Managing and Migrating Large Imaging Databases

Fred M. Behlen, Ph.D.
LAITEK Inc., Homewood, IL USA
LAITEK Medical Software S.R.L., Cluj-Napoca
Work in data migration provides opportunity to witness the varied ways that the DICOM Standard is interpreted, misinterpreted or ignored in real-world implementations. ... the work is part consultancy, part janitorial.
• Feedback from the Real World of DICOM Implementations
  – The Right
  – The Wrong
  – The Weird
  – The Troublemakers
  – The Ambiguous
The Wrong

- Attributes in retrieved data not the same as in C-FIND
  - Study, Series and SOP Instance UIDs
  - On-the-fly transformations (“tag morphing”)
- Non-Unique UIDs
  - Standard says no two UIDs can be the same
  - Hierarchical Query Model can tolerate duplicates if
    - Series UIDs are unique within Study
    - Instance UIDs are unique within Series
- Over-length fields
  - (0018,0015) Body Part Examined (CS) > 16 char
- C-MOVE Final Status returned before sub-operations complete
• **C-FIND**
  – Multiple return records for the same Study
    • Different values in (0020,1208) Number of Study Related Instances, that may not add up to the total number of images in the Study

• **Duplicate copies of images**
  – Same Image, Same UID
  – Same Image, Different UID
  – Different Image, Same UID

• **Changing Study UID to match HL7 order**
  – Study is sent & stored, but can’t be found
The Troublemakers

- **Studies with Zero images**
  - Created by HL7 interface & forgotten
- **Closing of Studies**
- **Can’t C-FIND studies with null Study Date**
- **Increasing use of proprietary annotation**
  - Expendable if merely illustrative
  - Must be preserved if it contains clinically significant information, e.g. corrections to laterality markers
Point out the obvious

Laterality correction

Measurement calipers
• **Transient errors**
  - No problem if detected (simply re-fetch)
  - May be insidious in migration
  - Automatic detection needed
The Ambiguous

Cardinality of Accession Number to Study UID – two camps

1. 1:1 Accession Number uniquely identifies Study
2. 1:N Multiple studies may share same Accession Number
   N:1 (Grouped Procedures) variously addressed

⇒ Migrating from Camp 2 to Camp 1 requires Study merges
San Andreas Fault

RIS  PACS
Conclusions

- Fidelity of migration is key to long term data integrity
- Overall fidelity of migrations is high, thanks to DICOM
- DICOM compliance is improving
- More annotation being used, needs DICOM implementations for portability
- Some tightening of DICOM specs on Query/Retrieve behavior will help
- There will always be stuff to fix when data moves to a new system