Standardizing file formats, security, and integration of digital chest image files for pneumoconiosis classification

David Clunie

CTO, RadPharm, Inc.

Outline

- Scope and assumptions
- Types of digital data
- File format – DICOM
- Transfer and workflow
- Software compatibility issues
- Integrating reference images
- Integrating classification results with images
- Security and privacy

Scope and assumptions

- The task is to classify
  - PA projection chest x-rays
  - for pneumoconiosis
  - by trained and certified human readers (B readers)
- X-Rays acquired and distributed in digital form
  - not digitized film-screen exposures
- Readers not affiliated with/credentialed by acquisition site
  - may not be providing direct patient care but independent review
  - may need to use their “own” reading equipment
  - different equipment vendors for acquisition & reading equipment
- Readers may or may not have duty of care to patient
  - i.e., may or may not have legitimate access to patient’s identity (SSN)

Types of digital data

- Projection X-Ray image technology
- Computed Radiography (CR)
  - cassette based workflow – cassette “reader”
- Digital X-Ray (DX)
  - sensor is in X-Ray path
  - direct or indirect (phosphor + detector)
- Makes little difference in this context

File format - DICOM

- DICOM is ubiquitous
  - supported by all modern devices in all countries
  - global standard – international committee, ISO 12052
- Why DICOM and not TIFF, JPEG, etc.?
  - bit depth suitable for sensors (> 8 bit)
  - patient demographics in header
  - management information in header
  - technique information in header
- The only inter-vendor standard in use
  - between acquisition devices (modalities) and PACS

Which DICOM “flavor”

- Old “CR” object (SOP Class)
  - from the beginning (1993)
  - designed for CR
  - very loosely constrained attributes, type and grayscale
- New “DX” SOP Class
  - supports direct & indirect DX + CR
  - clearly distinguishes two types of image
    - for processing (raw, requires processing to view)
    - for presentation (processed, ready to view)
  - standard grayscale output space (PS 3.15 GSDF P-Values)
    - consistency on calibrated displays
  - makes mandatory many attributes
    - orientation, laterality, etc.

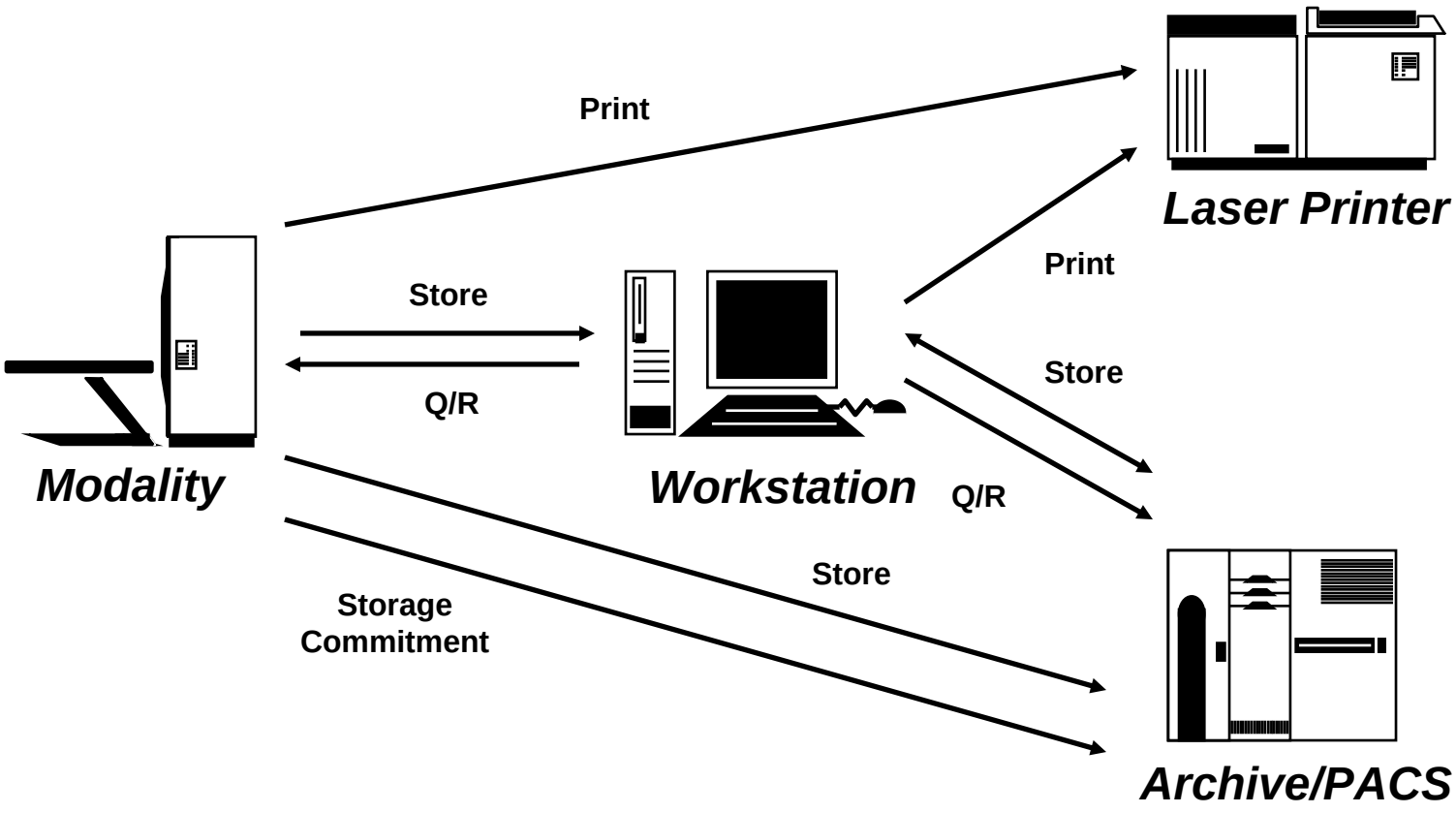
DICOM display implications

- DX “for presentation” images
  - can be reliably and consistently displayed on other vendor’s equipment
- CR images consistency depends on
  - type of image configured at acquisition
  - may need further “processing” by display
    - nice feature if have PACS from CR vendor
    - kills interoperability
- There are no standards for interoperable image processing
- Bottom line
  - insist on “processed” images (or both raw and processed)
  - requires *a priori* choice of processing type & parameters

DICOM – more than a file format

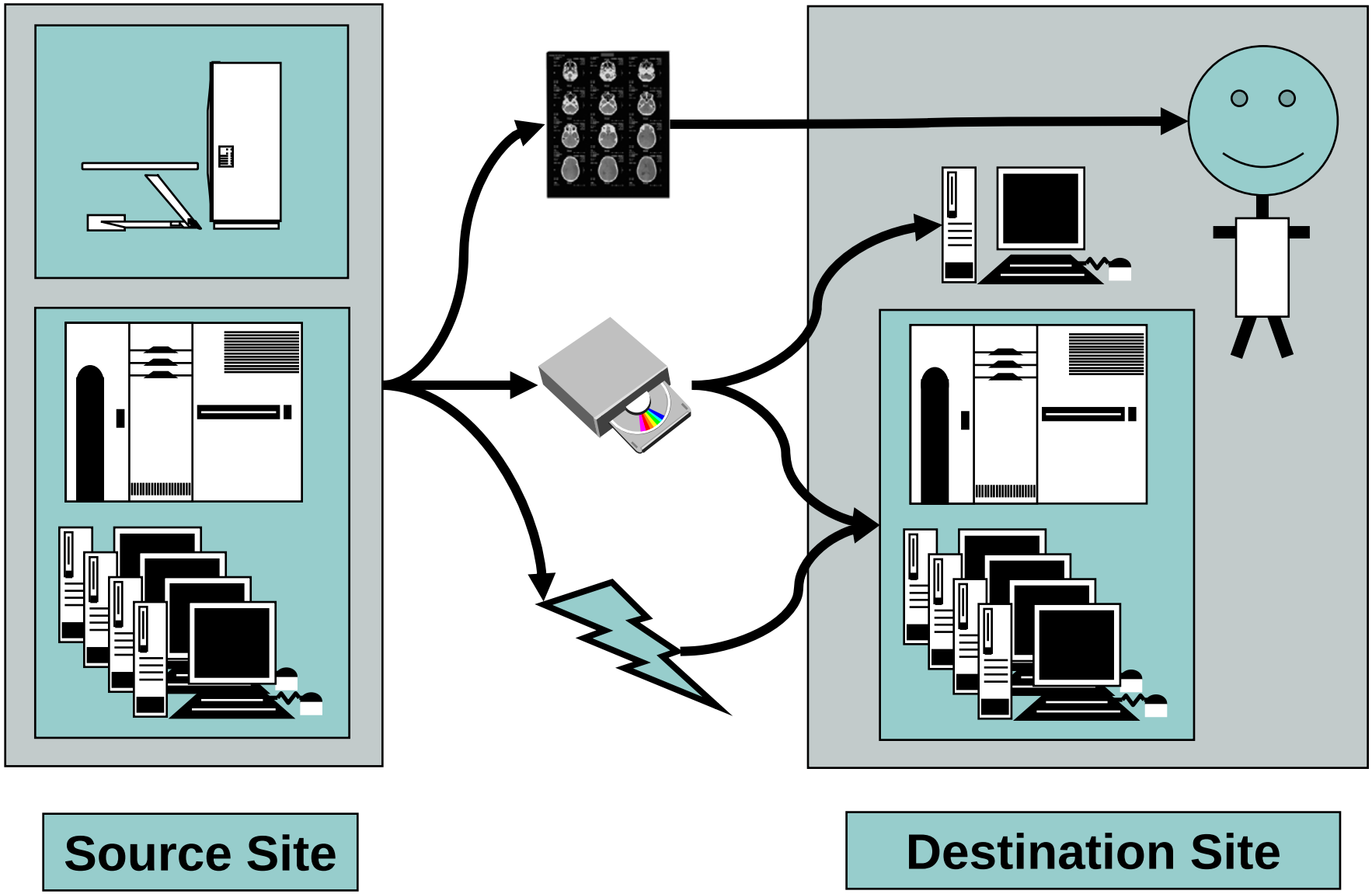
- Standard defines many services
  - transfer across network (local, Internet)
  - query for list of patients/studies etc.
  - retrieve patients/studies/series/images
  - work lists
  - printing
- Interchange Media (CD, DVD, USB)
  - “sneaker net” – carry or mail images on media
- Services allow automated interoperability
  - more than manual loading/dragging/dropping image “files”

DICOM Services



Transfer and workflow

- Modality acquires digital images
  - if PACS, DICOM transfer to PACS
  - if no PACS, burn DICOM files to CD
- Read images locally
  - read on workstation built-in to PACS
  - read on 3rd party DICOM workstation attached to PACS
- Read elsewhere - export from PACS
  - burn to DICOM CD
  - send via network to remote site
  - make available over network for viewing remotely

