Databases in Clinical Research

Harvard-MIT Division of Health Sciences and Technology
HST.951J: Medical Decision Support

Overview

- Background: History and utility of clinical data repositories
- Strategies: Integrating the outcomes tracking database into clinical workflow
- Brigham and Women’s Catheterization Laboratory Database: Architecture, Advances, Limitations
- Examples of Data Exploration: Risk models, “drilling down”, and device safety monitoring

Need for Clinical Data Repositories

- Randomized clinical trials are gold standard for testing a hypothesis, but there are significant limitations:
  - generalizability
  - timeliness
  - cost $$$

Cost of Randomized Clinical Trials

- Estimated cost of RCT:

  Drug Trial: $15,000/patient

  - 1000 patient trial: $15MM

- Simply too expensive to answer every relevent clinical question with prospective blinded RCT.

Clinical Registries

- **While RCT's test hypotheses, the real world of clinical practice is a registry.**

  All patients (generalizability)
  Dynamic (timeliness)

- **Significant *Potential* cost savings when automated clinical registry (database system) bundled with other functional requirements**

  clinical reporting, billing, inventory control

History of Successful Clinical Registries

- Duke Database
- Washington Heart Center
- Beth Israel Hospital, Boston
- Cleveland Clinic
- Mayo Clinic
- Massachusetts General Hospital

Why Clinical Cardiology?

- High volume clinical sites
- High event rates – death, MI, revascularization, rehospitalization, etc.
- High profile
- High cost to study

Applications of Clinical Databases:

- **Clinical Research:**
  - **Retrospective "Hypothesis Generator"**
  - **Data mining**
  - **Prospective automated CRF**
  - **Risk prediction modeling**
- **Quality Assurance:**
  - **Interprovider variability**
  - **Benchmark review – ACC NCDR**
- **Business and Operations Review –** Turnover times, referral patterns
- **Regulatory Requirements –** State DPH

Overview

- Background: History and utility of clinical data repositories
- Strategies: Integrating the outcomes tracking database into clinical workflow
- Brigham and Women's Catheterization Laboratory Database: Architecture, Advances, Issues
- Examples of Data Exploration: Risk models, "drilling down", and device safety monitoring

Strategies for Maintaining Clinical db

Three Strategies:

- **Prospective/retrospective off-line chart review**
- **Data extraction w/ supplemental chart review**
- **Complete integration into electronic record system**

Clinical Database Strategies:

Parallel Chart Review

Hybrid Strategy

Fully Integrated

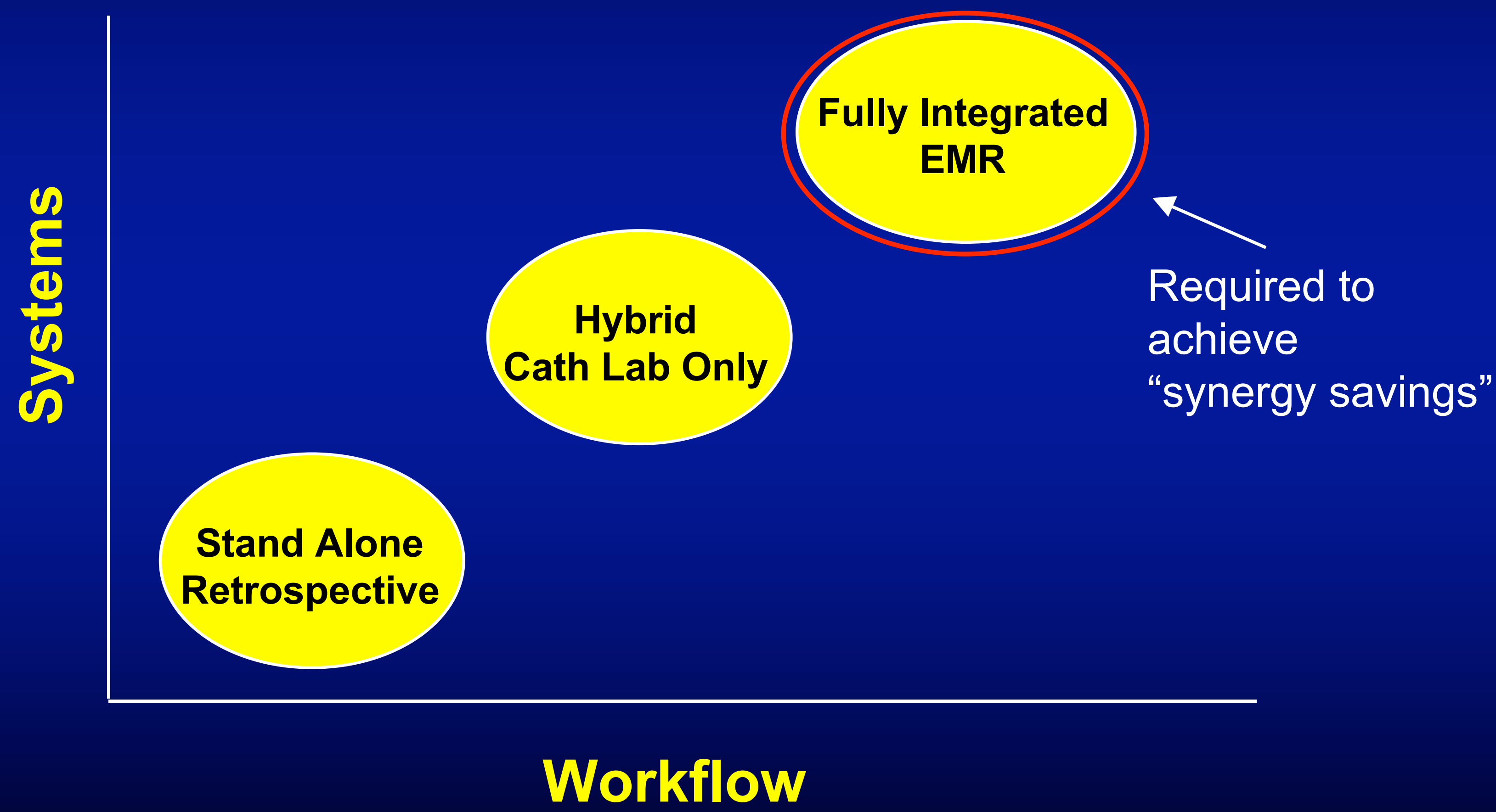
-- independent of clinical process
-- focus on data quality
-- maintain current workflow

-- requires team of coders
-- COST $$$

-- purely prospective
-- integral part of routine workflow
-- lowest cost (??)

