NAACCR Cancer Informatics Symposium

Moving Towards the Vision: Integrating the Cancer Registry into Clinical Research

Joyce C. Niland, Ph.D.
Chair, Division of Information Sciences
City of Hope National Medical Center

*The Registry Profession of the Future: No More Dinosaurs**

"Given the changes we are experiencing in medicine, the threats to the viability of the registry profession, and the potential opportunities for the future, the characteristics of successful cancer registrars most likely will include:

- **flexibility, adaptability, ability to be proactive, assertiveness,**
- **willingness to learn new things**
- **willingness to collaborate and work in partnerships**
- **and an entrepreneurial outlook."**

***D.S. Miller, MD, MPH, *Journal of Registry Management*, 26:2, 1999**

Outline:

- City of Hope National Medical Center
  - IAIMS at COH
  - Cancer Registry
  - Clinical Trials Data
  - Outcomes Research
- Integrating the Registry into Clinical Research
  - Analysis of Data Flow
  - Re-engineering of Workflow
- Future Plans

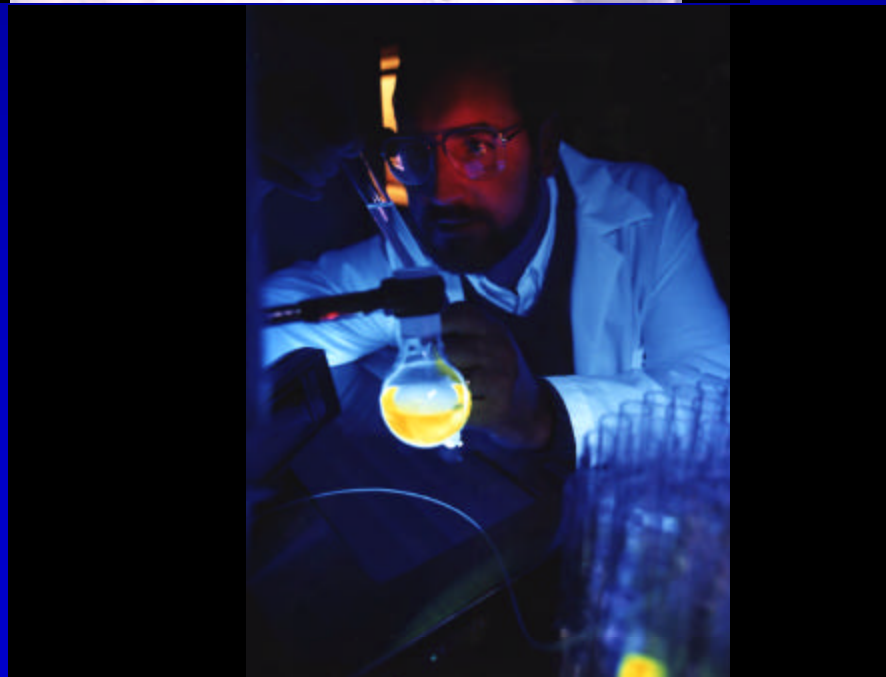
City of Hope National Medical Center

- Began as 2 tents in the desert as a TB sanatorium in 1913
- Established as a hospital, moved to cancer emphasis ~ 1950s
- Now designated by NCI as a Comprehensive Cancer Center

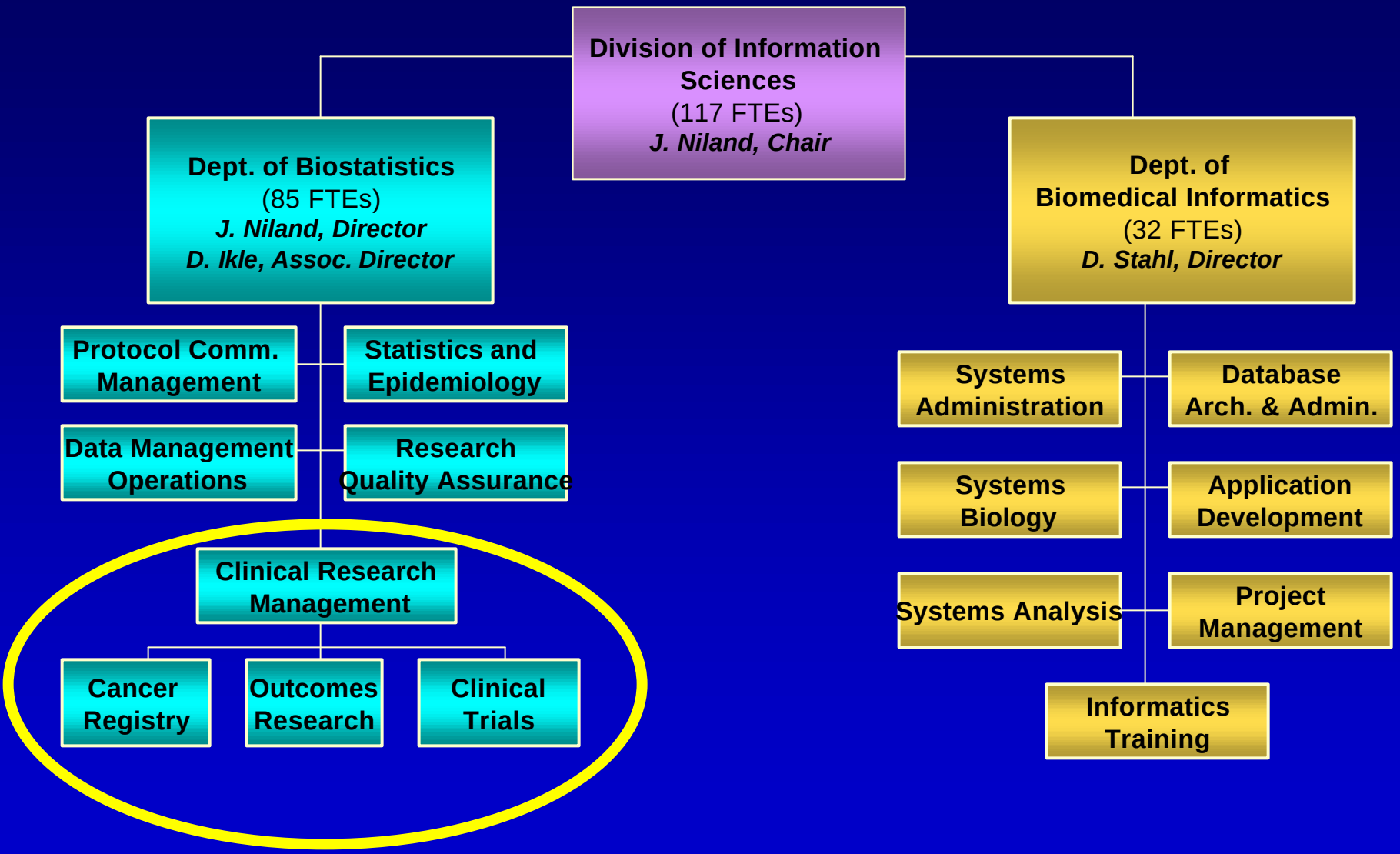
City of Hope National Medical Center

- State-of-the-art care to patients with cancer & other life-threatening diseases
  - 200 beds, 110 MDs
  - 2700 cancer patients/year

- First of 5 endowed Beckman Research Institutes in the nation
  - Leading edge research into causes, prevention, and cure of such diseases
  - 154 research scientists
  - $45 million in grants



Division of Information Sciences Organizational Chart



3 Operational Systems for Cancer Patient Data at COH

- **Cancer Registry: *CNext***
- **Clinical Trials: *BITS***
- **Outcomes Research: *NCCN***

Ultimate Goals for Integrated Information Systems

- Capture highest quality data one time at the source of origin
- Share data electronically as appropriate and as needed
- Create information out of data in near real time, with quality control “up front”
- Apply tools to create knowledge out of the information

Challenges in Data Integration

- Operational systems were created to meet different functional needs of an institution
- Clinical research systems are developed independently over time ("silo" approach)
  - Systems, rules, processes, and vocabularies for data collection vary
  - Standards for patient data collection are needed
- Semantic/conceptual barriers pose substantial challenges to database interconnectivity
  - Leads to the requirement for "metadata"

CNext:

State Developed Cancer Registry Software System

PATIENT IDENTIFICATION

DIAGNOSIS C770 LYMPH NODES R SUPRACLAV
1960 LATERALITY: NOT PAIRED
96903/ 9 FOLLICULAR LYMPHOMA

HOSPITAL: CITY OF HOSPE
FIRST SEEN: 2001
ACCESSION NO.: 200101238/00
DOC NO.: 05/07/2002
REG NO.:
VER: CN1.2.2

DIAGNOSIS DATE: 06/13/1998
1st ADMISSION DATE: 11/21/2001
DATE INPAT ADMIT: 11/25/2001
DATE INPAT DSCH: 01/22/2002

LAST: [redacted]
FIRST: [redacted]
M: [redacted]

MEDICAL RECORD NO.: 1050558
NAME SUFFIX:
TYPE REPORTING SOURCE: HOSPITAL/CLINIC
PHONE: [redacted]

ADDRESS AT DIAGNOSIS: [redacted]
COUNTY OF RESIDENCE: [redacted]

MARITAL STATUS: MARRIED
SPANISH: NON-SPANISH
SEX: FEMALE
AGE: 055
DATE OF BIRTH: [redacted]

RACE 1: WHITE 2: NOT CODED 3: NOT CODED 4: NOT CODED 5: NOT CODED

PLACE OF BIRTH: UNKNOWN
SOCIAL SECURITY NO.: 350344911
MAIDEN:
ALIAS:
LAST:

PLACE OF DIAGNOSIS: Kaiser San Diego

HOSPITAL REFERRED FROM: KAISER SUNSET
HOSPITAL REFERRED TO: UNREFERRED PATIENT
CLASS OF CASE: 3 RX ELSEWHERE

PHYS.