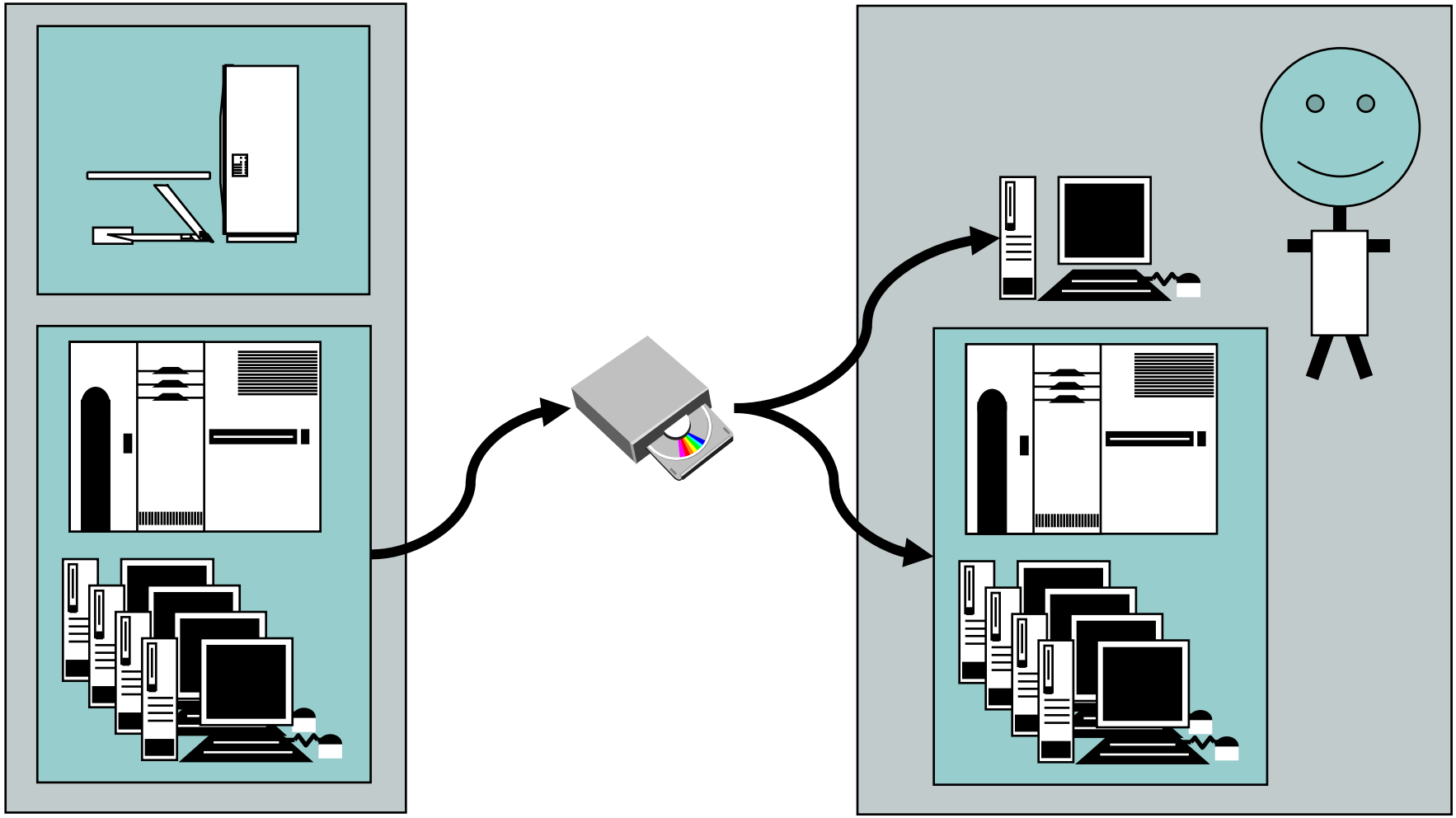


Software compatibility issues

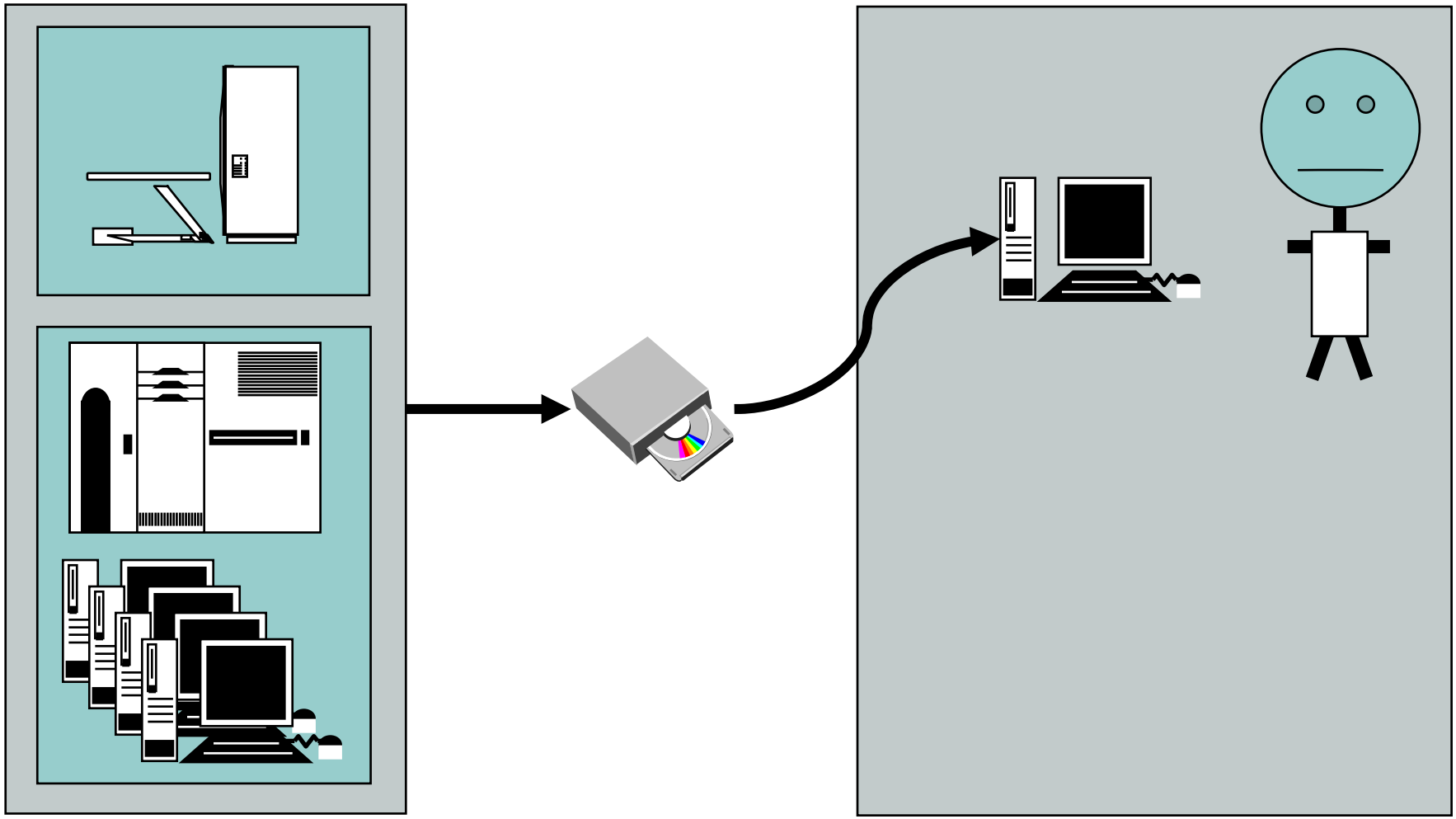
- Acquisition
  - different DICOM SOP Classes (CR, DX)
  - different type of image (raw vs. processed)
  - inconsistent grayscale contrast
    - incorrect encoding of lookup tables
  - other configuration issues
    - many CR/DX vendors make highly configurable
    - intent is to adapt to vagaries/limitations of PACS
    - unintended consequences when sent off site

Software compatibility issues

- Transfer
  - local DICOM network transfer rarely an issue
    - widely tested
    - pre-requisite for PACS to work at all
  - DICOM CD compatibility not as universal
    - some vendors default to proprietary CD formats
    - some vendors write “bad” DICOM CDs
    - use of compression may cause issues
    - improving, but a lot of old, bad equipment in field
  - DICOM supports ISO standard compression schemes
    - lossless (reversible) – JPEG, JPEG-LS, JPEG 2000
    - lossy (irreversible) – JPEG, JPEG 2000



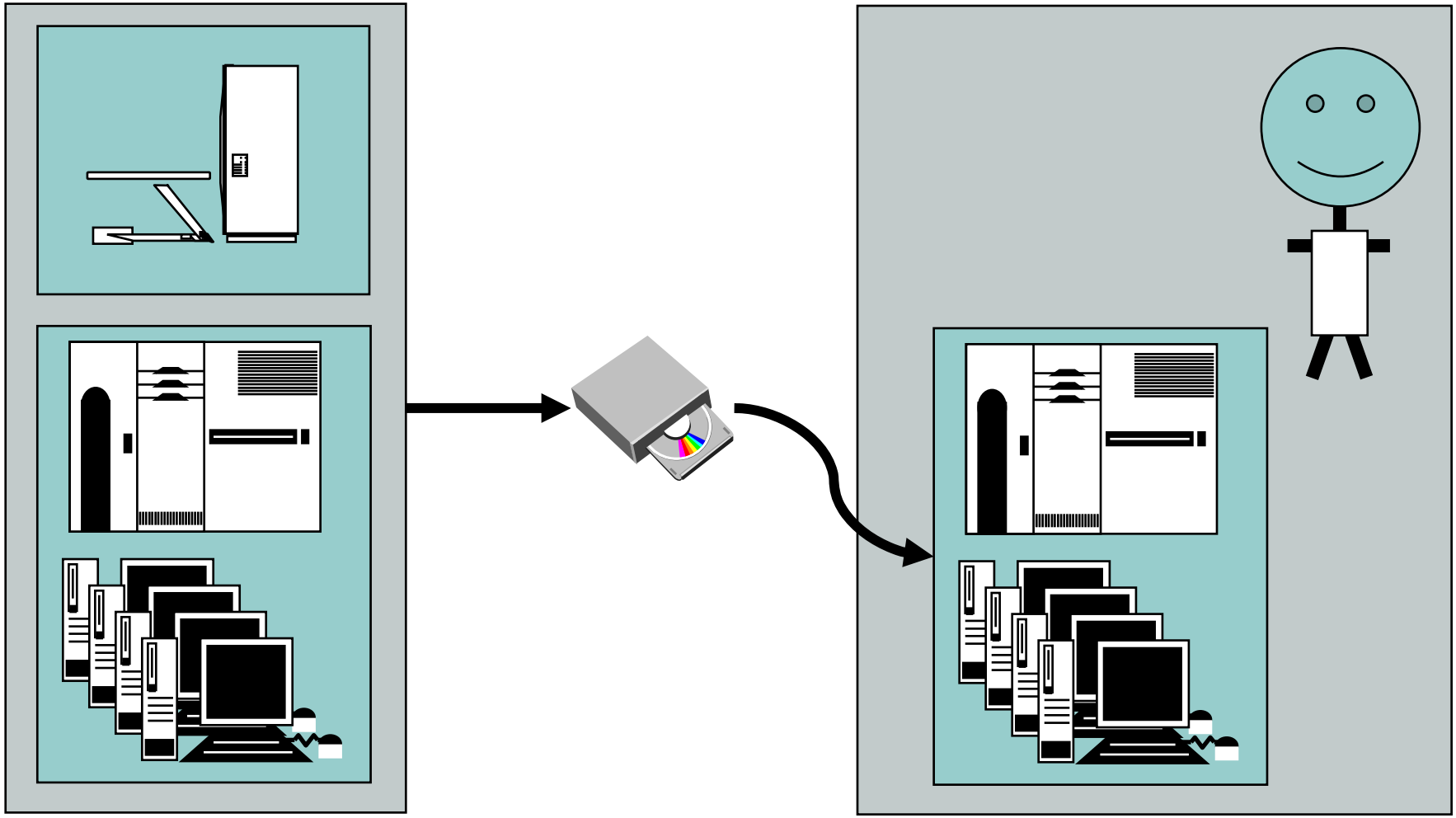
PACS -> Media -> PC Viewer or PACS Import