-- data quality issues
-- data management

Integration Dimensions:

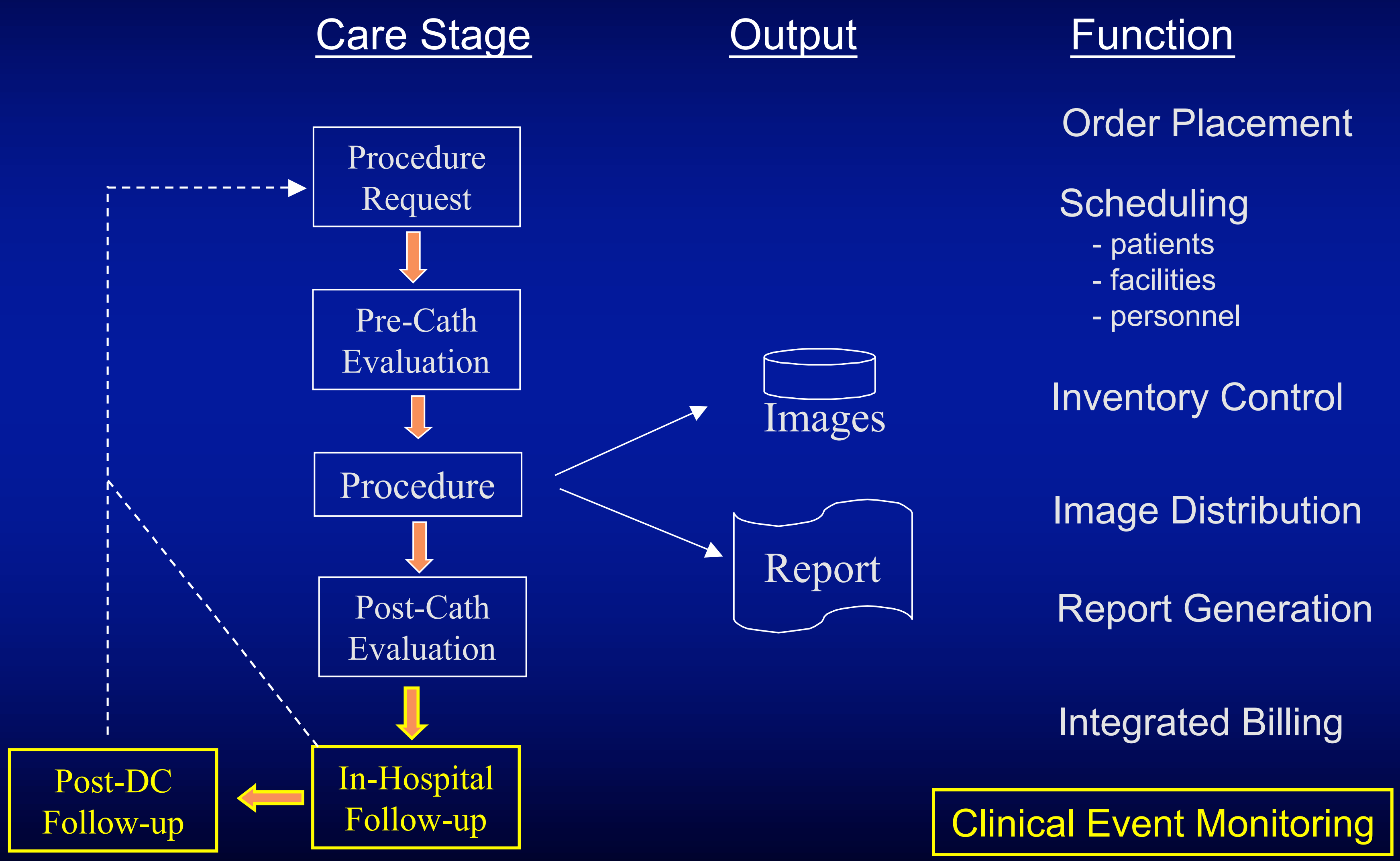


Multi-Use Function of Clinical Cath Lab Databases:

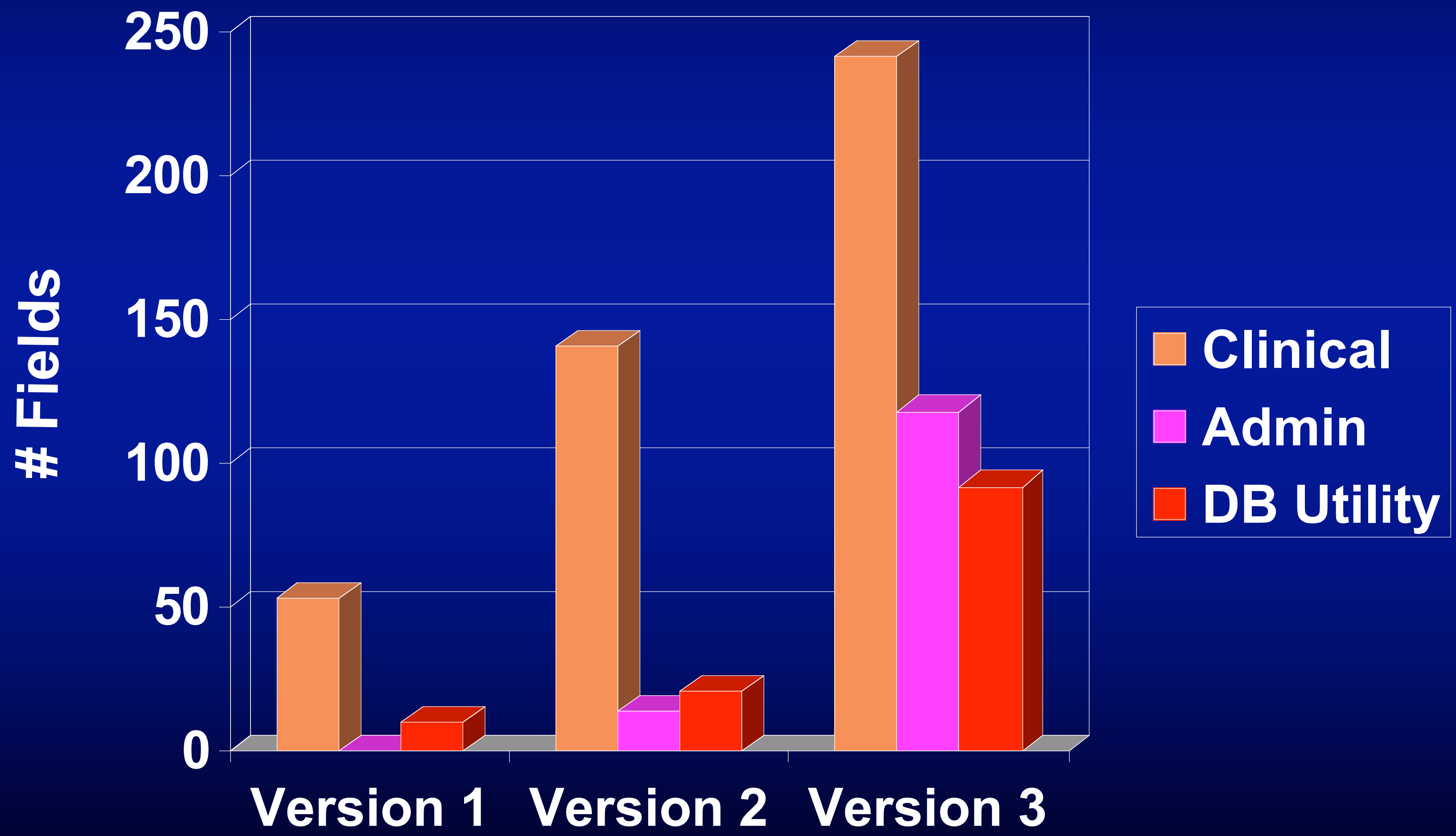
- Clinical Outcomes Tracking Database:
  - Retrospective Clinical Research
  - Quality Assurance
  - Administrative reporting
- Clinical report generation (structured reporting; transcription templates)
- Technical and Professional Billing
- Inventory Management

- Increased complexity of database with each additional functional layer.

Information Flow Integrated Into Care Process



Evolutionary Growth in DB Design: BWH CCL DB



Functions Supported in Cath Lab:

- Clinical Documentation
- Clinical Outcomes Database (Research)
- Technical Billing
- Professional Billing
- Inventory Management
- Clinical/Quality Improvement Database
- Administrative Database Functionality
- State Reporting (DPH)

DB: Core to Supporting Multiple Functions

- Clinical Documentation
- Technical Billing
- Professional Billing
- Inventory Management
- Clinical/Quality Improvement Database
- Procedure Scheduling
- Administrative Database Functionality
- Image archiving

Tension within Medical Informatics

Database Requirements

-- structure data entry
-- limited vocabulary
-- fixed meaning
-- no free text entry
-- focus on consistency