ATTENDING: SPIELBERGER, R MD O0016022
REFERRING: ----------
SURGEON: ----------
FOLLOWING: ----------
MED ONC: SPIELBERGER, R MD O0016022
RAD ONC: RAUBITSCHEK, A MD A0034427

TUMOR MARKER 4: NOT APPLIC
TUMOR MARKERS: UNKNOWN, NO INFORMATION
UNKNOWN, NO INFORMATION
UNKNOWN, NO INFORMATION

SUMMARY STAGE: DISTANT MET
Dx CONFIRMATION: POS HISTOL

SIZE EOD: 002
EXTENSION: 80 / 99
LN: 0
NODES POS/EXAMINED: 99 / 99
PED STAGE SYS/CODER: 88 / 88 / 0

TNM EDITION: 5TH ED
TNM CODER: C MANAGING P ---------- O NOT STAGED

AJCC STAGE TNM C: ------- / ------ / ------ P: ------- / ------ / ------ O: ------- / ------- / ------
TNM
TNM
AJCC SUM STAGE C: 4A P: ------- O: -------

SITE OF DIST METS 1: NONE 2: NONE 3: NONE
SUSPENSE: COMPLETE

DIAGNOSTIC PROCEDURES

PHYSICAL EXAM
Hx Stage 4A follicular lymphoma, tx w/chemo. Relapse 6/22/01, tx w/chemo. Consult @ COH 8/01 for HD chemo/PSCR. Returned to COH 11/01 for PSCR.

X-RAYS/SCANS
All lymph node workup was neg except for the supraclav nodal area (Per OS infor in chart). 6/20/01 CT Abd & Pel: Large infilt mass involv L retroperitoneum & extending into pel.

SCOPES

LABORATORY TESTS
7/10/01 HIV Neg.

OPERATIVE FINDINGS

PATH/AUTOPSY (GROSS & MICRO) PATH Bx: S01-6419 PATH SURG
S016419 (OS slides) 6/13/98 R supraclav bx (OS Path site): Follicular lymphoma, Gr 3. 7/2/98 BM core bx: Involve w/mal lymphoma. 6/22/01 (OS path) Soft tis L retroperitoneal area core bx: Non-Hodgkins lymphoma. 6/27/01 BM bx: NED.

FINAL DIAGNOSIS
FOLLICULAR LYMPHOMA

CONFIDENTIAL

SUM SURGERY

DATE	SURGERY	REASON NO	SCOPE LN SURG	# NODES	OTHER SITE	HOSP NO.
00/00/0000	00	1	9 UNKNOWN	99	0	
8 NOT APPLIC			Approach: 0 NOSURGERY			
PROCEDURE 1 00/00/0000	00		9 UNKNOWN	99	0	
8 NOT APPLIC						
PROCEDURE 2 00/00/0000	00		9 UNKNOWN	99	0	
8 NOT APPLIC						
PROCEDURE 3 00/00/0000	00		9 UNKNOWN	99	0	
8 NOT APPLIC						

DATE NON CA: 06/13/1998 NON CA DIR: 02 / 00 RECONST SURG: 9 / 9

RAD.

00/00/0000 0 / 0 1
RAD SEQ 0
9 9

CHM.

06/18/1998 3 / 0 0 CHOP x 6cyc thru 10/98 (Kaiser).

HO.

00/00/0000 0 / 0 1

IM.

00/00/0000 0 / 0

OTH.

00/00/0000 0 / 0

REMARKS

Fhx Ca: Neg. Smok/ETOH: Neg.

RACE - WHITE

DATE PAT CONT: 02/28/2002
VITAL STAT: ALIVE
DATE L TUM ST: 02/28/2002
TUM STAT: FREE
TYPE LAST PAT FUP: READMISSION
TYPE LAST TUM FUP: READMISSION

DATE 1st ENTERED	DATE COMPLETED	DATE PRINTED	ABS. INIT	COD PROC	TUM REC NO.
03/12/2002	04/17/2002	06/04/2002	CAC	20	01

E-Path:

Automated Case Finding

ISIS-Registrar:

Electronic Path Reports

ISIS-Registrar 1.0

File Reports Window Help

Pathology Reports

Chart/MRN

084113
250492
005633
269126
005633
140974
323506
047860
391717
452616
008772
311264
178928
456750
225729
123508
459006
190143
352635
00463113
093780
093780
267667
253544
00466144
267667
00466902
373072
146693
1091586
9054644
1089325

Not Reviewed | Updated by:

PATHOLOGY REPORT

Report Identification		Patient Information			
Institution:	CCO	Chart/MRN:	452616	Address:	Address Sample15
Pathology ID:	97:S2532	SSN/SIN:			
Report Date:	1997-04-01	Surname:	SAMPLE15	City/Town:	
Report Type:	Correction	Given Name:	WALTER	State/Prov.	
Requester ID:		Sex:	M	Zip/Post Code:	
Requester:	,	Date of Birth:	1936-09-01	Country:	
Procedure Date:	1997-04-01	Age:	60 (at procedure date)		
Surgeon ID:	063362	Insurer:			
Surgeon:	AMURAWAIYE, EMMANU	Account No.			
Pathologist ID:	051987	Race:			
Pathologist:	LAFARGA, A	Ethnicity:			

DIAGNOSIS

A. SEGMENT OF SIGMOID COLON. WELL DIFFERENTIATED ADENOCARCINOMA INFILTRATING FULL

THICKNESS OF THE BOWEL WALL AND FOCALLY THE SEROSA.

ALL THE MESENTERIC LYMPH NODES ARE FREE OF METASTATIC TUMOR.

TUBULAR ADENOMA.

ADENOMATOUS POLYP.

B. PROXIMAL AND DISTAL RINGS OF ANASTOMOSIS. FREE OF TUMOR.

/jlm

Clinical History Not available on the requisition

Company Name
Address/Telephone

32 reports | ADMIN [SYS] @copath_agent_1 HOME | 210 MB Pagefile available

Start ISIS-Registrar 1.0 9:44 AM

BITS:

Biostatistics Information Tracking System

Clinical Trials at City of Hope

- **Approximately 350 open trials at City of Hope**
  - **Over 150 treatment trials open to accrual**

- **Sponsorship of protocols:**
  - **~1/3 National Cooperative Groups**
  - **~1/3 Pharmaceutical Sponsors**
  - **~1/3 Intramural Trials**

- **Developed centralized in-house data repository / data collection application in 1989:**
  - ***BITS: Biostatistics Information Tracking System***

6 6 7 9

day 100 = 10/19/01

AUTOLOGOUS BMT
DATA COLLECTION FORM:
Hodgkin's Disease & Non-Hodgkin's Lymphoma

Missing value codes for fill-in blanks: -8=N/A and -9=missing or unknown
Please do not use missing value codes in date fields.

Date Form Completed: 2 / 12 / 02 Completed By: Mudra Nathwani.

A. DEMOGRAPHICS First Name: [redacted] MI: - Last Name: [redacted]

Treating MD: Molina

Medical Record Number: [redacted] Protocol Number: 5 5 5 5 5

B. PATHOLOGY AT DIAGNOSIS Date of Diagnosis: 1 / 19 / 01 Missing ☐ (-9)

Diagnosis:

Hodgkin's Disease	☐ (4)	Non-Hodgkin's Lymphoma	☒ (3)
Composite HD & NHL	☐ (10)	Missing	☐ (-9)
Other _____ Specify	☐ (99)	N/A	☐ (-8)

Date of Biopsy: 1 / 19 / 01 * Site: 2, 98 → Clavicle, ___, ___, ___, ___ Missing ☐ (-9)