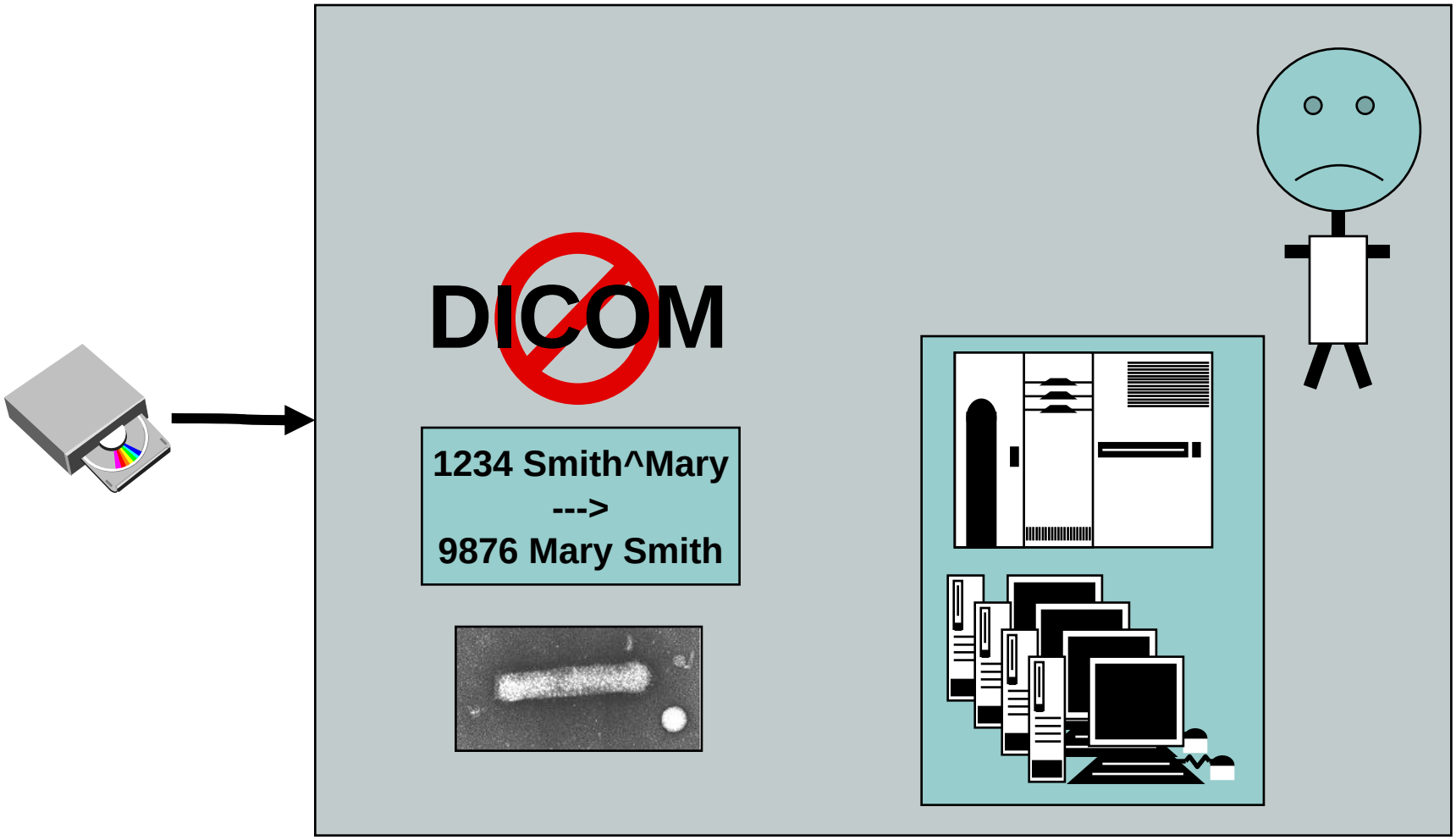
Is the Standalone PC Viewer a solution ?

CD Viewer Issues

- CDs often burned with Windows viewer on it
  - PC and operating system version issues
  - software install issues
  - speed issues if run from CD
  - display compatibility issues (especially calibration)
- Hospital IT security policy
  - risk of viruses
  - risk of interference with local applications
- Training and usability
  - need to learn to use dozens of viewers
- Designed for review not primary interpretation
  - viewers may lack full grayscale pipeline support



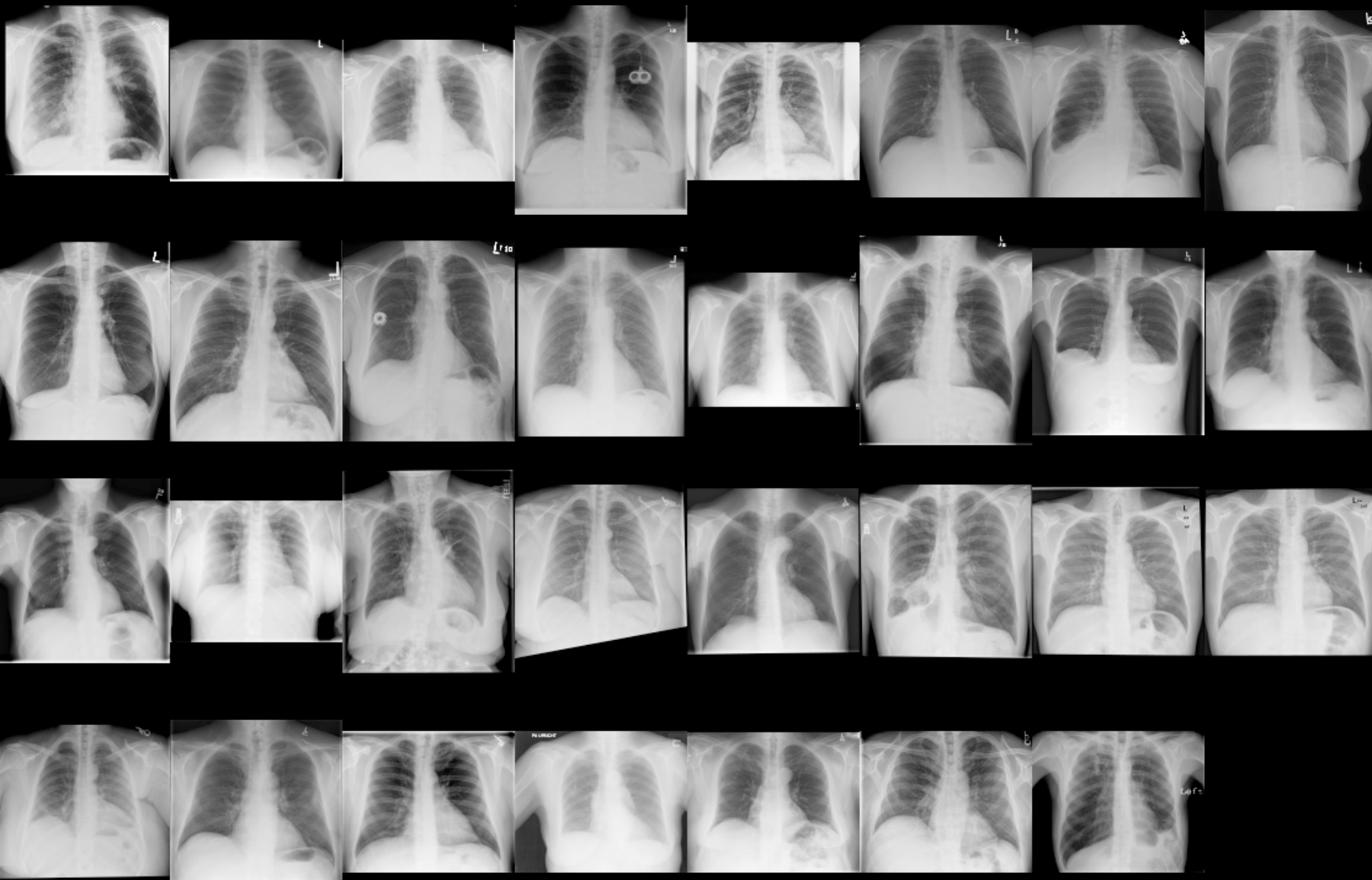
The best solution: Import standard media into the PACS