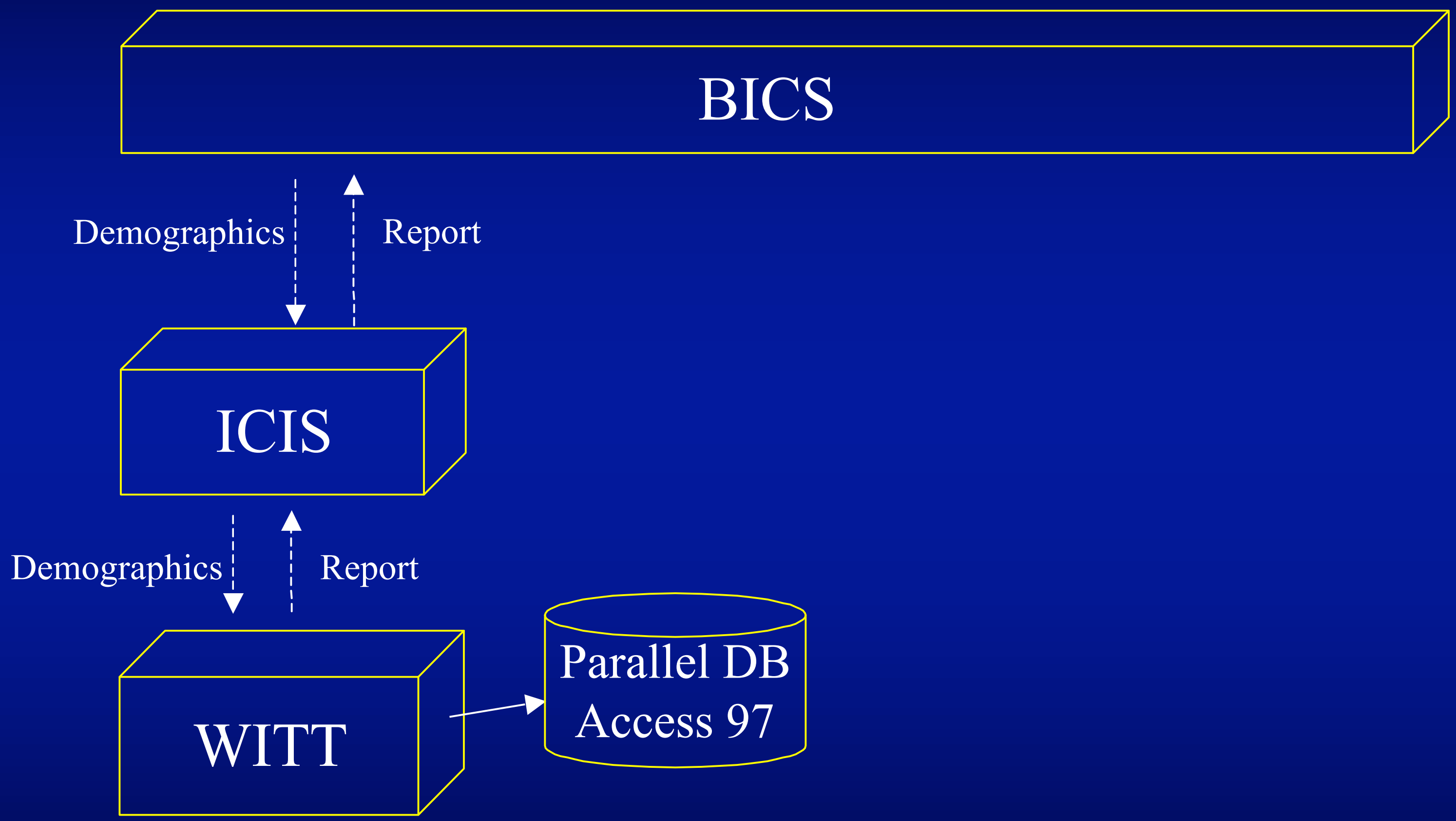
Clinical Communication Requirements

-- unstructured
-- unlimited vocabulary
-- variable meaning
-- frequent revision
-- focus on interpretation

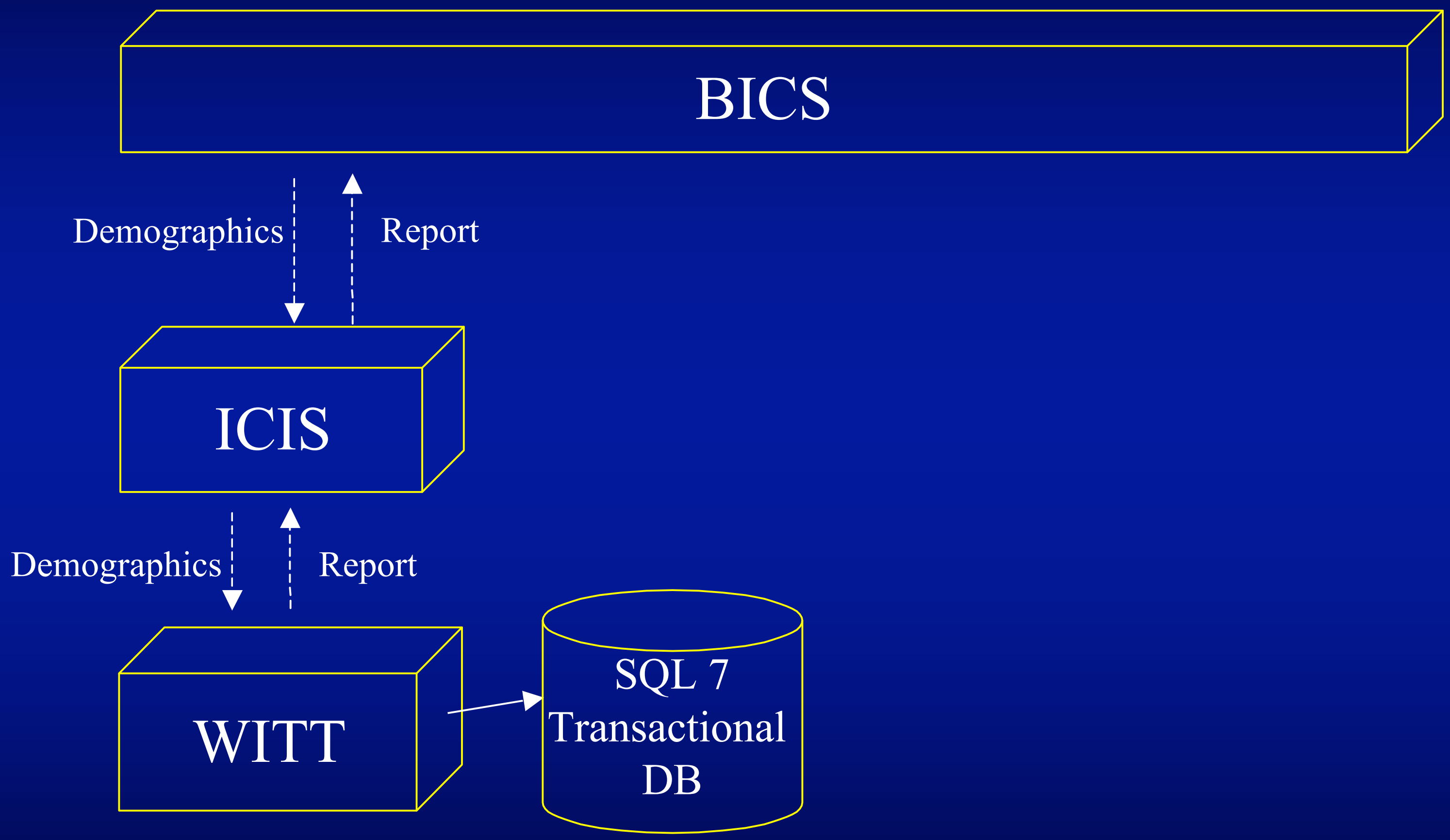
Overview

- Background: History and utility of clinical data repositories
- Strategies: Integrating the outcomes tracking database into clinical workflow
- Brigham and Women's Catheterization Laboratory Database: Architecture, Advances, Issues
- Examples of Data Exploration: Risk models, "drilling down", and device safety monitoring