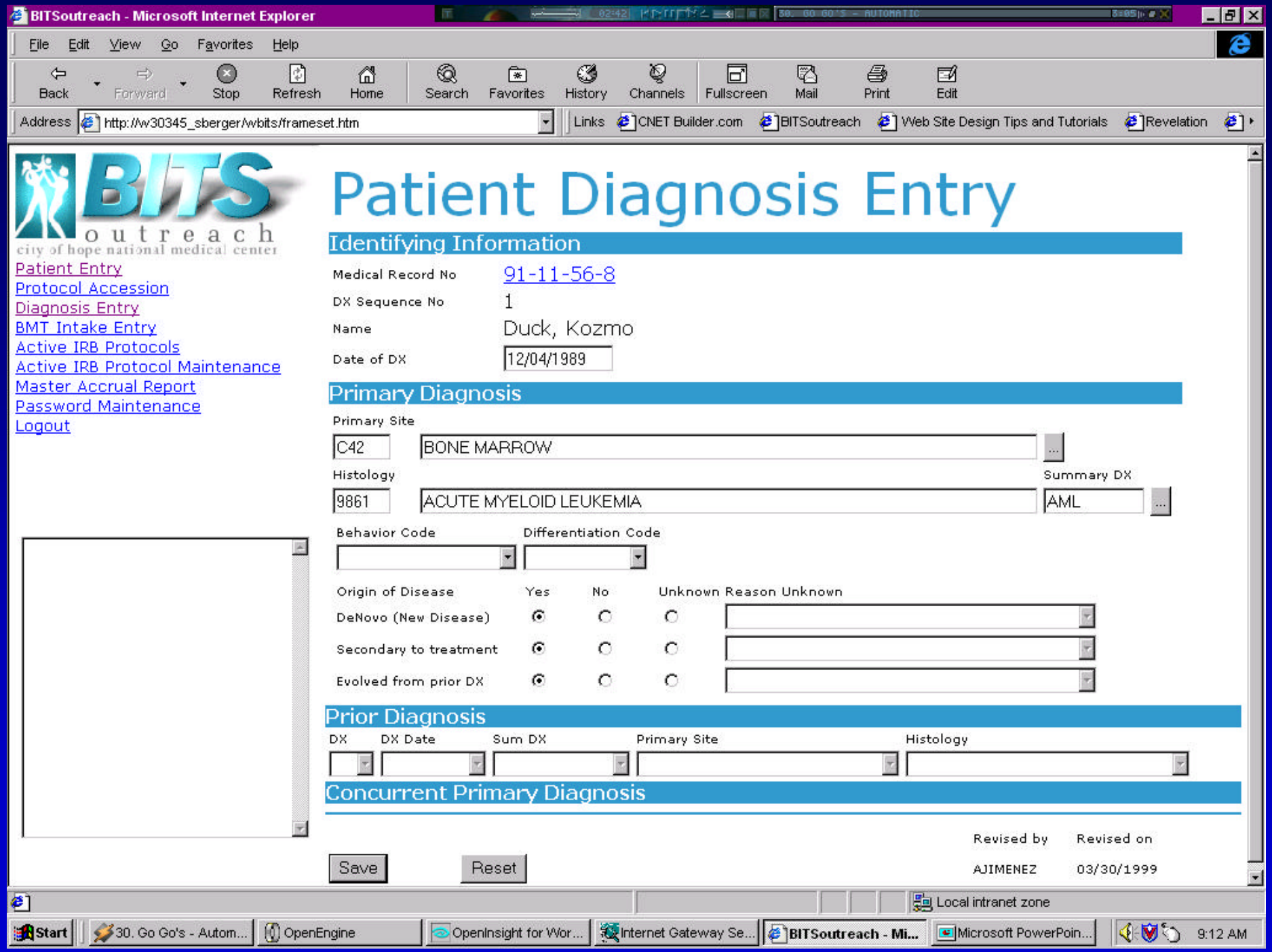
* Use the Site Codes on page 3

Hodgkin's Disease:		Non-Hodgkin's Lymphoma:		
Lymphocyte predominance	☐ (1)	Low grade	☐ (5)	Specify*: _____
Nodular sclerosis	☐ (2)	Intermediate grade	☒ (6)	Specify*: E (Mantle cell)
Mixed cellularity	☐ (3)	High grade	☐ (7)	Specify*: _____
Lymphocyte depletion	☐ (4)	Other _____ Specify	☐ (99)	
Other _____ Specify	☐ (99)	Missing	☐ (-9)	
Missing	☐ (-9)			

BITS II

*“Son of BITS” **

***Caution:** speak slowly & enunciate *very* clearly


BITSoutreach - Microsoft Internet Explorer
File Edit View Go Favorites Help
Back Forward Stop Refresh Home Search Favorites History Channels Fullscreen Mail Print Edit
Address http://w30345_sberger/wbits/frameset.htm
Links CNET Builder.com BITSoutreach Web Site Design Tips and Tutorials Revelation
BITS outreach
city of hope national medical center
Patient Entry
Protocol Accession
Diagnosis Entry
BMT Intake Entry
Active IRB Protocols
Active IRB Protocol Maintenance
Master Accrual Report
Password Maintenance
Logout
Patient Diagnosis Entry
Identifying Information
Medical Record No 91-11-56-8
DX Sequence No 1
Name Duck, Kozmo
Date of DX 12/04/1989
Primary Diagnosis
Primary Site
C42 BONE MARROW
Histology
9861 ACUTE MYELOID LEUKEMIA
Summary DX
AML
Behavior Code
Differentiation Code
Origin of Disease Yes No Unknown Reason Unknown
DeNovo (New Disease)
Secondary to treatment
Evolved from prior DX
Prior Diagnosis
DX DX Date Sum DX Primary Site Histology
Concurrent Primary Diagnosis
Revised by AJIMENEZ
Revised on 03/30/1999
Save
Reset
Local intranet zone
Start
30. Go Go's - Autom...
OpenEngine
OpenInsight for Wor...
Internet Gateway Se...
BITSoutreach - Mi...
Microsoft PowerPoin...
9:12 AM

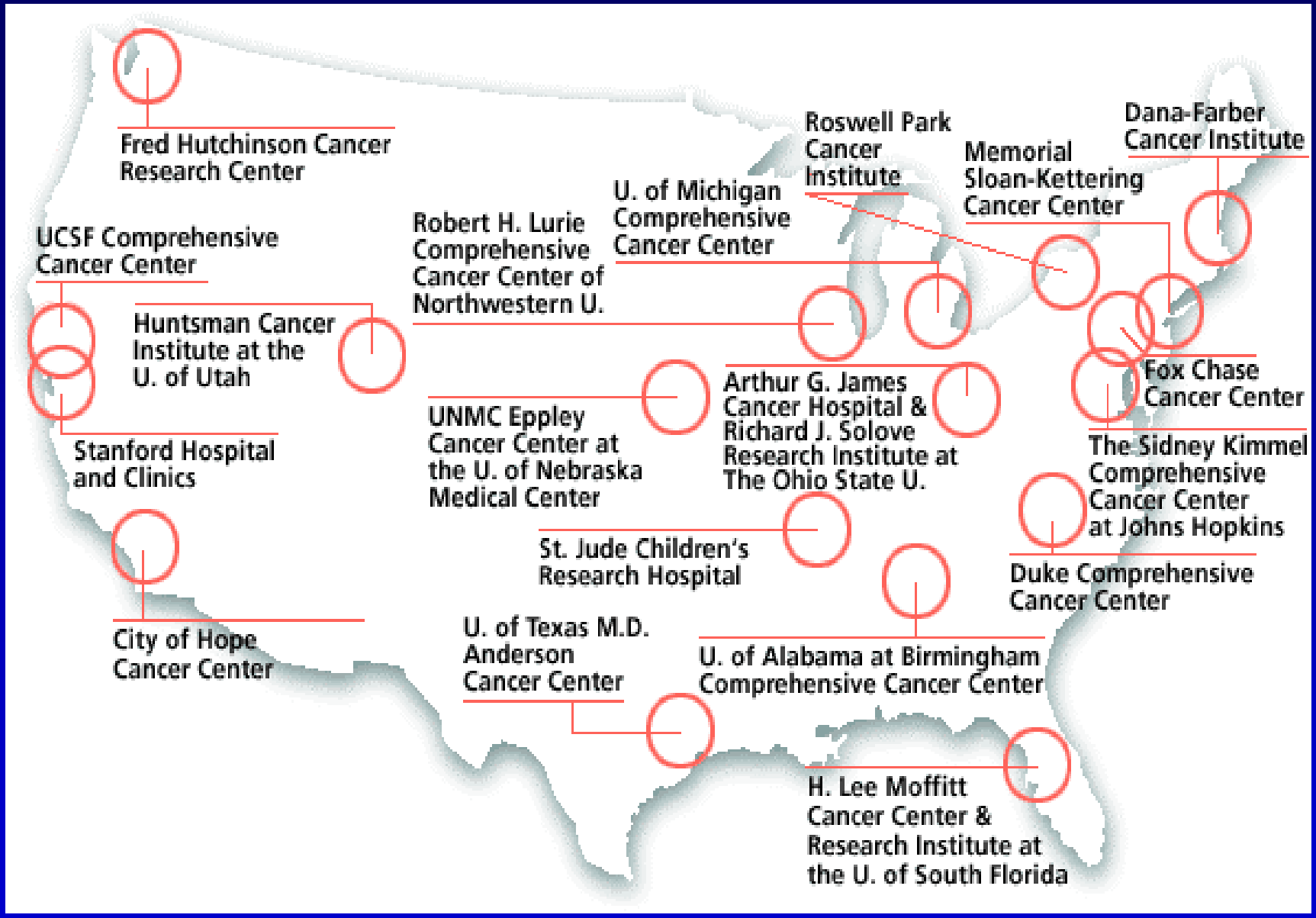
Enhancements to BITS II

- **Migrating from AREV to MS SQL Server 7.0**
- **Web-browser application for direct data entry from local or remote sites**
  - **Allow direct data entry from multiple sites**
- **Developing additional electronic interfaces**
  - **IRB database system, HLA data, cytogenetics**
- ***Re-engineering of data collection to include Cancer Registry and outcomes staff***

Internet-based Data System:

The NCCN Outcomes Research Database

National Cancer Center Network (NCCN)