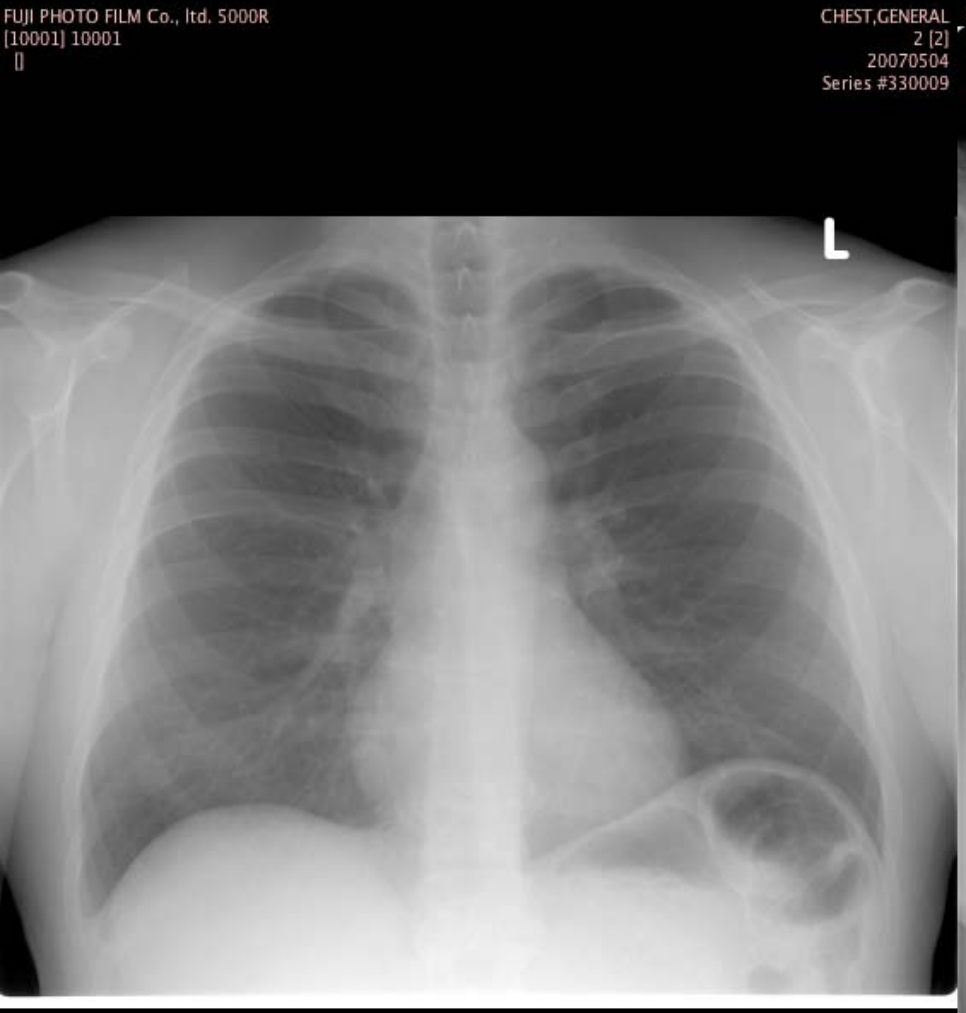


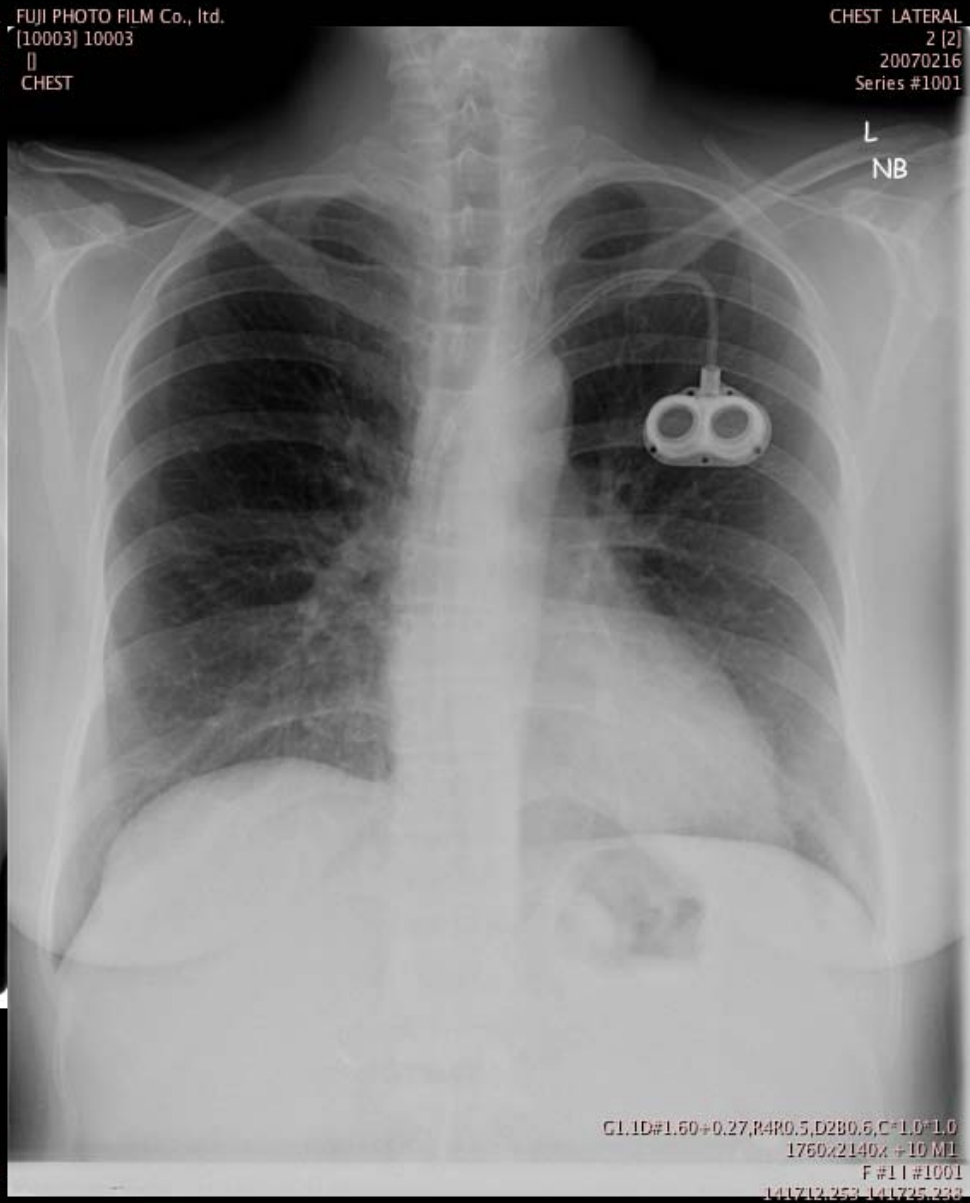
Barriers to import: format, ID reconciliation, viruses

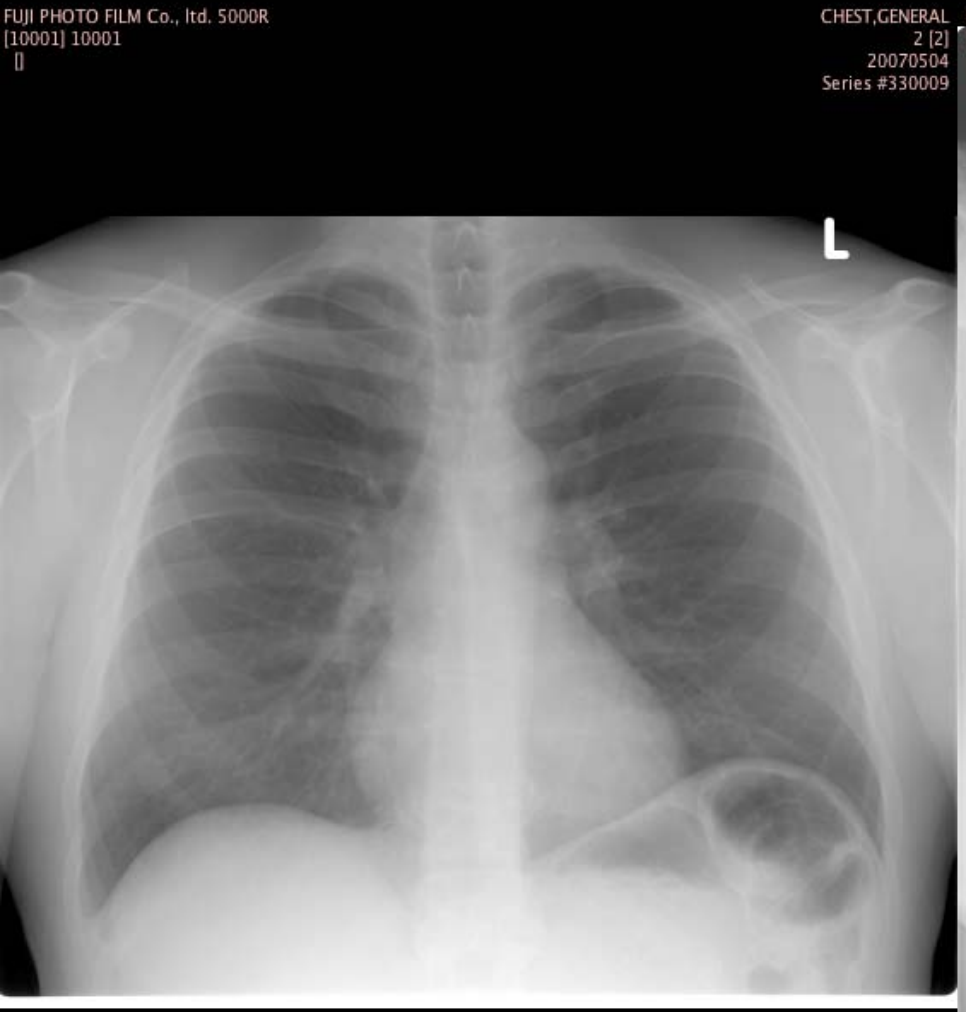
Software compatibility issues

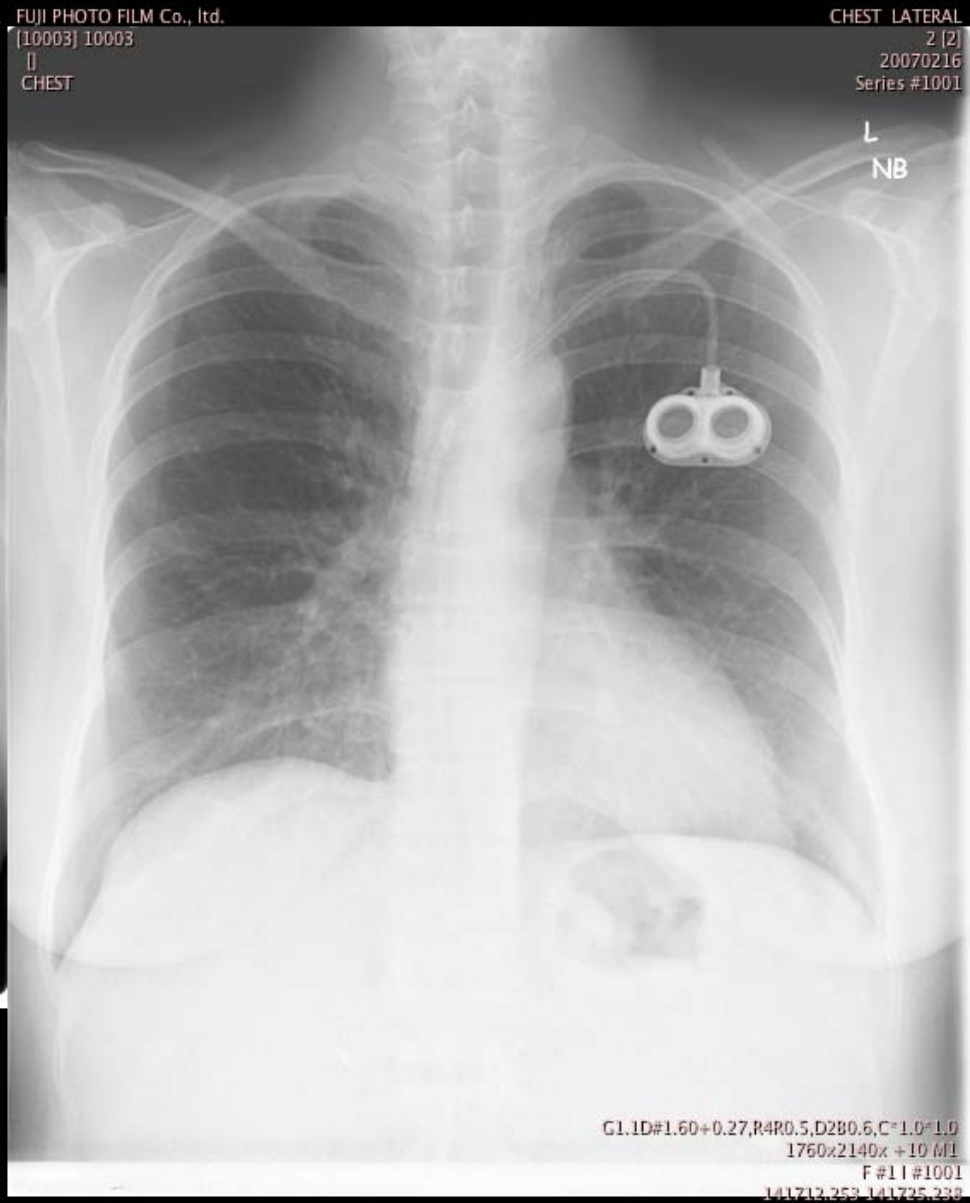
- Display
  - different DICOM SOP Classes
    - need to support both CR and DX
  - need images that are “ready to view”, not raw
  - consistent appearance (grayscale contrast)
    - correct application of “lookup tables” in image header
    - a GSDF calibrated display is necessary (but not sufficient)
  - base set of features is universal
    - zoom, pan, window
  - may or may not have additional features
    - enhancement and image processing
    - workaround known “bugs” in acquisition modalities




FUJI PHOTO FILM Co., ltd. 5000R
[10001] 10001
[]
CHEST,GENERAL
2 [2]
20070504
Series #330009
L
2140x1760x +10 M1
F #1 I #1009
124312.0000 124312.0000