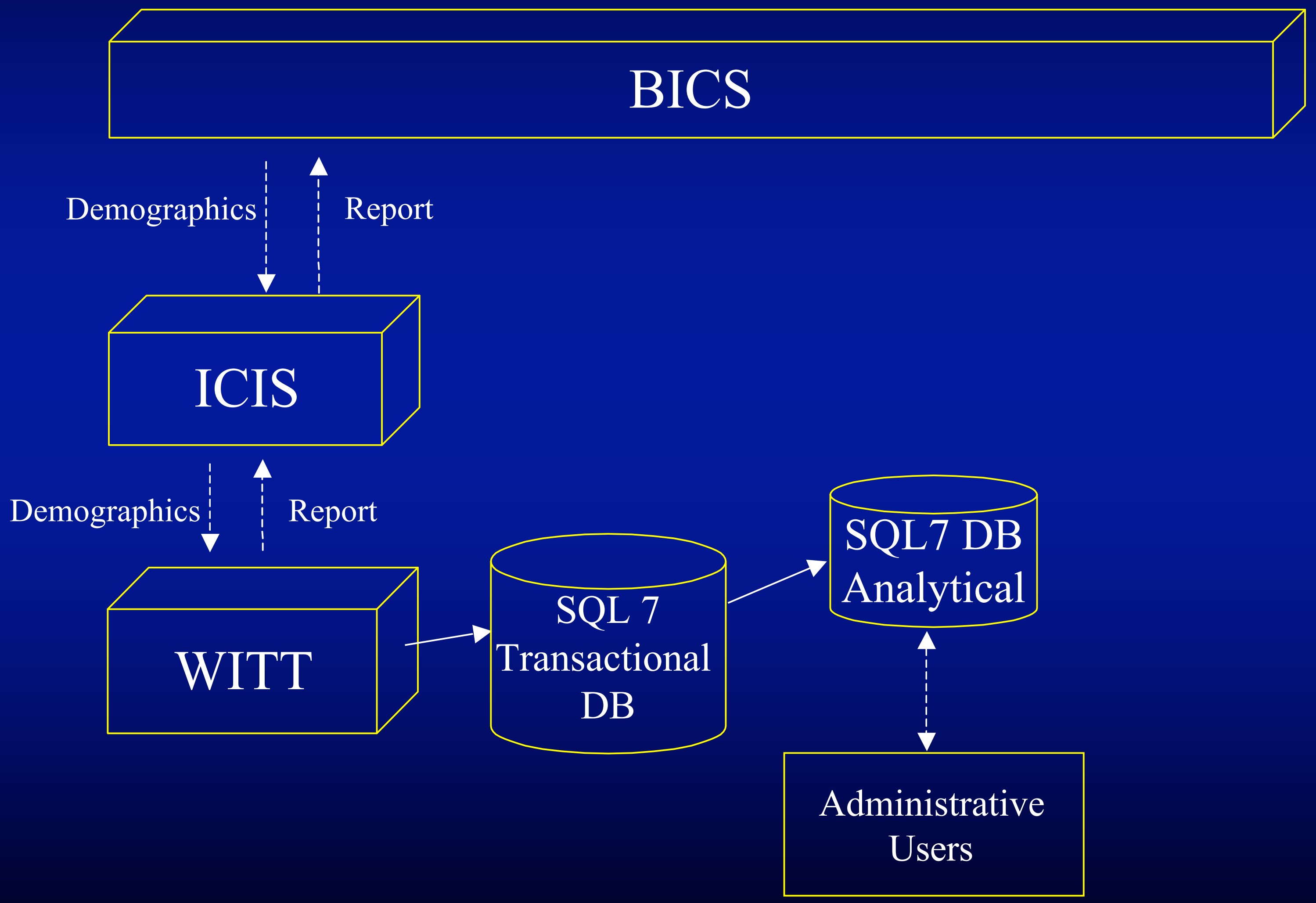
System Architecture: Phase I



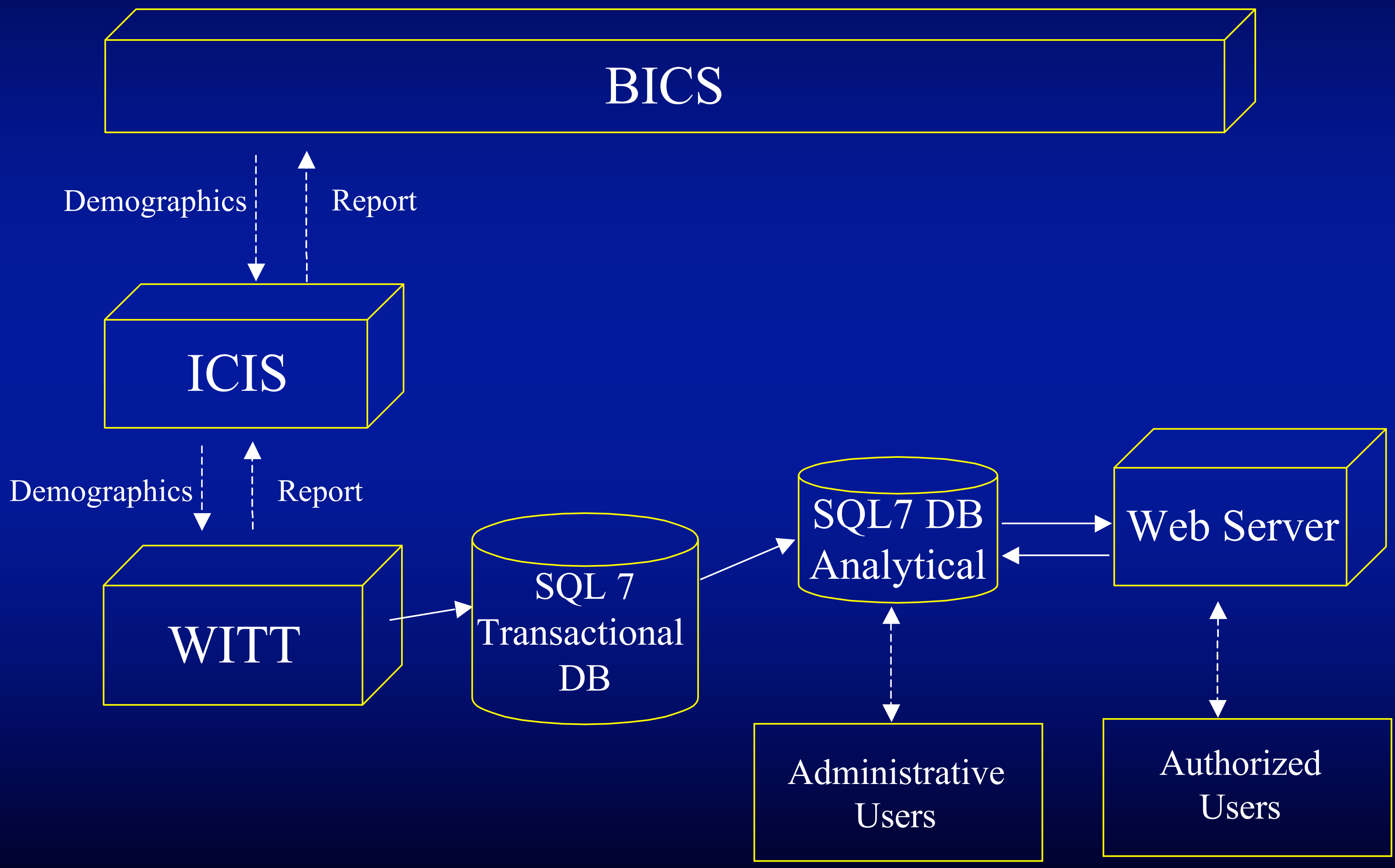