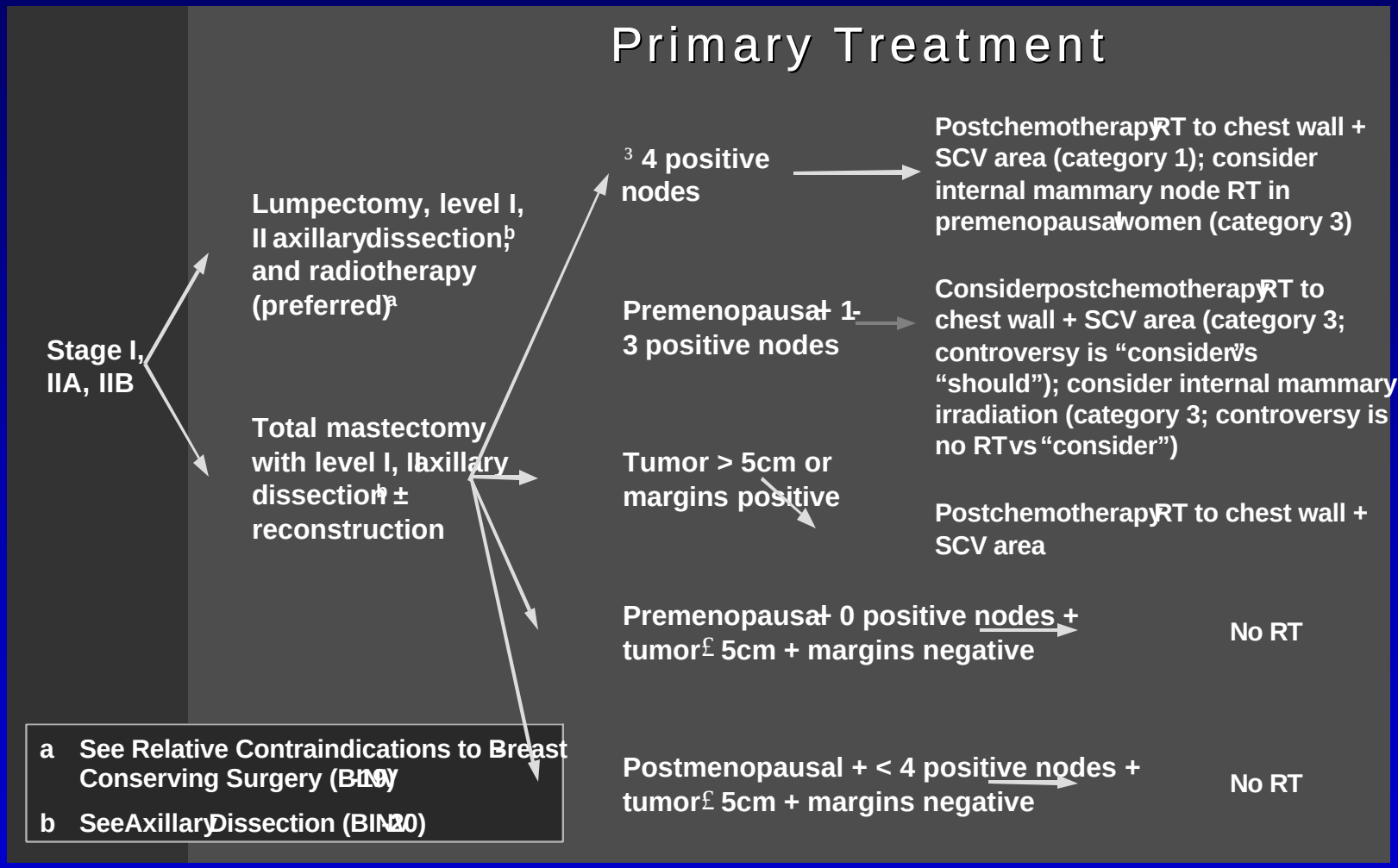
NCCN Outcomes Research Data System

- **Alliance of 20 national cancer centers**
- **Experts developed evidence-based clinical guidelines**

Sample Page from the NCCN Breast Cancer Guidelines

Breast Cancer (Invasive)

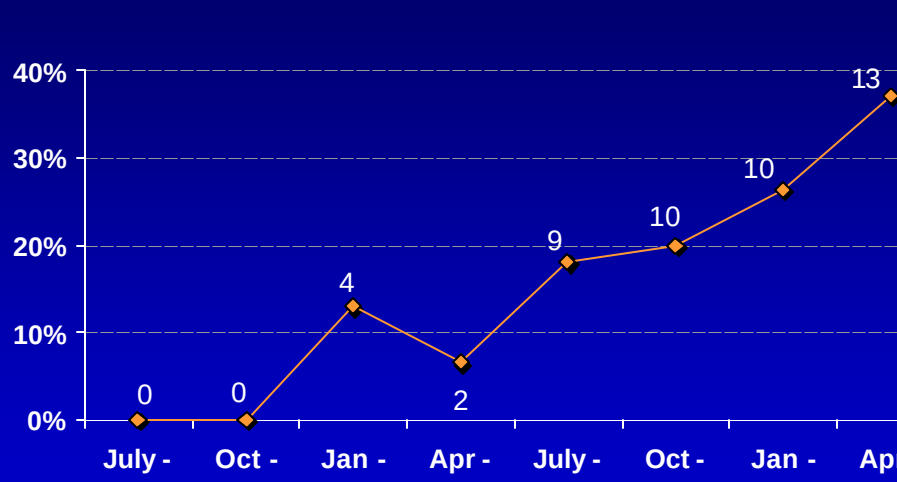
NCCN® VERSION 1.1999 PRACTICE GUIDELINES



NCCN Outcomes Research Data System

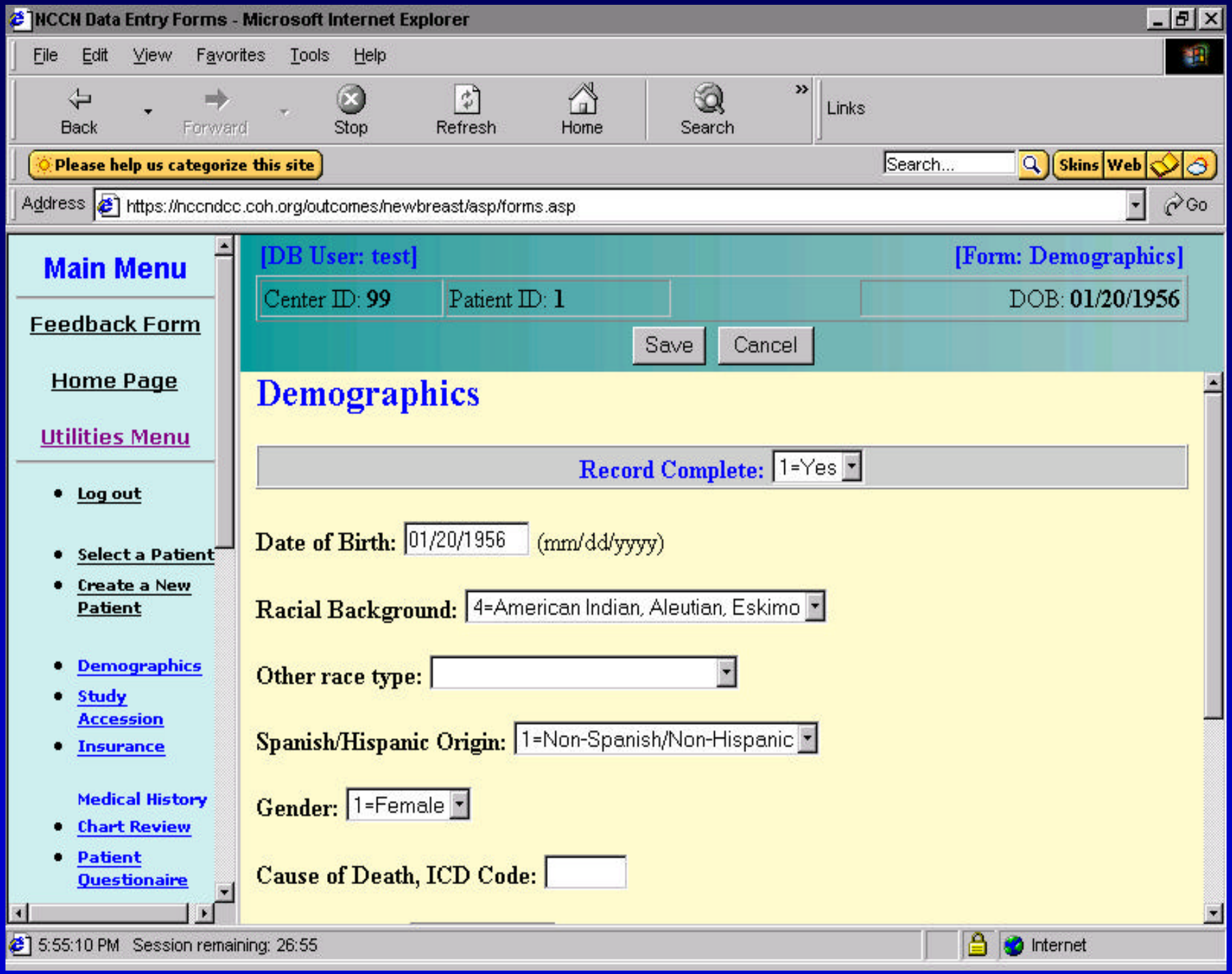
- Alliance of 20 national cancer centers
- Experts developed evidence-based clinical guidelines
- Conducting research to measure treatment patterns and outcomes
- COH serves as Data Coordinating Center
- Created an Internet-based data system, deployed nationwide for past 5 years