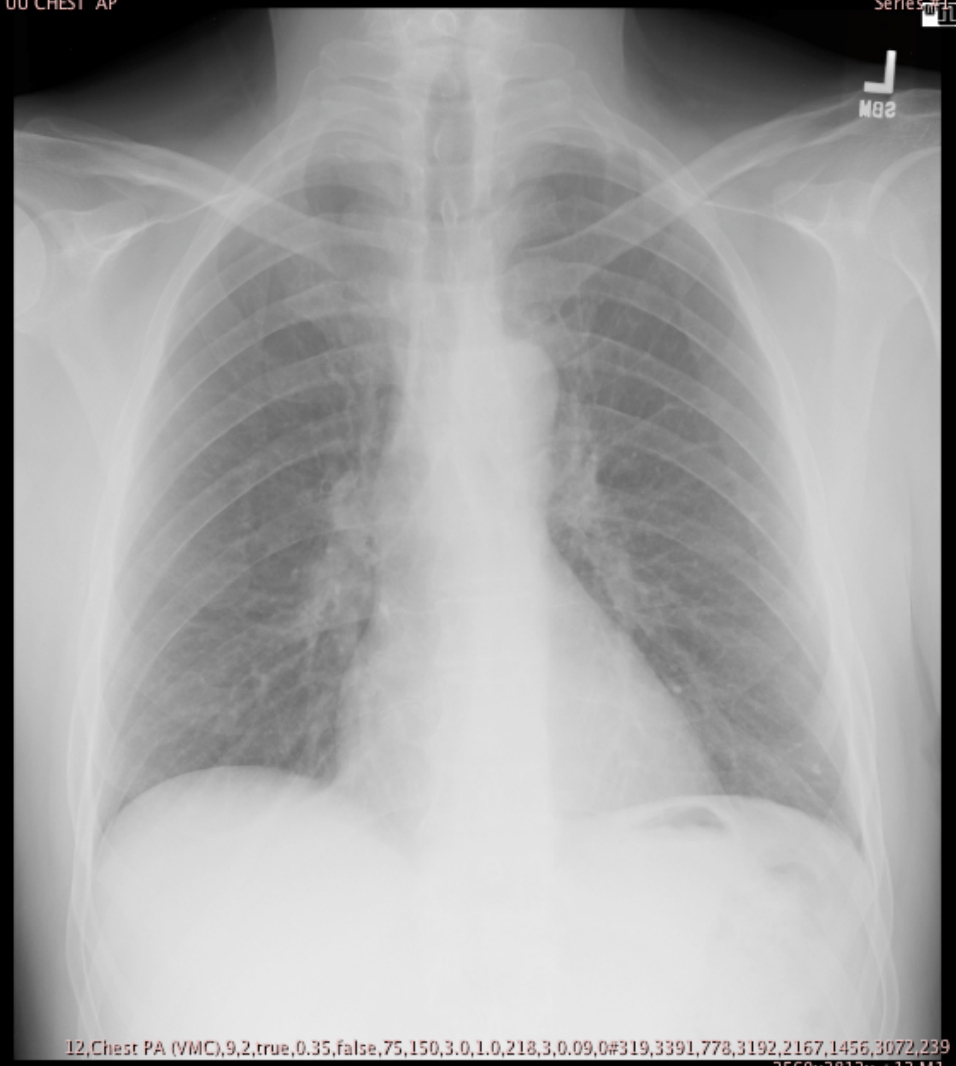

FUJI PHOTO FILM Co., ltd.
[10003] 10003
[]
CHEST
CHEST LATERAL
2 [2]
20070216
Series #1001
L
NB
G1.1D#1.60+0.27,R4R0.5,D280.6,C*1.0*1.0
1760x2140x +10 M1
F #1 I #1001
DERIVED\POST_PROCESSED


FUJI PHOTO FILM Co., ltd. 5000R
[10001] 10001
[]
CHEST,GENERAL
2 [2]
20070504
Series #330009
L
2140x1760x +10 M1
F #1 I #1009
124312.0000 124312.0000


FUJI PHOTO FILM Co., ltd.
[10003] 10003
[]
CHEST
CHEST LATERAL
2 [2]
20070216
Series #1001
L
NB
G1.1D#1.60+0.27,R4R0.5,D2B0.6,C*1.0*1.0
1760x2140x +10 M1
F #1 I #1001
141712.253 141725.238
DERIVED\POST_PROCESSED

HOLOGIC, Inc. DROC2000IC
[10011] 10011
[]
UU CHEST AP

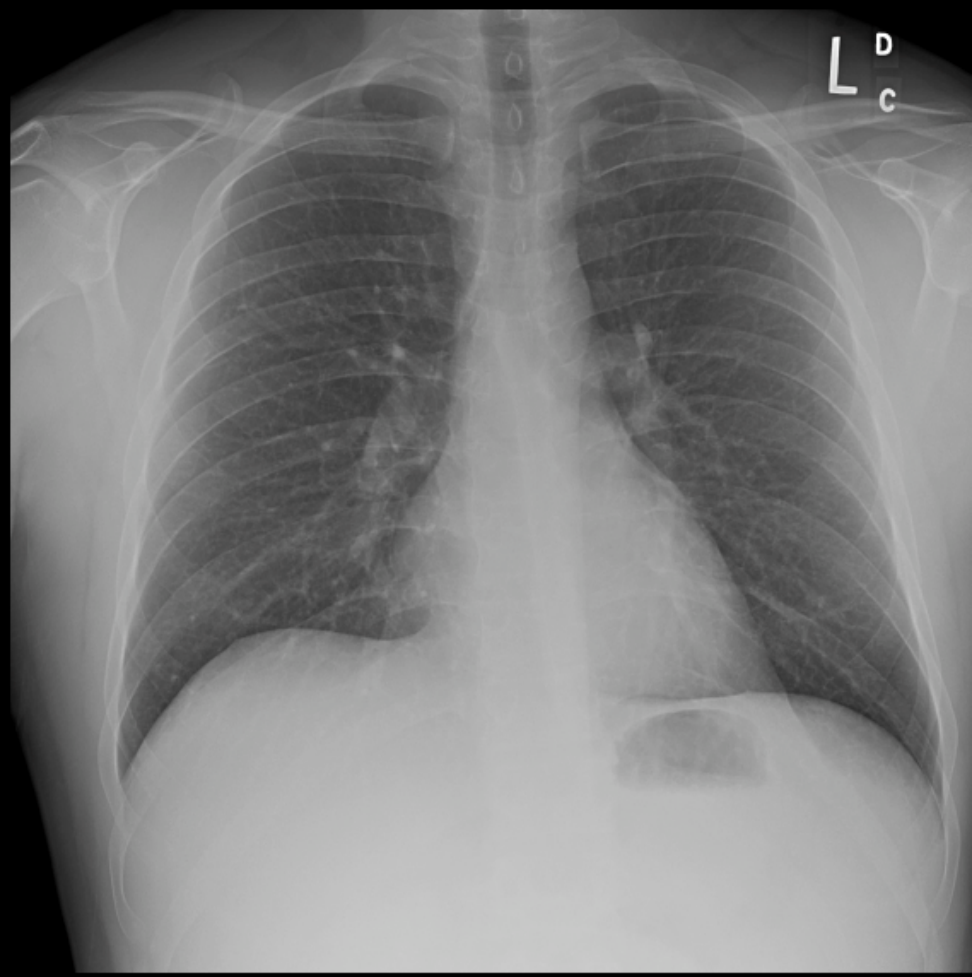
CHST2
CHEST
2 [2]
20050906
Series #1



12,Chest PA (VMC),9,2,true,0.35,false,75,150,3.0,1.0,218,3,0.09,0#319,3391,778,3192,2167,1456,3072,239
2560x2812x +12 M1
F #1 I #1
163946 163946
DERIVED\IT

GE MEDICAL SYSTEMS Revolution XQi ADS_27.5
[10005] 10005
[]
U CHEST Chest PA postero-anterior

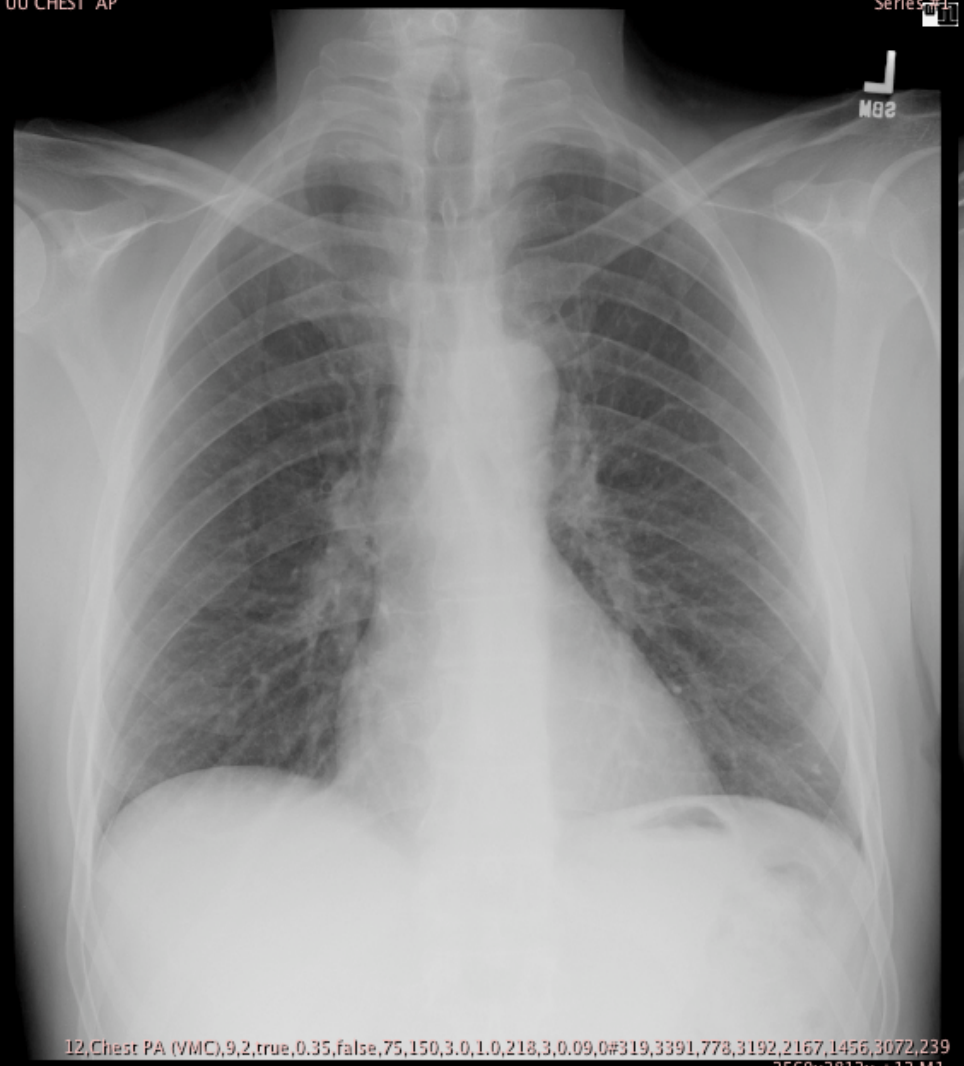
CHEST, PA & LATERAL
CHEST, PA & LATERAL
2 [2]
20040920
Series #35534



2022x2022x +14 M2
F #1 I #1
081345.000000 081401.000000

HOLOGIC, Inc. DROC2000IC
[10011] 10011
[]
UU CHEST AP

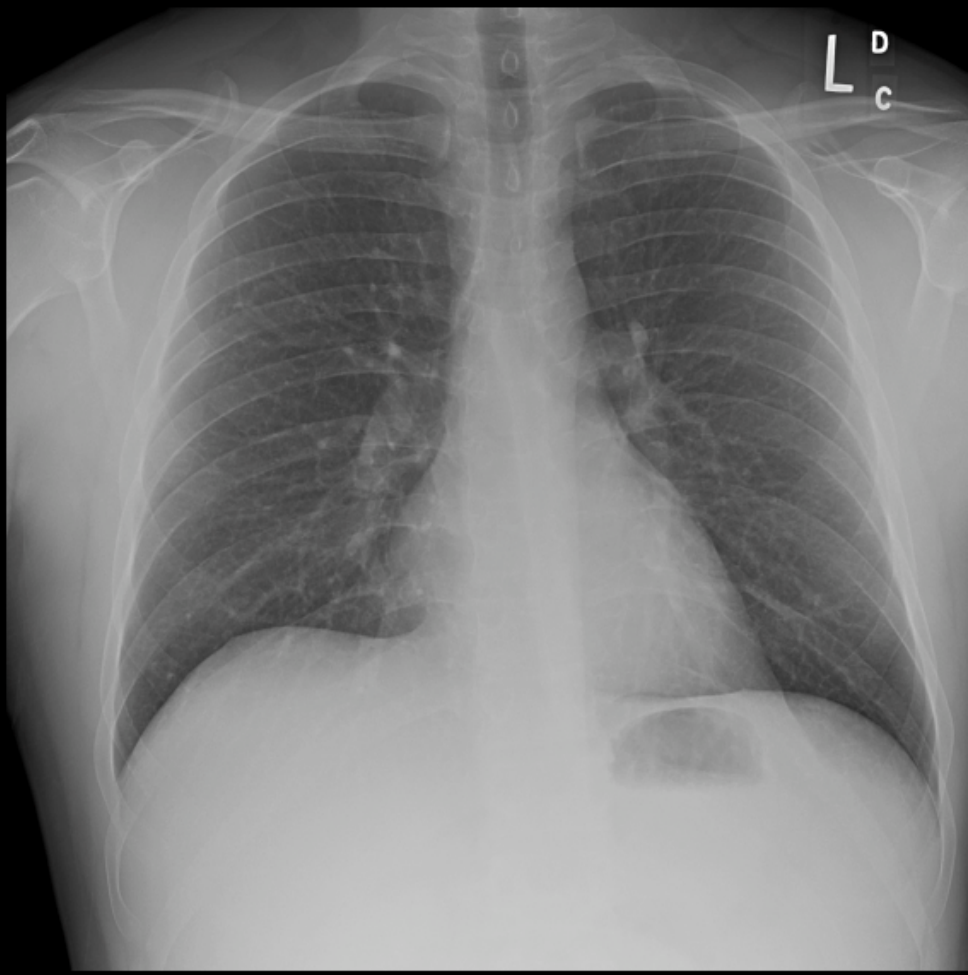
CHST2
CHEST
2 [2]
20050906
Series #1



12,Chest PA (VMC),9,2,true,0.35,false,75,150,3.0,1.0,218,3,0.09,0#319,3391,778,3192,2167,1456,3072,239
2560x2812x +12 M1
F #1 I #1
163946 163946
DERIVED\IT