System Architecture: Phase I



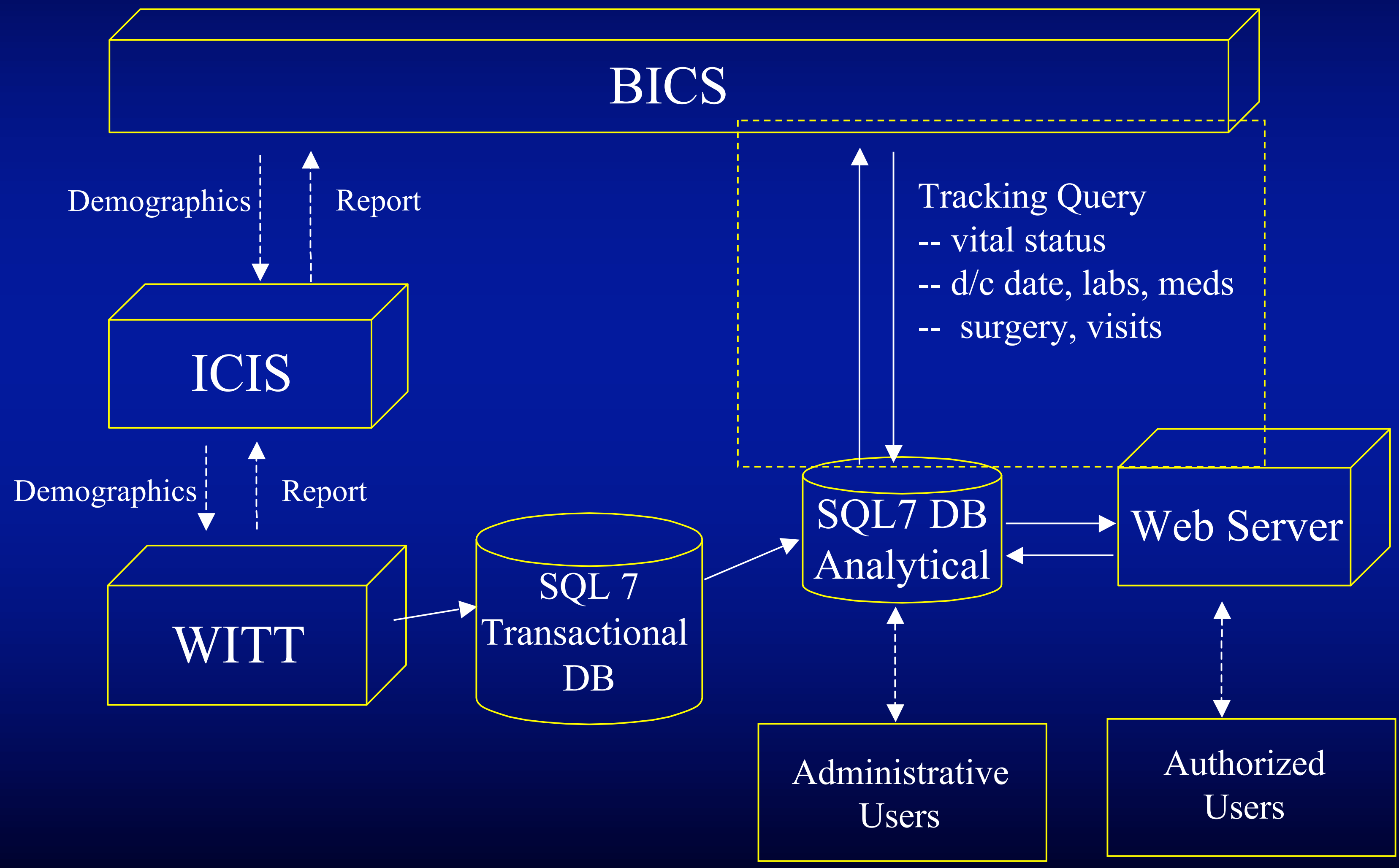
System Architecture: Phase II