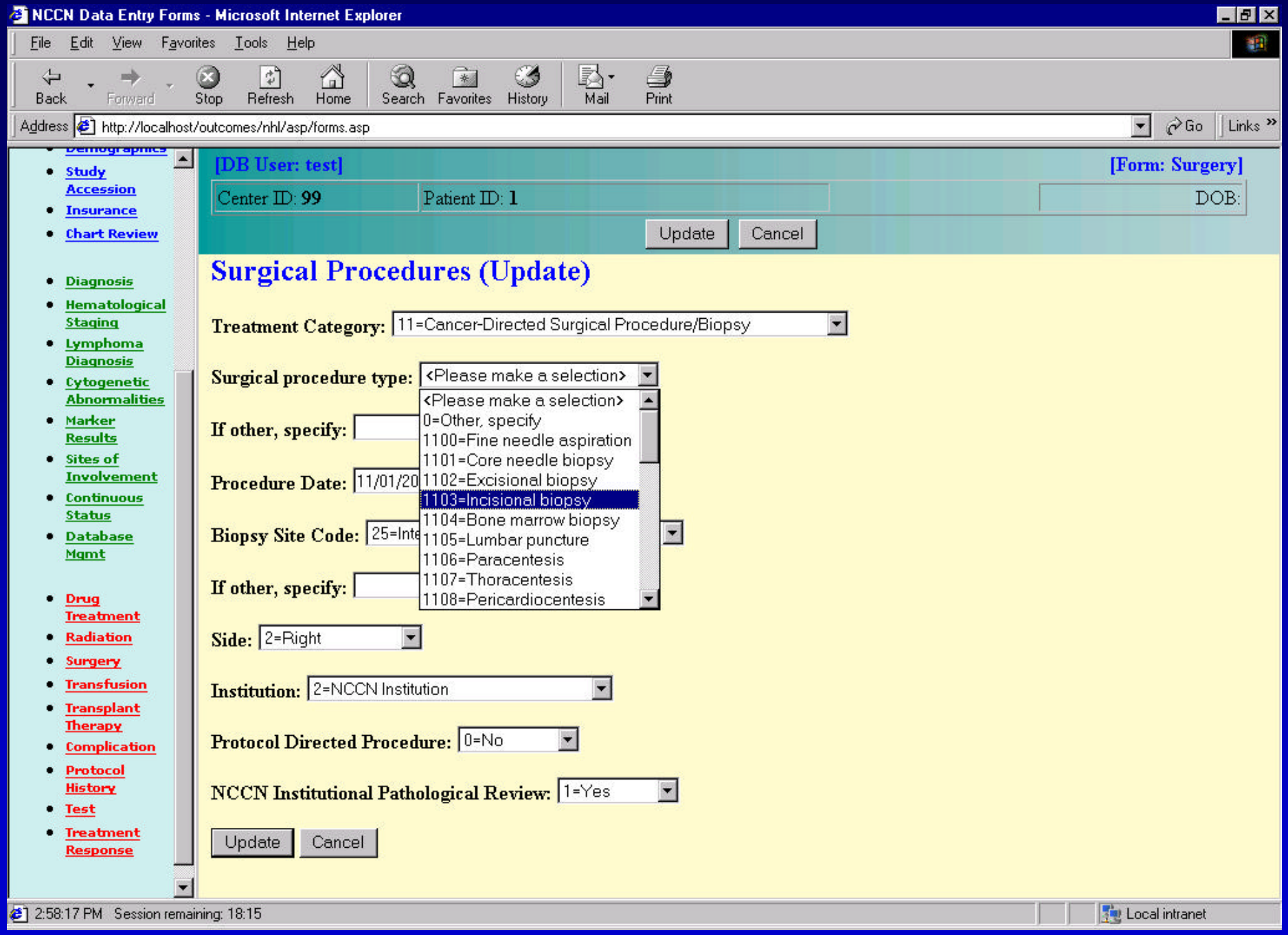
Use of Tamoxifen in Patients with DCIS Within 180 Days of Diagnosis



** DCIS: Ductal Carcinoma in Situ*

Tamoxifen time trend is statistically significant controlling for center (p <0


NCCN Data Entry Forms - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Forward Stop Refresh Home Search Links
Please help us categorize this site
Search... Skins Web
Address https://nccndcc.coh.org/outcomes/newbreast/asp/forms.asp
Go
Main Menu
Feedback Form
Home Page
Utilities Menu
Log out
Select a Patient
Create a New Patient
Demographics
Study Accession
Insurance
Medical History
Chart Review
Patient Questionaire
[DB User: test]
[Form: Demographics]
Center ID: 99
Patient ID: 1
DOB: 01/20/1956
Save
Cancel
Demographics
Record Complete: 1=Yes
Date of Birth: 01/20/1956 (mm/dd/yyyy)
Racial Background: 4=American Indian, Aleutian, Eskimo
Other race type:
Spanish/Hispanic Origin: 1=Non-Spanish/Non-Hispanic
Gender: 1=Female
Cause of Death, ICD Code:
5:55:10 PM Session remaining: 26:55
Internet


NCCN Data Entry Forms - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Forward Stop Refresh Home Search Favorites History Mail Print
Address http://localhost/outcomes/nhl/asp/forms.asp
Go Links
Demographics
Study Accession
Insurance
Chart Review
Diagnosis
Hematological Staging
Lymphoma Diagnosis
Cytogenetic Abnormalities
Marker Results
Sites of Involvement
Continuous Status
Database Mgmt
Drug Treatment
Radiation
Surgery
Transfusion
Transplant Therapy
Complication
Protocol History
Test
Treatment Response
[DB User: test]
[Form: Surgery]
Center ID: 99
Patient ID: 1
DOB:
Update
Cancel
Surgical Procedures (Update)
Treatment Category: 11=Cancer-Directed Surgical Procedure/Biopsy
Surgical procedure type: <Please make a selection>
<Please make a selection>
0=Other, specify
1100=Fine needle aspiration
1101=Core needle biopsy
1102=Excisional biopsy
1103=Incisional biopsy
1104=Bone marrow biopsy
1105=Lumbar puncture
1106=Paracentesis
1107=Thoracentesis
1108=Pericardiocentesis
If other, specify:
Procedure Date: 11/01/20
Biopsy Site Code: 25=Inte
If other, specify:
Side: 2=Right
Institution: 2=NCCN Institution
Protocol Directed Procedure: 0=No
NCCN Institutional Pathological Review: 1=Yes
Update
Cancel
2:58:17 PM Session remaining: 18:15
Local intranet

John S. Silva Marion J. Ball Christopher G. Chute
Judith V. Douglas Curtis P. Langlotz
Joyce C. Niland William L. Scherlis *Editors*