GE MEDICAL SYSTEMS Revolution XQi ADS_27.5
[10005] 10005
[]
U CHEST Chest PA postero-anterior

CHEST, PA & LATERAL
CHEST, PA & LATERAL
2 [2]
20040920
Series #35534

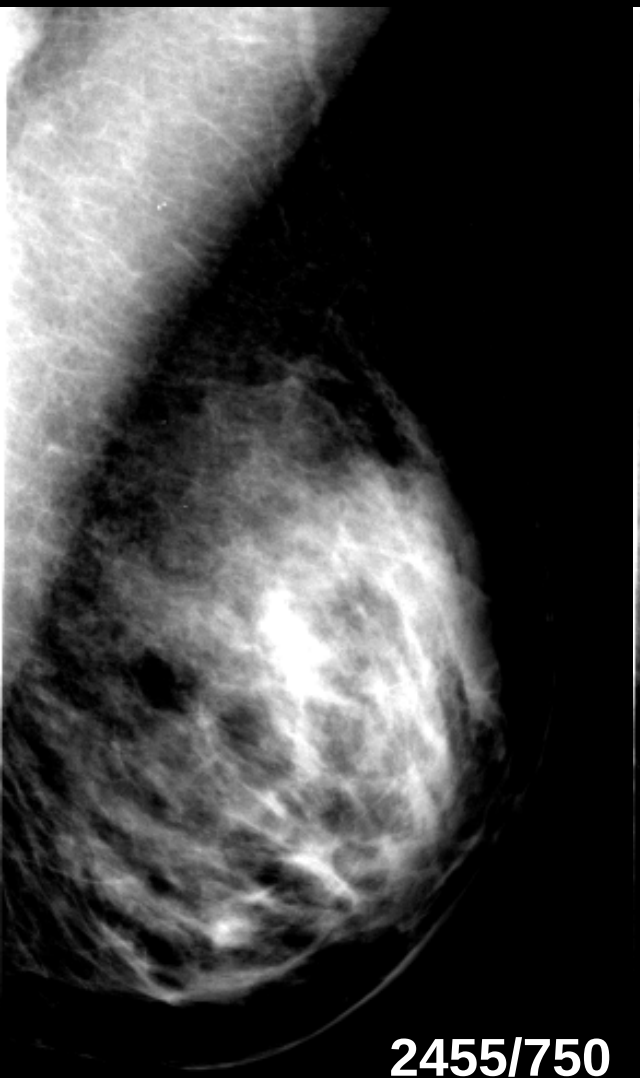


2022x2022x +14 M2
F #1 I #1
081345.000000 081401.000000

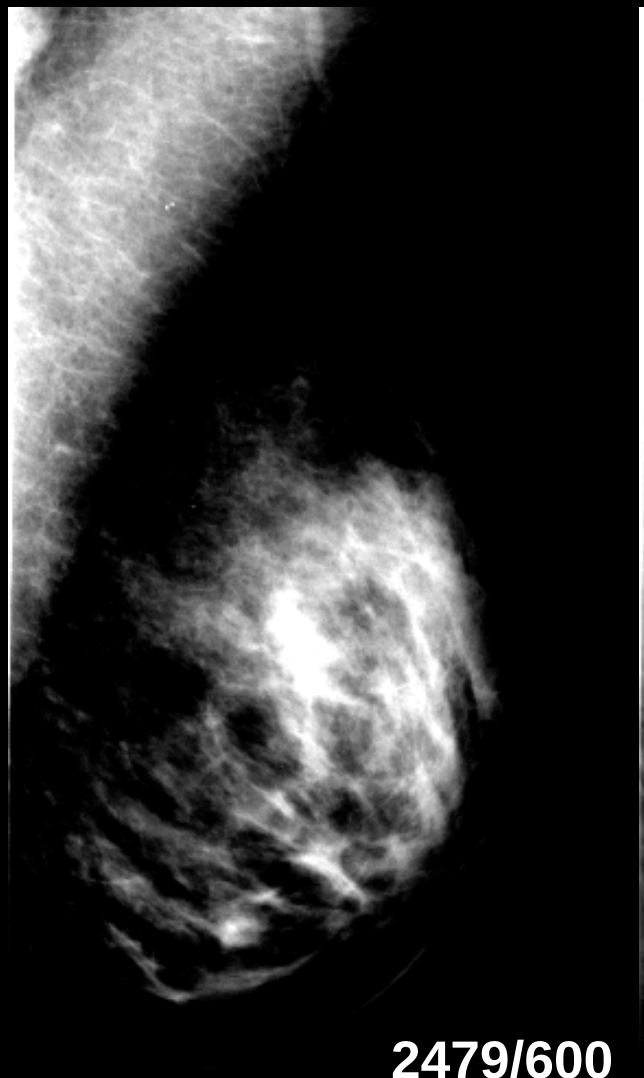
Image contrast adjustment

- A single default presentation of image contrast
  - usually not sufficient
  - adjust the image for light and dark areas
- Traditionally
  - linear window center and width
- Non-linear contrast adjustment
  - lookup table (LUT)
  - function
- DICOM supports all three mechanisms
  - image (or presentation states) may contain multiple