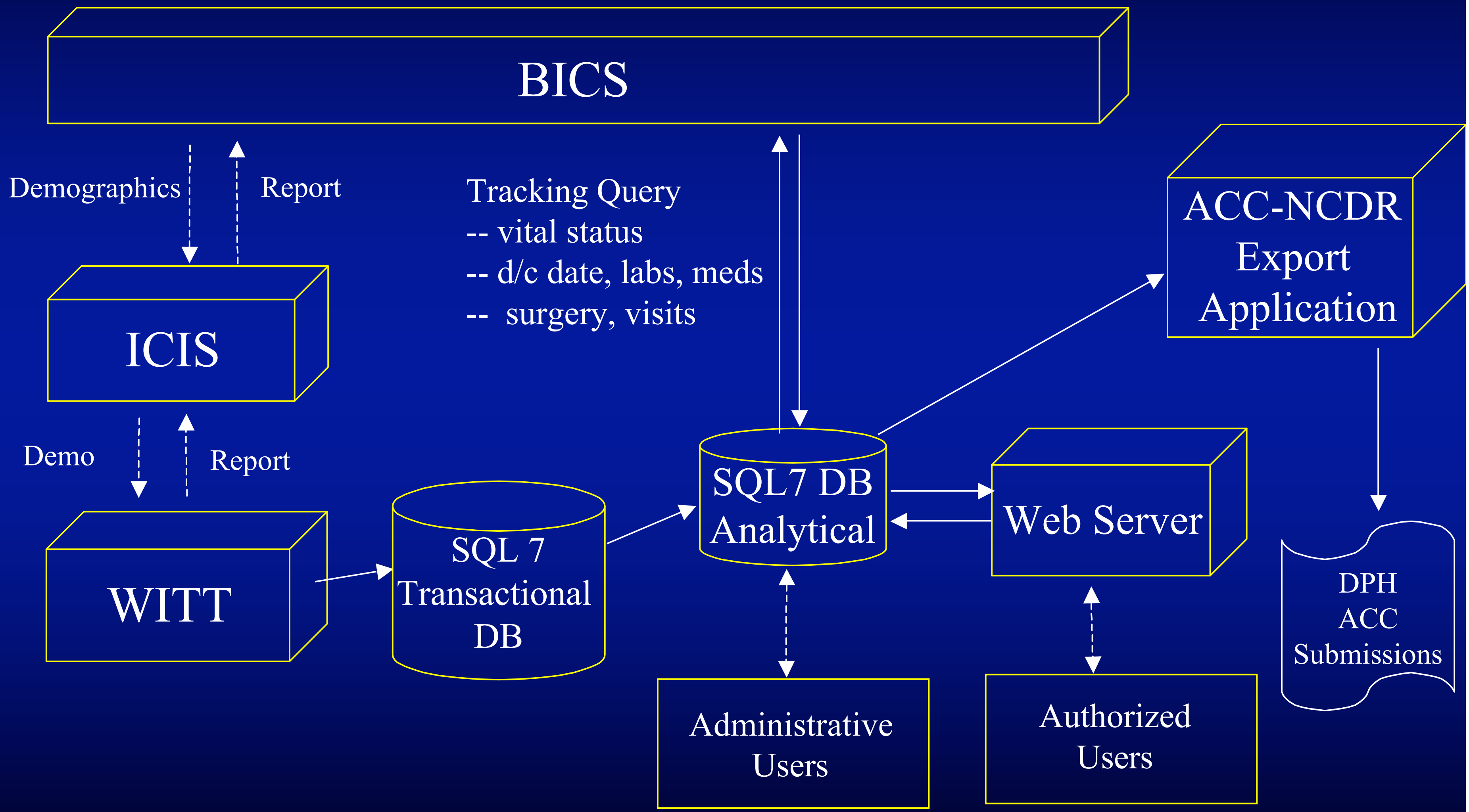
System Architecture: Phase II



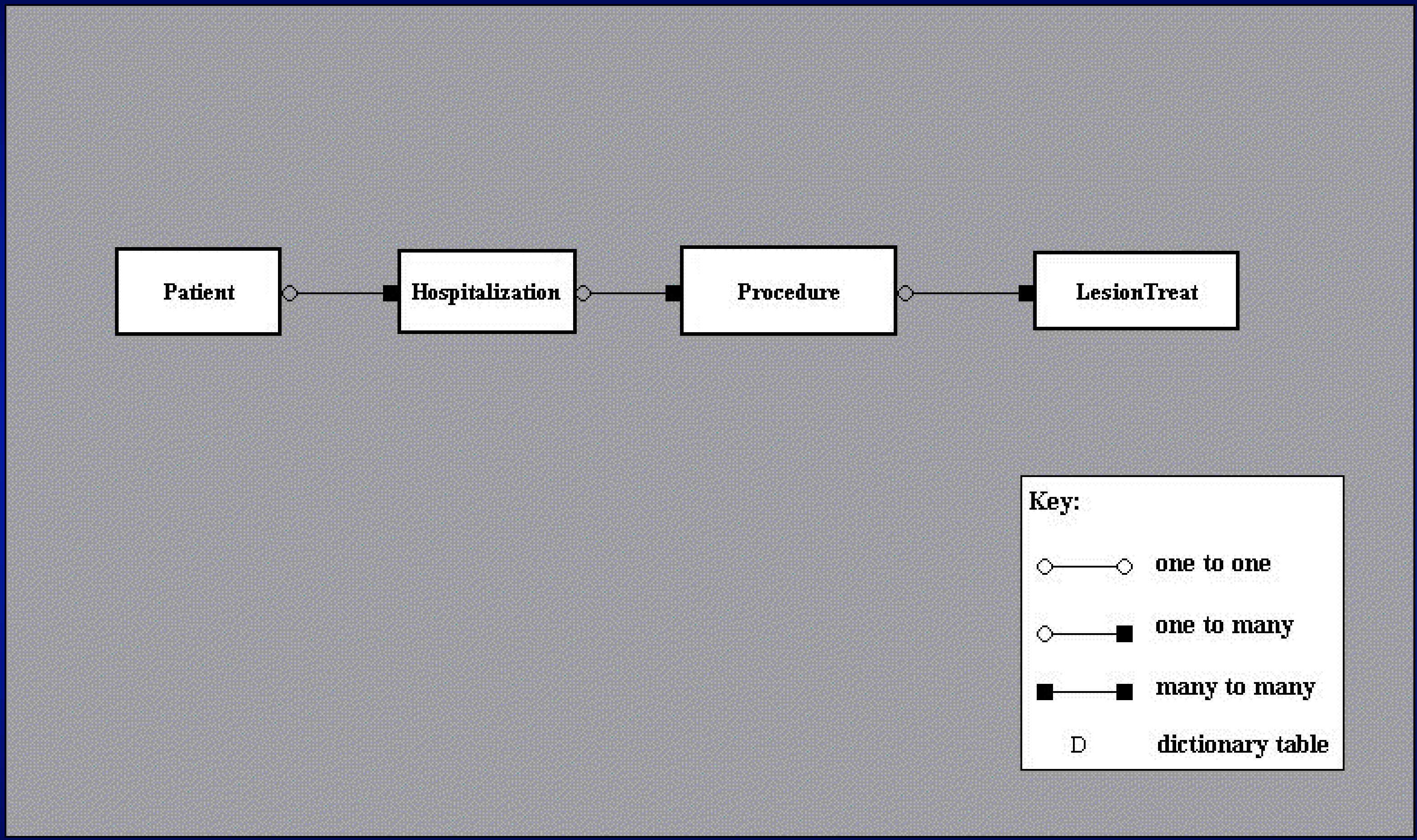