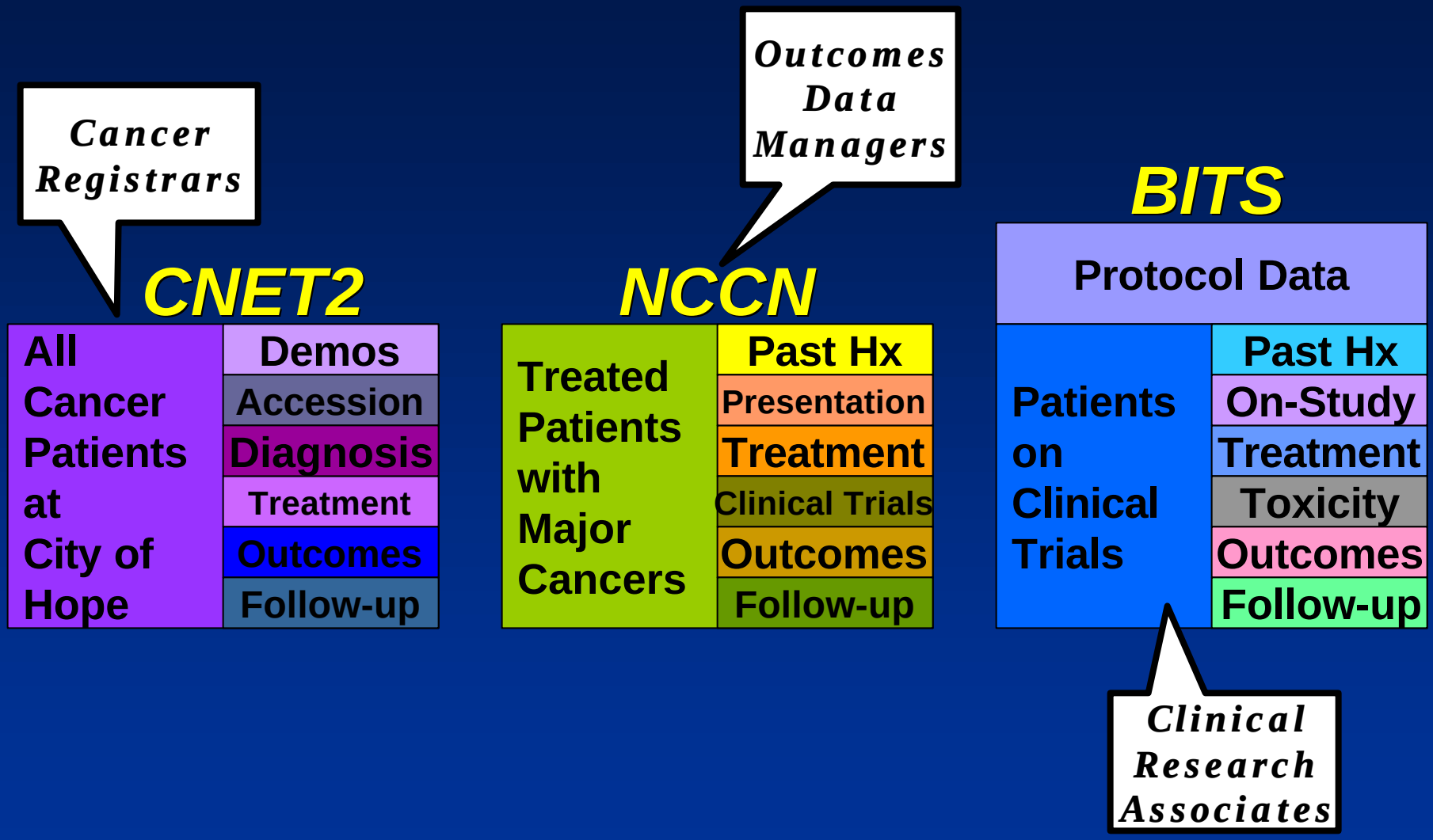
Foreword by Richard D. Klausner

Cancer Informatics

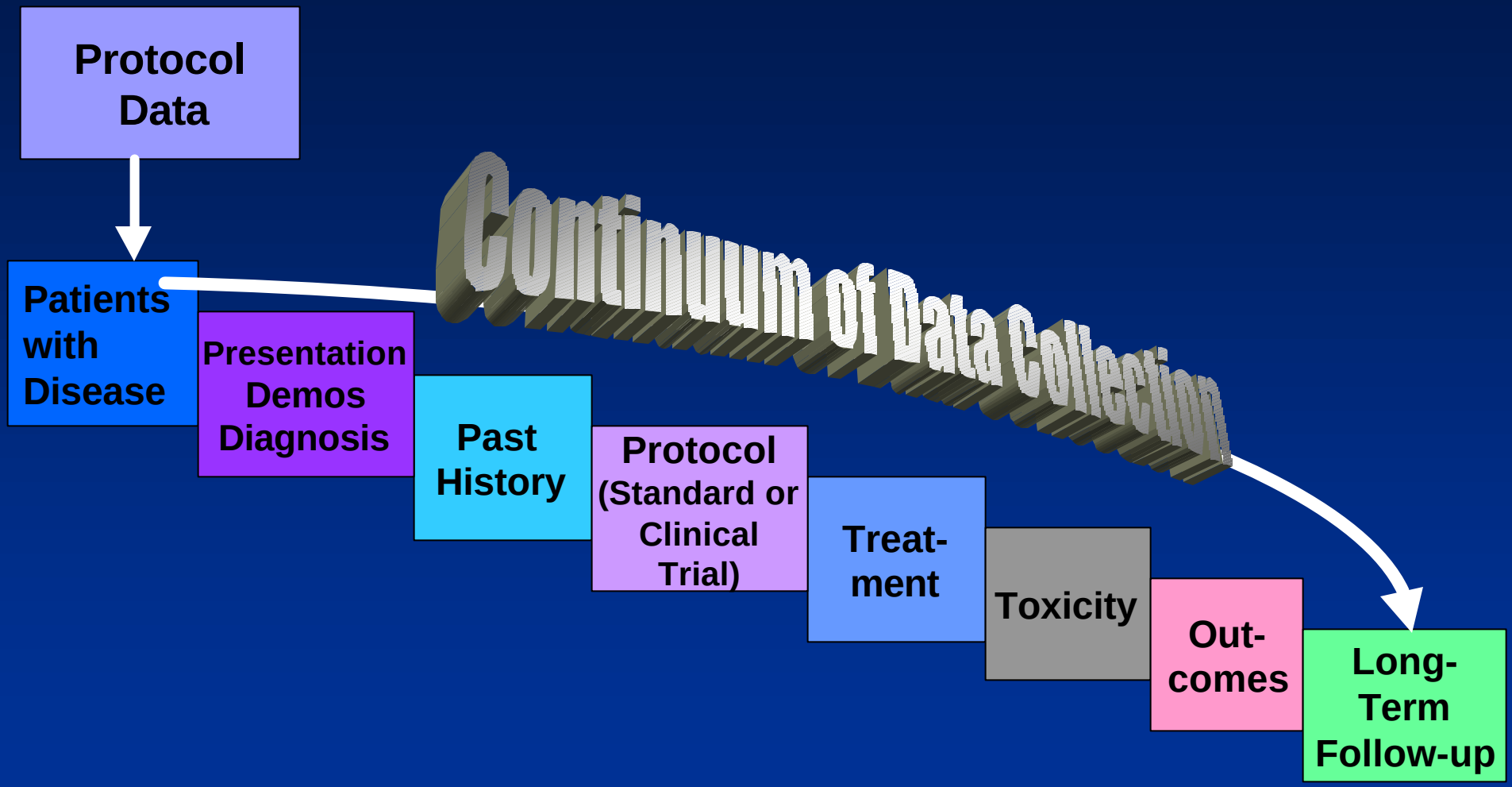
Essential Technologies for Clinical Trials

HEALTH INFORMATICS SERIES

Structure of Clinical Data Systems within the Division of Information Sciences



Future Clinical Data Systems within the Division of Information Sciences



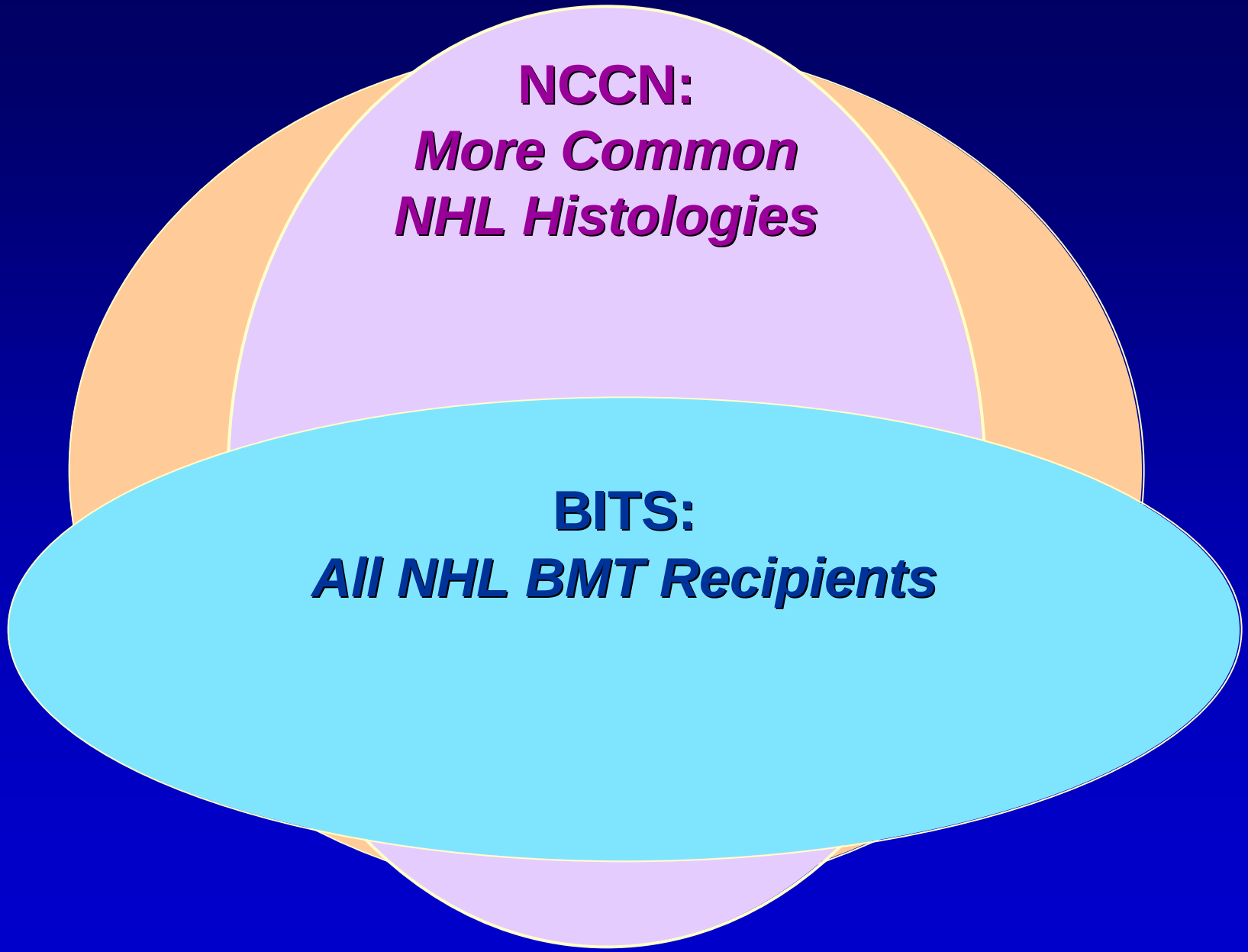
How Do We Get There from Here?

The Analysis……

A Case Study:

Shared Data for Non-Hodgkins Lymphoma (NHL) Patients

Intersecting Data Collection for NHL Patients



Reporting Requirements

Cancer Registry:	**State of California Department of Health Services**
	American College of Surgeons
	SEER
NCCN:	**Disease-specific Executive Committee**
BITS:	**COH Investigators**
	Study Sponsor
	National Cancer Institute

NHL Inclusion Criteria Across the 3 Databases

Inclusion Criterion:	Cancer Registry	NCCN Outcomes	BITS Clinical Trials
Histology	Any	Common only	Any
Age	Any	> 18 years	> 18 years
Prior Cancer	No exclusion, all recorded	No concurrent cancer	No previous cancer
Prior Treatment	Analytic & non-analytic	Primary rx at institution	Any (record all prior rx)
Current Treatment	Any	Any	BMT

Processes & Timing of Data Collection

Cancer Registry: **Data collection starts after completion of 1st treatment course.**
Treatment prior to presentation requested, not required.
No consent required.

NCCN: **Data collection starts after consent.**
Treatment prior to presentation and outside treatment requested.
Consent required prior to data collection.

BITS: **Data collection starts 100 days post-BMT.**
All treatments and relapses prior to presentation to COH must be recorded.
Outside treatment required from physician offices or hospitals.
Consent required.

NHL Data Collected Across the 3 Databases

Data Element	Cancer Registry	NCCN Outcomes	BITS Clinical Trials
Name	☑	○	☑
Medical Rec #	☑	○	☑
DOB	☑	☑	☑
Gender	☑	☑	☑
Race	Very Detailed	Mod Detail	5 categ's
Hisp Origin	☑	☑	☑
Diagnosis	Clinical	Pathological	Pathological
Disease / Vital Status	Follow Annually	6 & 12 Mths, then Annually	Restage @ 30 60, 100 days*

** After 1 year post-BMT annual follow-up done through COH Long-term Followup Office*

Entity-Attribute-Value (EAV) Map

Super Type Entity	Sub Type Entity	Attribute (Data Element)	Description	Domain
Person				
	Patient			
		Gender	Patients Gender	
				BITS
				CNET
				NCCN

System Data Dictionary

CNET

Technical Directory:

This area describes the field format, length, storage, etc.