Linear window - bright and dark areas clipped

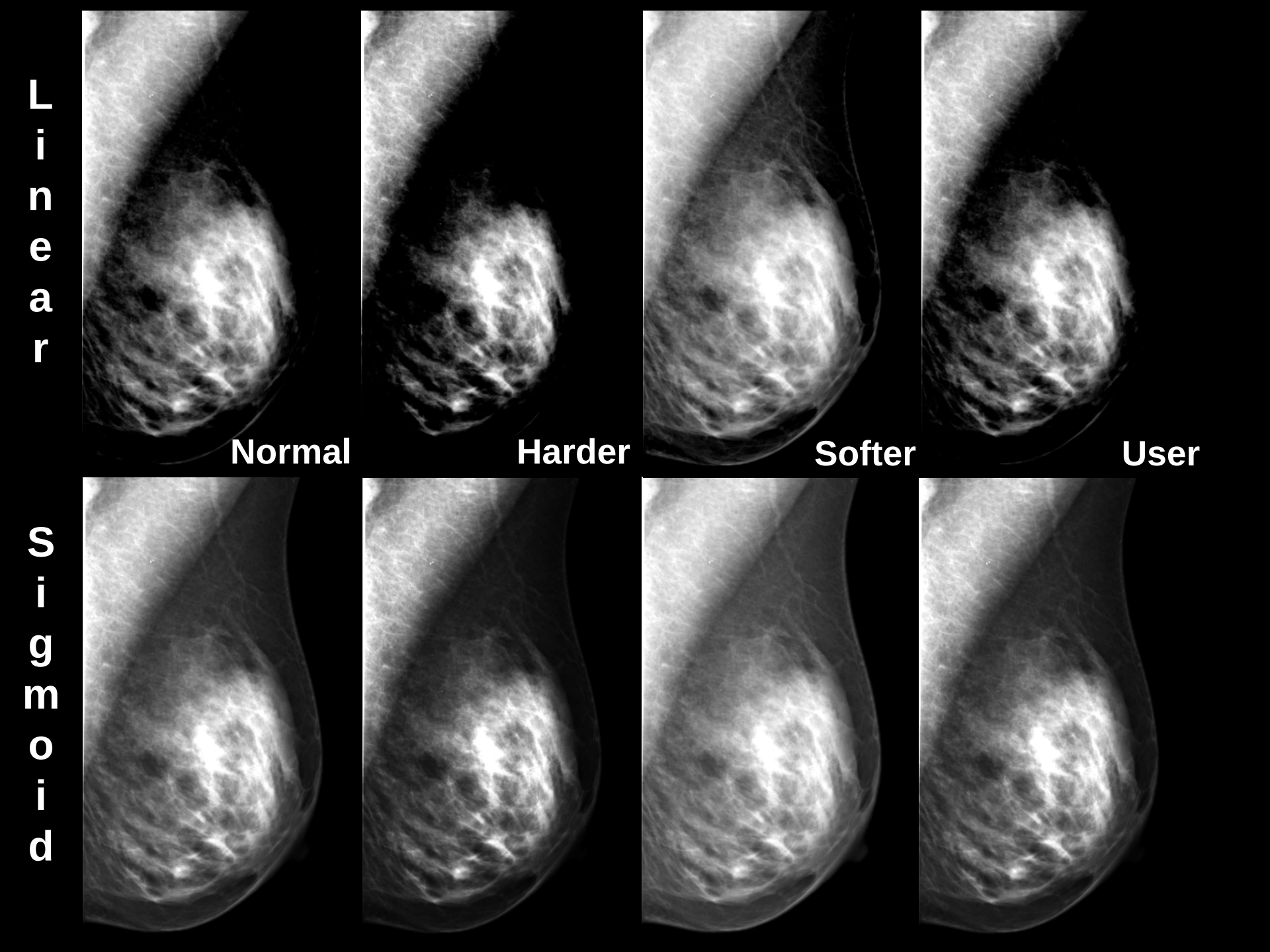


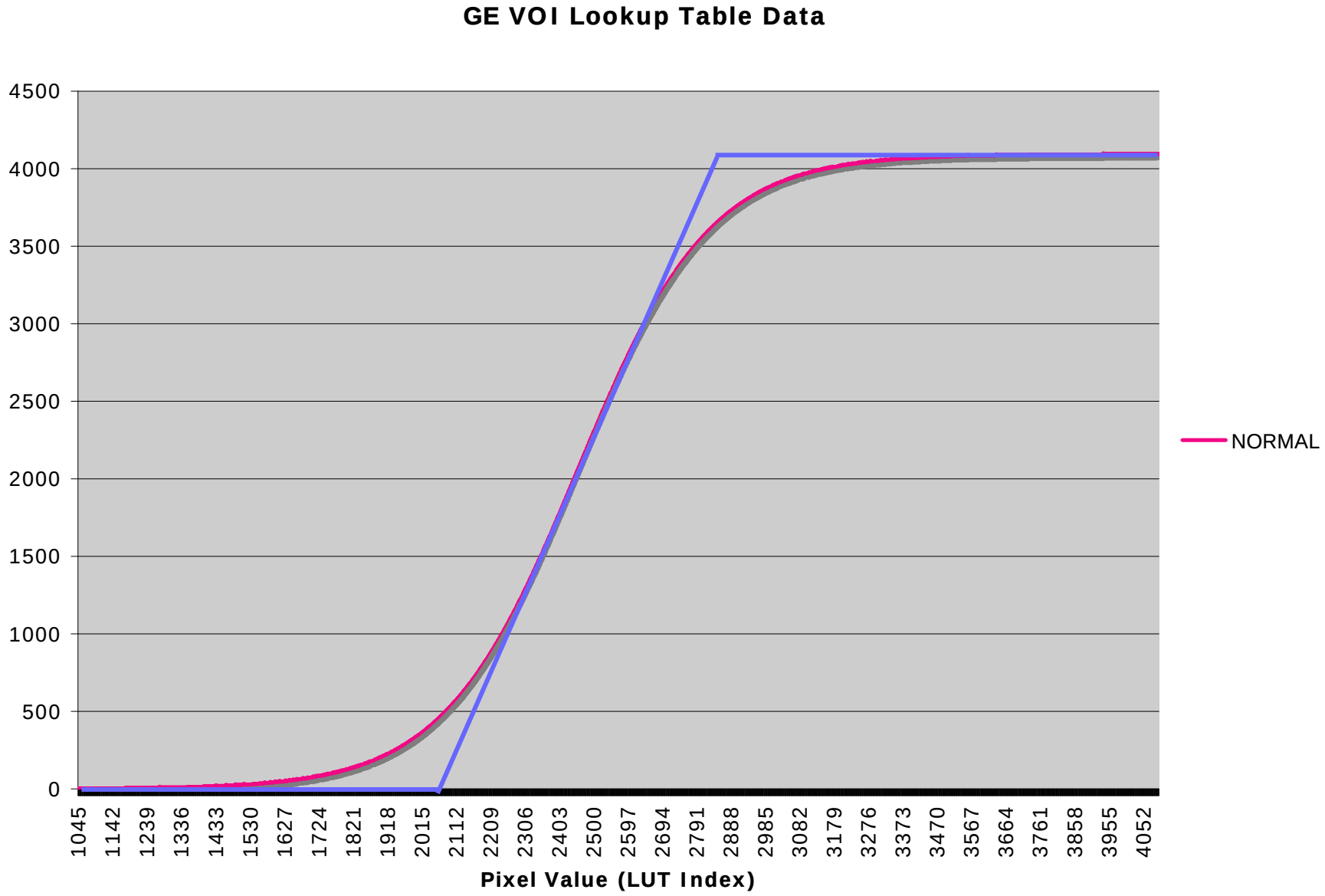
2455/750



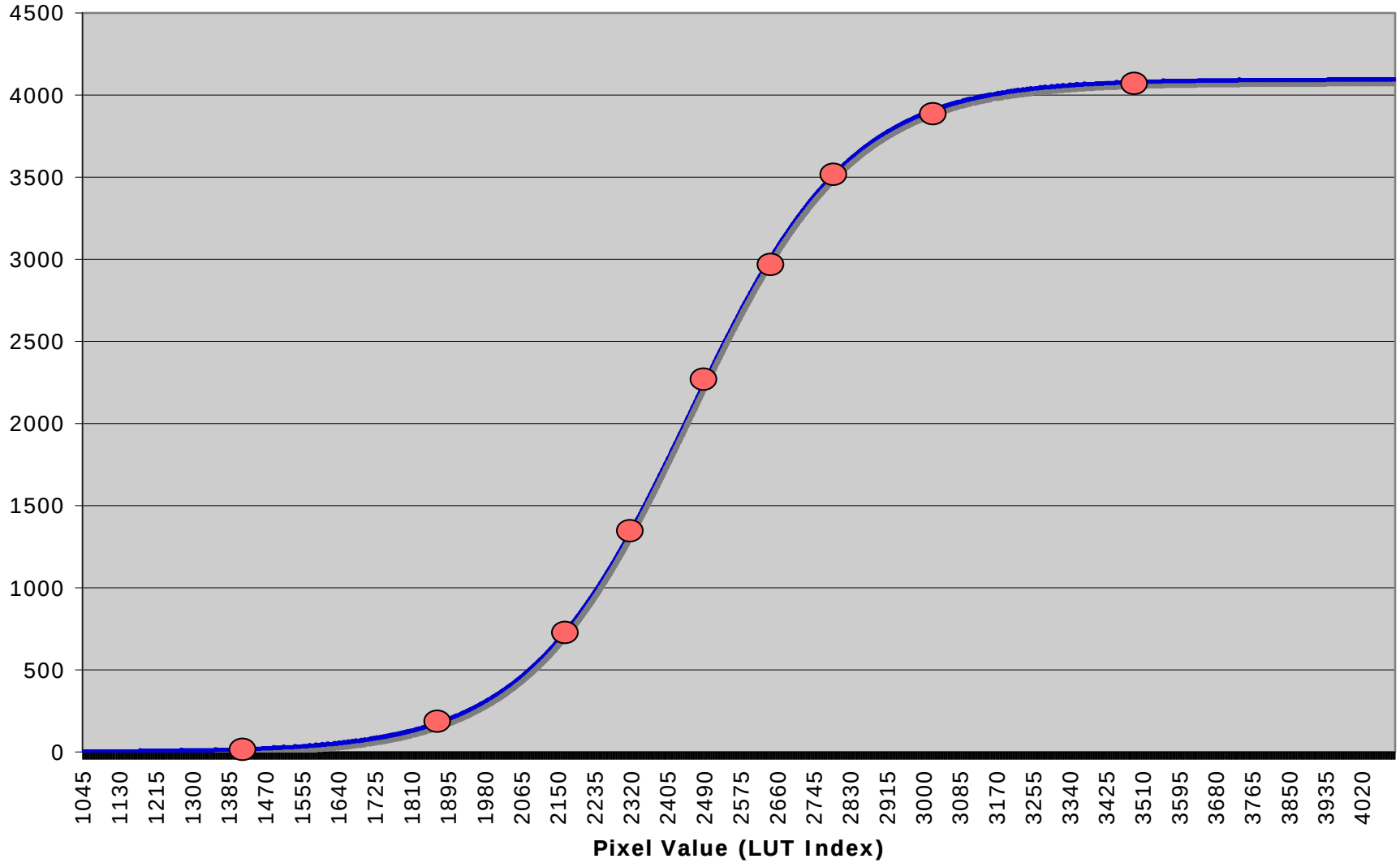
2479/600

2413/900


Linear
Sigmoid
Normal
Harder
Softer
User



Sigmoid curve encoded as VOI Lookup Table Data



Sigmoid LUT function

$$OUT = \frac{Output\_range}{1 + \exp\left(-4\frac{IN - WC}{WW}\right)}$$

Reference images

- ILO reference set
  - required to be available/used by regulations
- Digitally displayable with patient images
  - separate light box impractical
  - degrades workflow and perception (too bright)
- Need a digital version of reference sets
  - with comparable contrast and processing to digital acquisitions
  - digitizing (previously copied) reference films -> quality degradation
  - freely and publicly available on Internet
- Need to be DICOM
  - storable in PACS
  - viewable in workstations
  - DX images for consistent grayscale contrast (P-Values)

Philips Medical Systems digital DIAGNOST
[123-45-6789] Smith^John
[]
CHEST PA

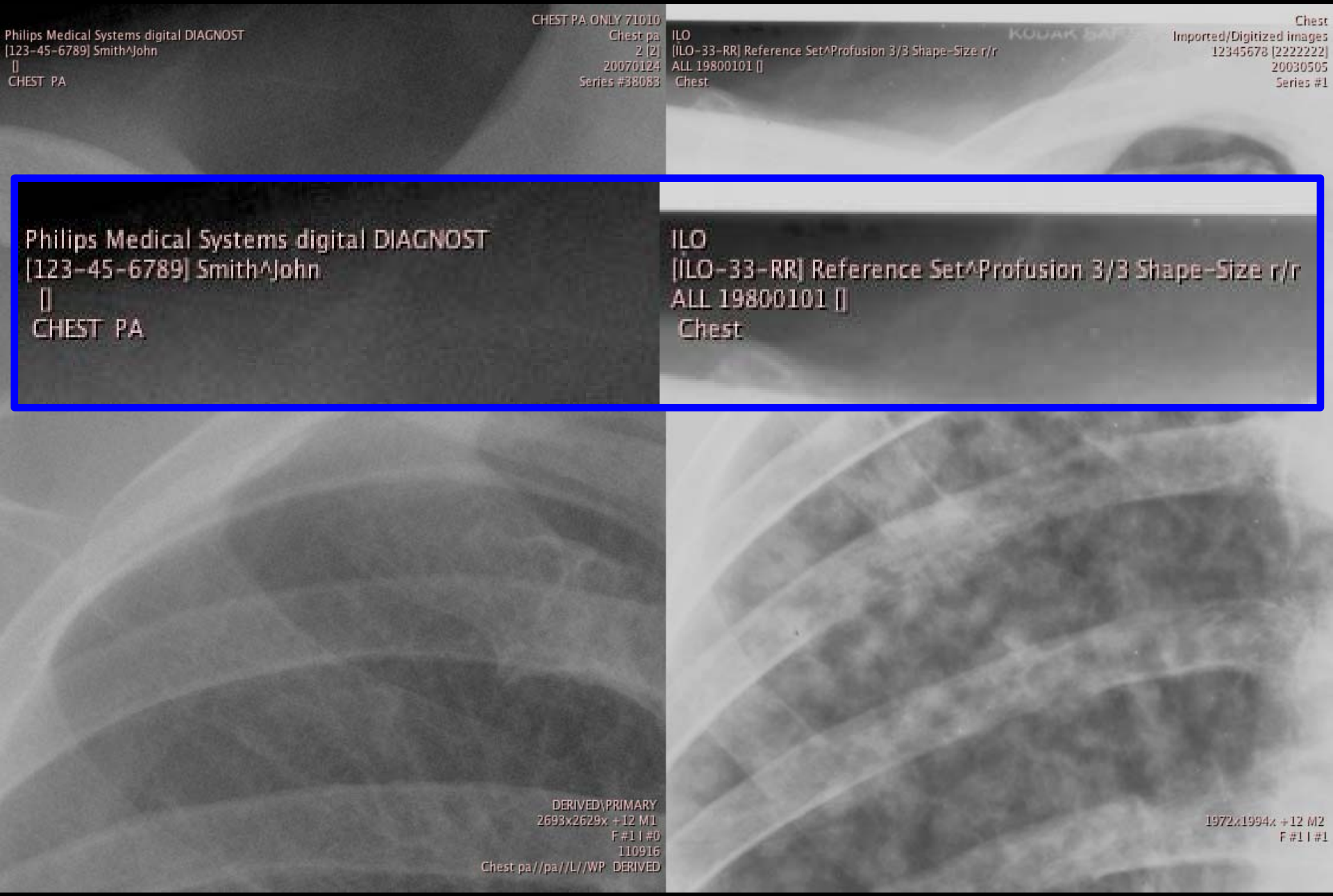
CHEST PA ONLY 71010
Chest pa
2 [2]
20070124
Series #38083

ILO
[ILO-33-RR] Reference Set^Profusion 3/3 Shape-Size r/r
ALL 19800101 []
Chest

Chest
Imported/Digitized images
12345678 [2222222]
20030505
Series #1

DERIVED\PRIMARY
2693x2629x +12 M1
F #1 I #0
110916
Chest pa//pa//L//WP DERIVED

1972x1994x +12 M2
F #1 I #1


Philips Medical Systems digital DIAGNOST
[123-45-6789] Smith^John
[]
CHEST PA
CHEST PA ONLY 71010
Chest pa
2 [2]
20070124
Series #38083
ILO
[ILO-33-RR] Reference Set^Profusion 3/3 Shape-Size r/r
ALL 19800101 []
Chest
Chest
Imported/Digitized images
12345678 [2222222]
20030505
Series #1
DERIVED\PRIMARY
2693x2629x +12 M1
F #1 I #0
110916
Chest pa//pa//L//WP DERIVED
1972x1994x +12 M2
F #1 I #1

Displaying reference images

- Make available as either
  - reference set “built in” to display software
  - “pseudo” patient
- Existing PACS workstations
  - no explicit mechanism to display “reference” images
  - many won’t show multiple “patients” simultaneously
- Dedicated workstations
  - can DICOM query PACS for patients
  - has local reference set available
  - off-the-shelf (OTS) 3rd party DICOM generic workstation
  - custom workstation specific to B reading task


