Monster'
- ❖ Changes to specs were not halted

Edits Interface to CSV2



What C/NET team needed to do

- ✓ Convert EDITS tools to 2010 before edits could be written
- ✓ Convert big databases to 2010 for testing.
- ✓ Integrate CS manual into SmartHelp
- ✓ Deliver front end software before July 1
- ❖ We released CNExT last week

Our User Interface Challenge

- How to make CSV2 more palatable for users?
 - Skip unused SSF's
 - Give users smart defaults for CSV2 fields
 - Shield date format weirdness from users
 - Integrate CSV2 manual into SmartHelp

Automatic defaults example

The screenshot shows a software interface with a table on the left and a dropdown menu on the right. The table has two columns: 'Field Name' and 'Data'. The dropdown menu is titled 'Lung CS Site-Specific Factor 1' and 'Separate Tumor Nodules/Ipsilateral'. It contains a list of codes and their corresponding descriptions.

Field Name	Data
HER2/neu, Other	
01: Separate Tumor Nodules/Ipsilateral	
02: Visceral Pleural Invasion (VPI)/EI	
CS Site Factor-3	999
CS Site Factor-4	999
CS Site Factor-5	999
CS Site Factor-6	999
CS Site Factor-7	999
CS Site Factor-8	999
CS Site Factor-9	999
CS Site Factor-10	999
CS Site Factor-11	999
CS Site Factor-12	999

Code	Description
000	No separate tumor nodules in ipsilateral lobe
010	Separate tumor nodules in ipsilateral lobe
020	Separate tumor nodules in ipsilateral lobe
030	(020 + 010) Separate tumor nodules in same and different lobe
040	Separate tumor nodules, ipsilateral same or different lobe

One-click coding

This screenshot is identical to the one above, but with a yellow circle highlighting the code '030' in the dropdown menu.

Field Name	Data
HER2/neu, Other	
01: Separate Tumor Nodules/Ipsilateral	
02: Visceral Pleural Invasion (VPI)/EI	
CS Site Factor-3	999
CS Site Factor-4	999
CS Site Factor-5	999
CS Site Factor-6	999
CS Site Factor-7	999
CS Site Factor-8	999
CS Site Factor-9	999
CS Site Factor-10	999
CS Site Factor-11	999
CS Site Factor-12	999

Code	Description
000	No separate tumor nodules in ipsilateral lobe
010	Separate tumor nodules in ipsilateral lobe
020	Separate tumor nodules in ipsilateral lobe
030	(020 + 010) Separate tumor nodules in same and different lobe
040	Separate tumor nodules, ipsilateral same or different lobe

Complexity revisited

- CVS2 became too large to be perfectly executed, ever.
- Other sciences have faced this dilemma
 - Can't know both position and speed of particle
 - Some theorems can never be proven.
- We have to live with it...

One-click coding

Field Name	Data
HER2/neu, Other	
01: Separate Tumor Nodules/Ipsilateral	030
02: Visceral Pleural Invasion (VPI)/EI	
CS Site Factor-3	.998
CS Site Factor-4	.998
CS Site Factor-5	.998
CS Site Factor-6	.998
CS Site Factor-7	.998
CS Site Factor-8	.998
CS Site Factor-9	.998
CS Site Factor-10	.998
CS Site Factor-11	.998
CS Site Factor-12	.998

Code	Description
000	No evidence of visce Tumor does not con layer (PL 0)
010	Invasion beyond the limited to the pulmo Tumor extends throu
020	Invasion to the surfa Tumor extends to th (PL 2)

Complexity in our future

- Flood of new bio-markers will challenge how we manage that information.
- We need a better way to manage incremental changes in our data requirements
- Must have thorough vetting by NAACCR committees, not just the CRSC as happened with Hematopoietic.

Managing Stress?

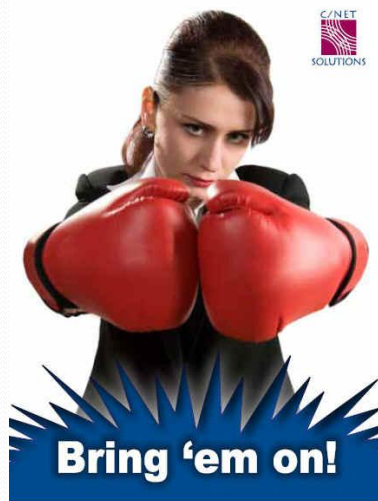
- Humor!



NCRA Registrar Comments While Punching 2010 Changes

- He's down!
- Somebody hasn't slit his throat?
- I think I have something sharp in my purse!
- I'm toast!

2010 Data Changes?



Thanks

- Kathleen Beaumont, for her crucial work on
 - the CS2-Edits interface
 - testing the CSV2 dll's and conversion specs
 - developing C/NET's conversion program and Edits Monster.