Business Directory:

This area defines the element from the user's perspective.

Domain Value Map

BITS			CNET		
Data Element	Code	Code Name	Data Element	Code	Code Name
Sex			Gender		
	M	Male		1	Male
	F	Female		2	Female
	X	Unknown		3	Hermaphrodite
	X	Unknown		4	Transexual
	X	Unknown		9	Unknown

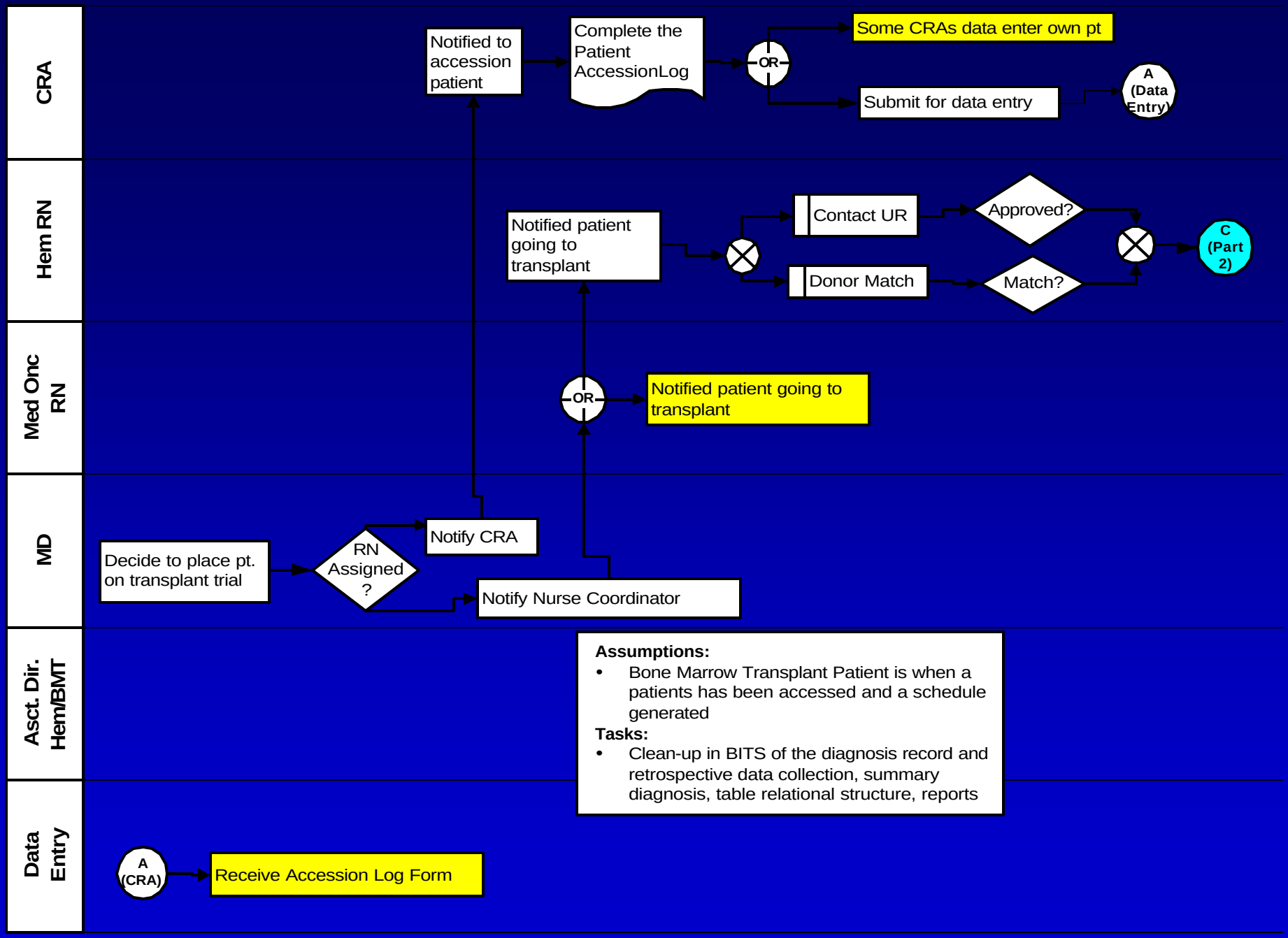
NHL Treatment and Complication Data

Data Element	Cancer Registry	NCCN Outcomes	BITS Clinical Trials
Chemo Start Date	☑	☑	☑
Chemo End Date	Text	☑	☑
Drugs Rec'd	Single vs. Multiple	Coded	Coded
Chemosensivity	○	☑	Inferred
Rx Complications	○	10 Major	Detailed CTC

** After 1 year post-BMT annual follow-up done through COH Long-term Followup Office*

Re-engineering of Workflow

Developing a system in which Cancer Registrars, Outcomes Data Managers and Clinical Research Associates combine efforts to eliminate redundant data, share data collection efforts


CRA
Hem RN
Med Onc RN
MD
Asct. Dir. Hem/BMT
Data Entry
Notified to accession patient
Complete the Patient AccessionLog
OR
Some CRAs data enter own pt
Submit for data entry
A (Data Entry)
Notified patient going to transplant
Contact UR
Donor Match
Approved?
Match?
C (Part 2)
OR
Notified patient going to transplant
Decide to place pt. on transplant trial
RN Assigned ?
Notify CRA
Notify Nurse Coordinator
Assumptions:
• Bone Marrow Transplant Patient is when a patients has been accessed and a schedule generated
Tasks:
• Clean-up in BITS of the diagnosis record and retrospective data collection, summary diagnosis, table relational structure, reports
A (CRA)
Receive Accession Log Form

Staging of Data Integration Across Data Systems

- **Initial Presentation**
  - **Diagnosis and Staging**
- **Long Term Follow-up**
  - **Relapse and Survival**
- **General Case Reviews**

1. Diagnosis and Staging

- **Change timing of CTR case review**
  - **CRA contacts assigned CTR when patient to go on clinical trial**
  - **Accept Cancer Registry diagnosis information as most accurate for all cancers**
- **Process created for rectification of disagreements**
  - **Provides additional quality assurance of data**
- **Fields added to track method of diagnosis**
  - **Different rules by reporting requirements**
- **Non-cancer diagnoses handled by CRA**

2. Follow-up Data

- **Phase in Cancer Registry collection of follow-up data, beginning with patients on national cooperative group studies**
  - **death and relapse only**
  - **contact outside sources for more detail**
    - **including some toxicity data**
  - **bring follow-up to within 12 months minimum**
- **CRAs continue to follow other patients**
- **Phase in additional Cancer Registry follow-up now that this segment is successful**

3. Sharing of Case Reviews

- **Initiator of case shares completed data**
  - **Generally NCCN or BITS data completed first**
  - **Central registration of cases to determine if data already abstracted for a given patient**
  - **Online view-only access to each data system**
- **Subsequent data abstractors utilize information collected to date**
  - **Saves time, must be mindful of differences in definitions, timing, context**
  - **Provides additional quality assurance**
  - **Process for resolving discrepancies**

Issues and Findings to Date

- **Most granular data should precede others**
- **Need to be continually mindful of differences in definitions, context of data collection**
  - **Accurate metadata!**
- **Improved accuracy achieved in many cases**
  - **Price is delays in collection, timing issues**
- **Able to bring CT follow-up into compliance through assistance by Registry staff**
- **Has led to improved communication and data flow between the 2 areas**
  - **Physical joint relocation of staff offices a plus**

Future Plans

Based on lessons learned through this process:

- **Provide electronic data exchange among CNext, BITS, and NCCN systems**
  - **No redundant data entry**
  - **Requires ability to add user-defined builds**
  - **Ideal is to import electronic files into CNET**
- **Explore enhancement of data sharing e.g. via coded treatment data for all modalities**
- **Develop Web application interface for data collection**
- **Include systems as operational data ‘feed’ into research data warehouse**
- **Lessons learned are informing our model for “Fully Integrated Research Standards & Tech”**

Prediction for the Future of Tumor Registrars...

Outcomes Research Associates of the Future!

Thanks to my collaborators….