L
pa
AP
BARCO


3/3
BARCO

Number of Displays

- Traditional PACS workstations
  - two portrait 3 megapixel
  - side-by-side current/prior or PA/lateral
  - sometimes four – ergonomically difficult
- ILO viewing standard
  - two films required – subject + reference
  - three recommended – between references
- Can simulate with two digital displays
  - rapidly toggle from one reference to another
  - order reference images by increasing profusion and size

Using existing infrastructure

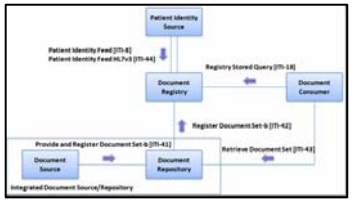
- B reading “outside” patients in one’s hospital office
- Importing patients for “consultation”
  - some PACS have explicit support for importation and reconciliation of foreign identifiers – IHE IRWF profile
  - may be able to view from CD inserted locally
- When not permitted by PACS/IT
  - whether by policy or technology
  - install another, separate, computer
  - share (expensive) high resolution grayscale displays via KVM (keyboard-video-monitor) switches – support up to 3MP displays

Remote reading approach

- Images are provided on central server
  - local access is via Internet connection
  - software is remotely accessed and managed and maintained
  - local machine provides Internet access and high quality monitors
- Implementation varies
  - browser applet or plug-in
  - installable local client (ActiveX or Java Web Start)
- Performance and satisfactory user experience
  - largely a function of speed of the connection
  - patient & reference images pre-loaded (work list look ahead)
  - lossy compression unlikely to be acceptable
- NIOSH could provide
  - central archive server & same client for all readers

Cross-enterprise sharing

- Problem of sharing patient related images and documents between loosely coupled enterprises is not new
- Integrating the Healthcare Enterprise (IHE)
  - initially focused on using existing standards (DICOM and HL7) and defining workflow within an enterprise
  - now expanding to cross-enterprise document sharing (XDS), including images (XDS-I)
  - is starting to address cross-enterprise user authentication (XUA) and patient identifier cross-referencing (PIX)



- NCI Cancer Biomedical Informatics Grid (caBIG)
  - secure access to shared resources & services
- In future, remote access by authorized B readers to images in acquisition site may be feasible and routine

Security and privacy

- Films, digital images, reports, and forms contain
  - individually identifiable health information (IIHI)
  - which is “protected” (PHI) under HIPAA Privacy Rule
- Unauthorized access to IIHI or PHI is a bad thing
- Solution: protect it or remove it
- Digital data is at risk when in physical form
  - e.g., CD, just like film and paper
- On-line digital information is also at risk
  - locally – access by unauthorized staff
  - remotely – access by other individuals when in transit

Protection of PHI

- Protection in transit
  - encryption on network
  - e.g. SSL on Internet, just like electronic commerce
- Authentication
  - login with username/password
- Access control
  - access rights constrained based on identity
- Audit trail
  - record of who saw/did what/when/where

Protecting PHI in digital images

- DICOM network
  - DICOM services built on top of existing security mechanisms
  - virtual private network (VPN) or SSL for privacy
  - user identity can be conveyed
  - PACS can constrain access and maintain audit trail
- Web-based access to DICOM images
  - normal browser security mechanisms
  - can also use VPN and SSL support
  - web server handles authentication, access, audit trail

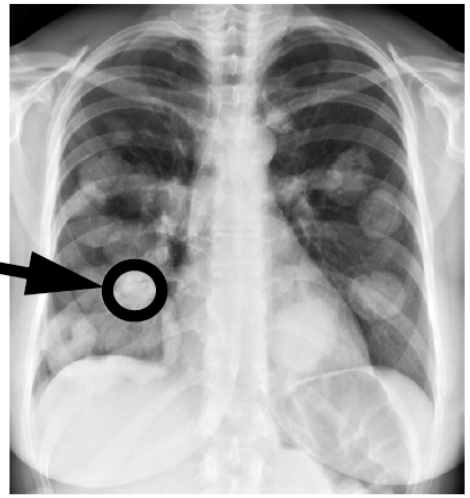
Removing PHI in digital images

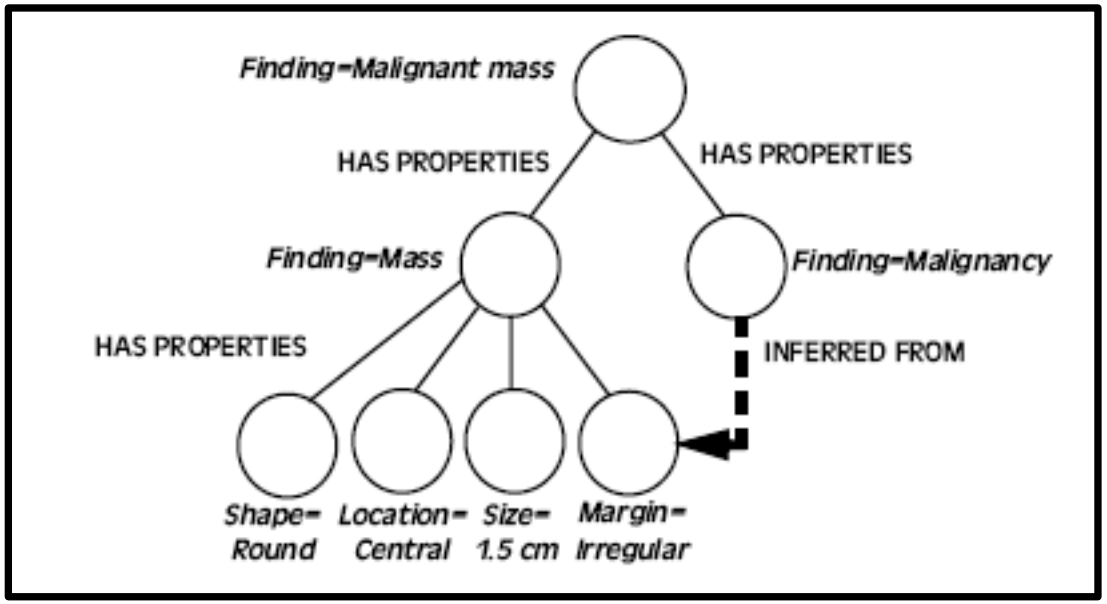
- If it isn't there, then there is less need to protect it
  - B readers do not need the patient's true identity
  - e.g., replace patient's name/SSN with pseudonymous number
  - maintain (secure) association of pseudonym and true identity
- This is the same process that is used in clinical trials
  - independent readers are "blinded" to subject's identity to reduce bias and protect privacy
  - facilitates secondary re-use – e.g., open archive for research, CAD
- The identity is in the digital image "file headers"
  - not the image pixels
  - relatively "easy" to automate such "de-identification"
  - digital x-rays should NOT have identity burned into pixels (but can be checked and blacked out if necessary)

Integrating results with images

- DICOM Structured Reporting (SR)
  - encoding of structured information
  - about images (+ other things, like waveforms)
  - codes, text, measurements, coordinates
  - references to images, locations, outlines
  - hierarchical organization
  - structure defined by “templates” for specific applications

Chest X-ray Report:
Observer: Clunie^David^A^Dr.
History: malignant melanoma excised 1Y
Findings:
- finding: multiple masses in both lung fields
- best illustration of findings:
Conclusions:
- conclusion: cannon-ball metastases
- conclusion: recurrent maligant melanoma
Diagnosis Codes:
- diagnosis: 172.9/ICD9
- diagnosis: 197.0/ICD9





DICOM Structured Reports

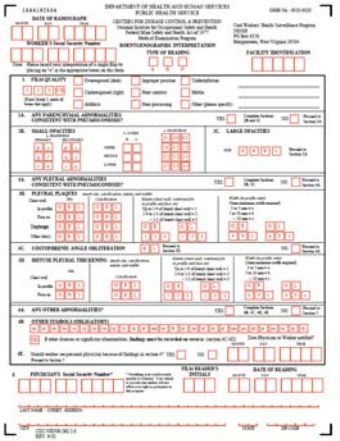
- Examples of templates in DICOM standard
  - Basic radiology report
  - Mammography and Chest CAD results
  - Echocardiography measurements
  - Obstetric measurements
  - Cardiovascular CT measurements
  - Radiation dose reports
  - …

Advantages of DICOM SR

- Same header structure as images
- Same patient/study/series model as images
- Widely used by modalities to encode measurements
- Easy to store in and retrieve from PACS
- Easy to convert into other forms
- Can be converted into HL7 CDA XML
- Can extract and render as plain text or PDF
- Can search contents and extract structure and codes
- A DICOM “form” that can point to images & locations

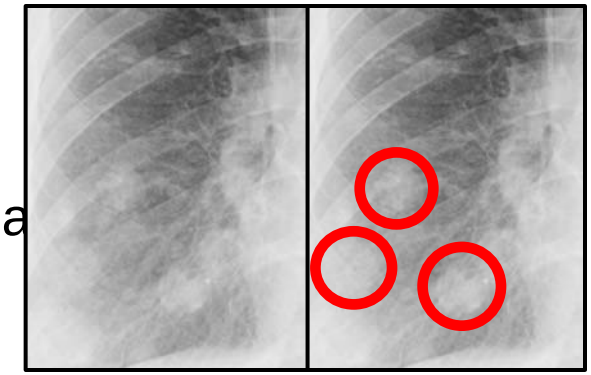
DICOM SR for classification

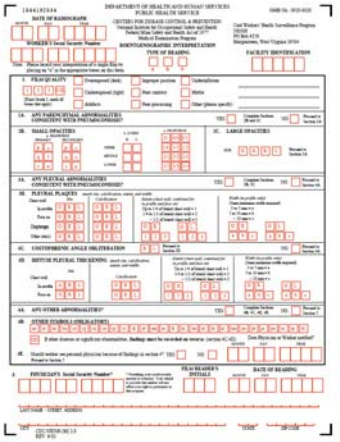
- Could easily encode both NIOSH
  - “Roentgenographic Interpretation”
  - “Miner Identification Document”
- DICOM could define in PS 3.16 of the standard
  - a template to match existing NIOSH form
  - standard codes for each concept (field)
  - a reference to UID of image being read
  - references to prior images for comparison
- Additionally, could save image “annotations”
  - pointers to locations of abnormalities in image
- SR can be digitally signed (as can images referenced)
  - standard cryptographic public key based mechanism



DICOM SR for classification

- Could easily encode both NIOSH
  - “Roentgenographic Interpretation”
  - “Miner Identification Document”
- DICOM could define in PS 3.16 of the standa
  - a template to match existing NIOSH form
  - standard codes for each concept (field)
  - a reference to UID of image being read
  - references to prior images for comparison
- Additionally, could save image “annotations”
  - pointers to locations of abnormalities in image
- SR can be digitally signed (as can images referenced)
  - standard cryptographic public key based mechanism





Conclusion

- An entire infrastructure already exists to support clinical use of digital projection X-Rays
- It is based on use of DICOM standard between modalities, PACS and workstations, using networks and CDs
- Most sites are now experienced with exporting and providing outside access to digital images (including “for presentation” digital X-Ray)
- Correct choice (or construction) of an appropriate image viewer should allow consistent display and reliable review of images, side-by-side with ILO or equivalent reference images
- Expensive displays already installed can easily be re-used
- Results can be stored as DICOM Structured Reports and DICOM will help with adding templates and codes as requested
- Matters of security and privacy can and should be addressed through conventional means that are already widely used clinically

Recommendations

- Both CR and DX DICOM images should be permitted
  - due to large installed base and some vendors who will not convert to DX
- Processed (“for presentation”) images should be required
  - must not be dependent on proprietary processing in display workstation
- Display workstations should be qualified and certified for B reading
  - must work with test images from different vendors and software
  - must support all variations of encoding and grayscale pipeline
  - must be able to display reference images side-by-side
- Images should be de-identified before sending for reading
  - must minimize leakage of IIHI/PHI
- Digital (not digitized film) reference set should be created & released
  - comparable in contrast and resolution to CR and DX images
- Explore creation of a managed distributed or centralized infra-structure
  - remote reading from a central PACS, or an XDS or caBIG network
  - open archive for research