Ina Ervin, Cancer Program Coordinator

Janet Nikowitz, Outcomes Research Coordinator

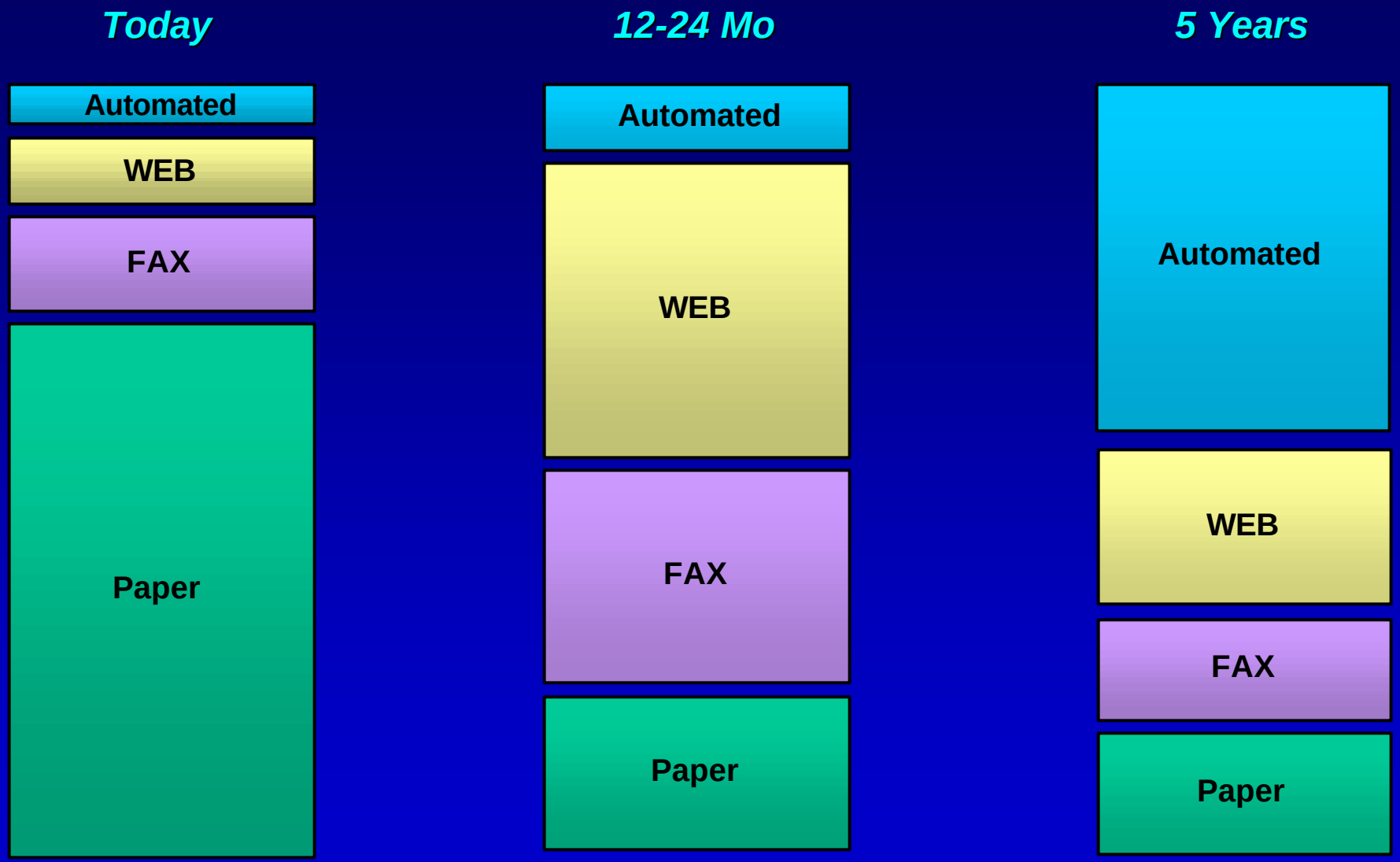
Joycelynne Palmer, Sr. Systems Analyst

Geri Connie, CTR, Outcomes Data Manager

Mudra Nathwani, Clinical Research Associate

QUESTIONS?

Transition from Today…



Our Ultimate IAIMS Integration Project:

Research Data Warehouse

Obstacles to Data Warehousing for Research

- **Semantic/conceptual barriers pose the most substantial challenges to database interconnectivity**
- **Standards for patient data are needed**
  - **Data collection methods, taxonomies, level of comprehensiveness, granularity, and knowledge representation schema**
- **Leads to the requirement for “metadata”, including unique concept identifiers and mapping tools**

Operational Systems:

Hosp Admin | Patient Care | State Reporting | Clinical Trials | Outcomes | Res

System:	A2K	OACIS	CNET	BITS	NCCN
Data:	<- - - -Vocabulary- - - ->				
Demos	Propr-1	Propr-1	ACOS	Propr-1	Propr-2
Diagnosis	ICD-9*	SNOMED*	ICD-0*	ICD-9, NCI	ROADS
Treatment	CPT	(Text)	(Text)	Propr-3	Propr-4
Toxicity	-----	(Text)	-----	CTC	Propr-5
Response	-----	(Text)	ACOS	Propr-6	ACOS

*In the UMLS

Operational Systems:

Unified clinical research system

System:	A2K	OACIS	CNET	BITS	NCCN
Data:	<- - - -Vocabulary- - - ->				
Demos	Propr-1	Propr-1	ACOS	Propr-1	Propr-2
Diagnosis	ICD-9*	SNOMED*	ICD-0*	ICD-9, NCI	ROADS
Treatment	CPT	(Text)	(Text)	Propr-3	Propr-4
Toxicity	-----	(Text)	-----	CTC	Propr-5
Response	-----	(Text)	ACOS	Propr-6	ACOS

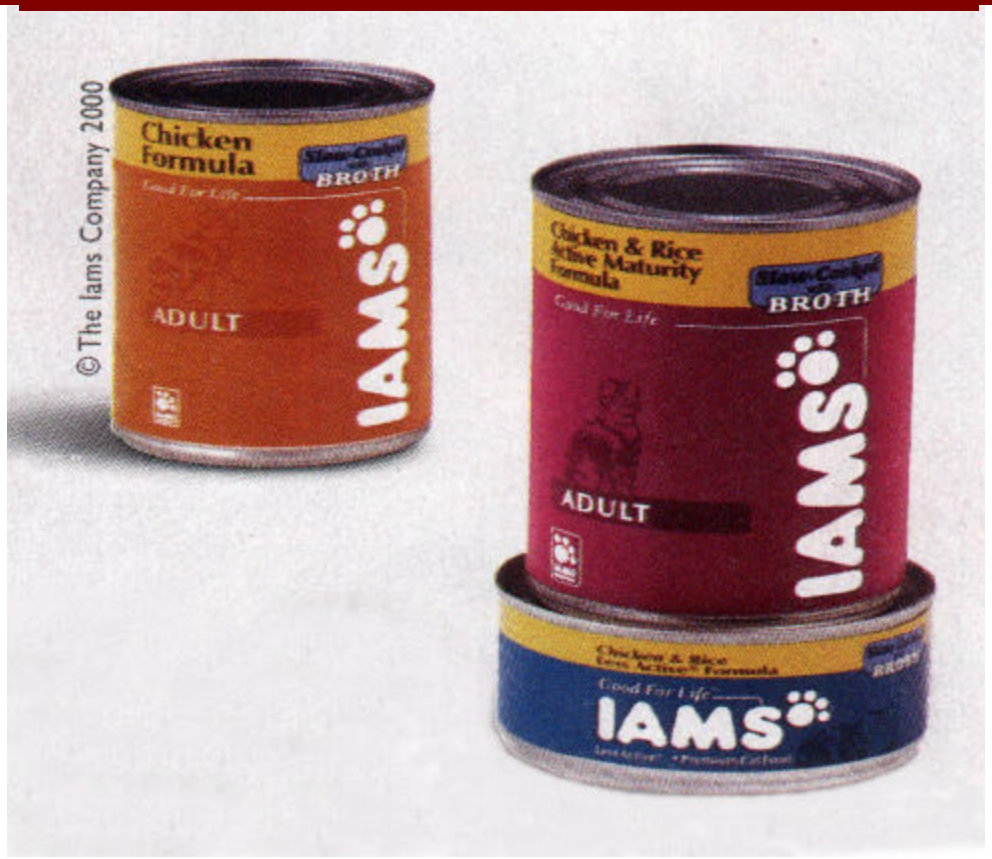
**In the UMLS*

Metadata Are Critical!

- **Data about the data:**
  - **What data are available?**
  - **Who is responsible for it? (“data stewards”)**
  - **Where is it stored? In what format? etc**
- **Technical Directory:**
  - **Field name, format (character/numeric), length….**
- **Business Directory:**
  - **Definition of data field, how is it collected, what is the context, where is it used, …..**

IAIMS:

Integrated Advanced Information Management Systems



Integrated Advanced Information Management System (IAIMS) Program

- **Founded by the National Library of Medicine in 1984 to facilitate and fund information systems and integration**

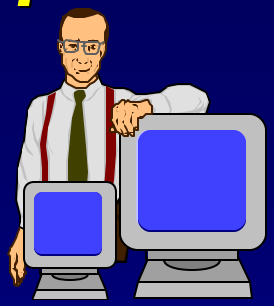
- **Goals:**
  - **Manage medical knowledge more effectively**
  - **Provide for a system of comprehensive and convenient information access**

- **City of Hope funded by *IAIMS* program:**
  - **Phase I Planning Grant: 1996-1998**
  - **Phase II Implementation Grant: 1998-2003**


allocate
collaborate
formulate
aggregate
automate
communicate
Integrate
consolidate
facilitate
accelerate
incorporate
motivate
activate
liberate

Biostatistics Information Tracking System (BITS)

CPRMC / IRB



Mainframe

Accession Logs

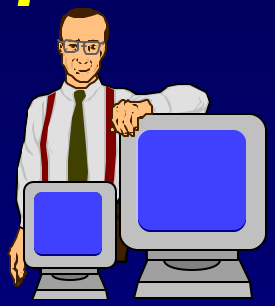
Protocol Data

Patient Data

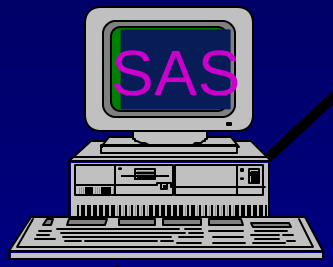
Reports (NCI)

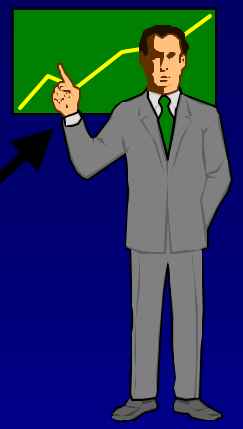
Biostatistics Information Tracking System (BITS)

CPRMC / IRB



Mainframe





Accession Logs

Protocol Data

Patient Data

Data Request Forms

Reports (NCI)

Clinical Trials Modules

Study Work-up Report