System Architecture: Phase III



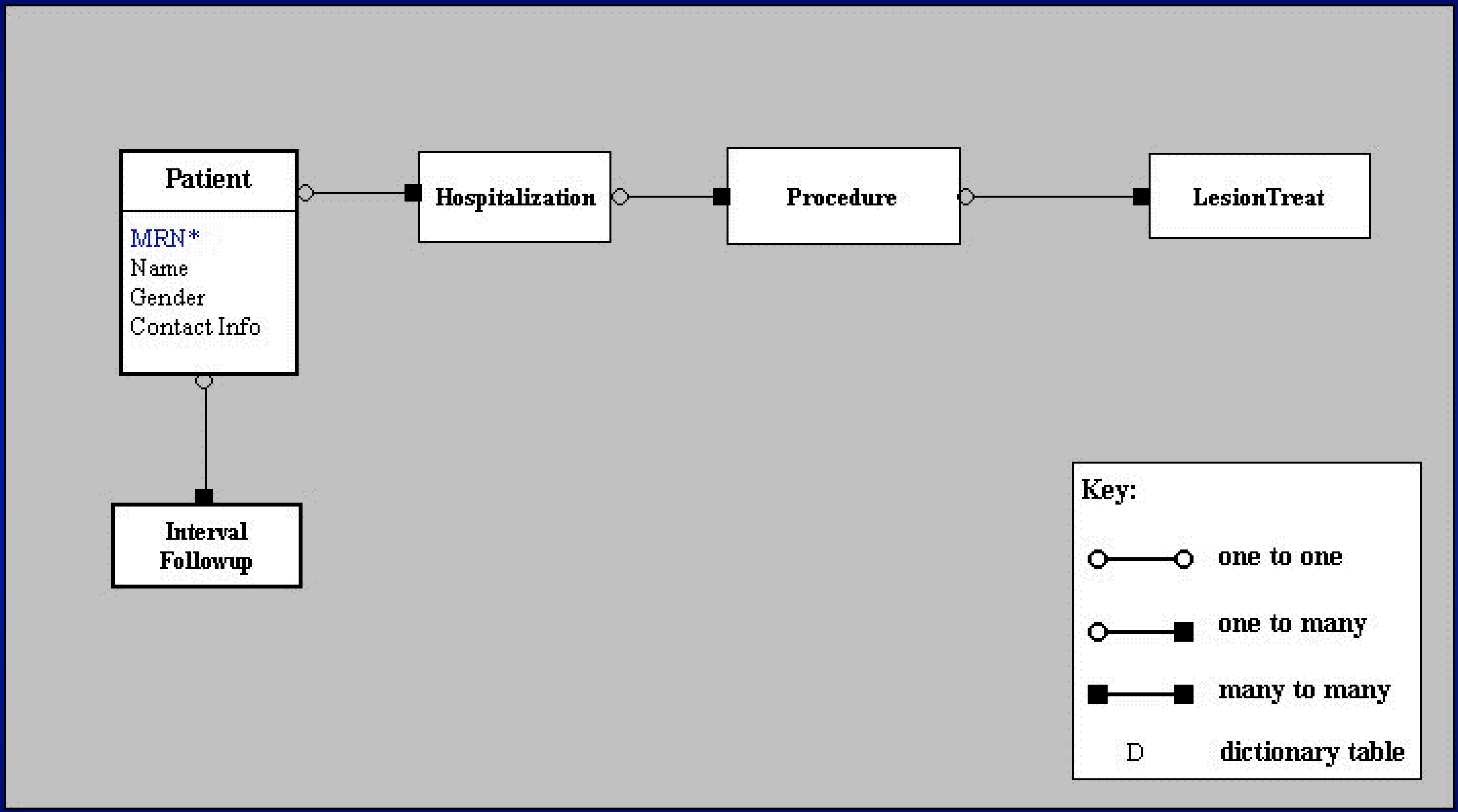
System Architecture: Phase IIIb



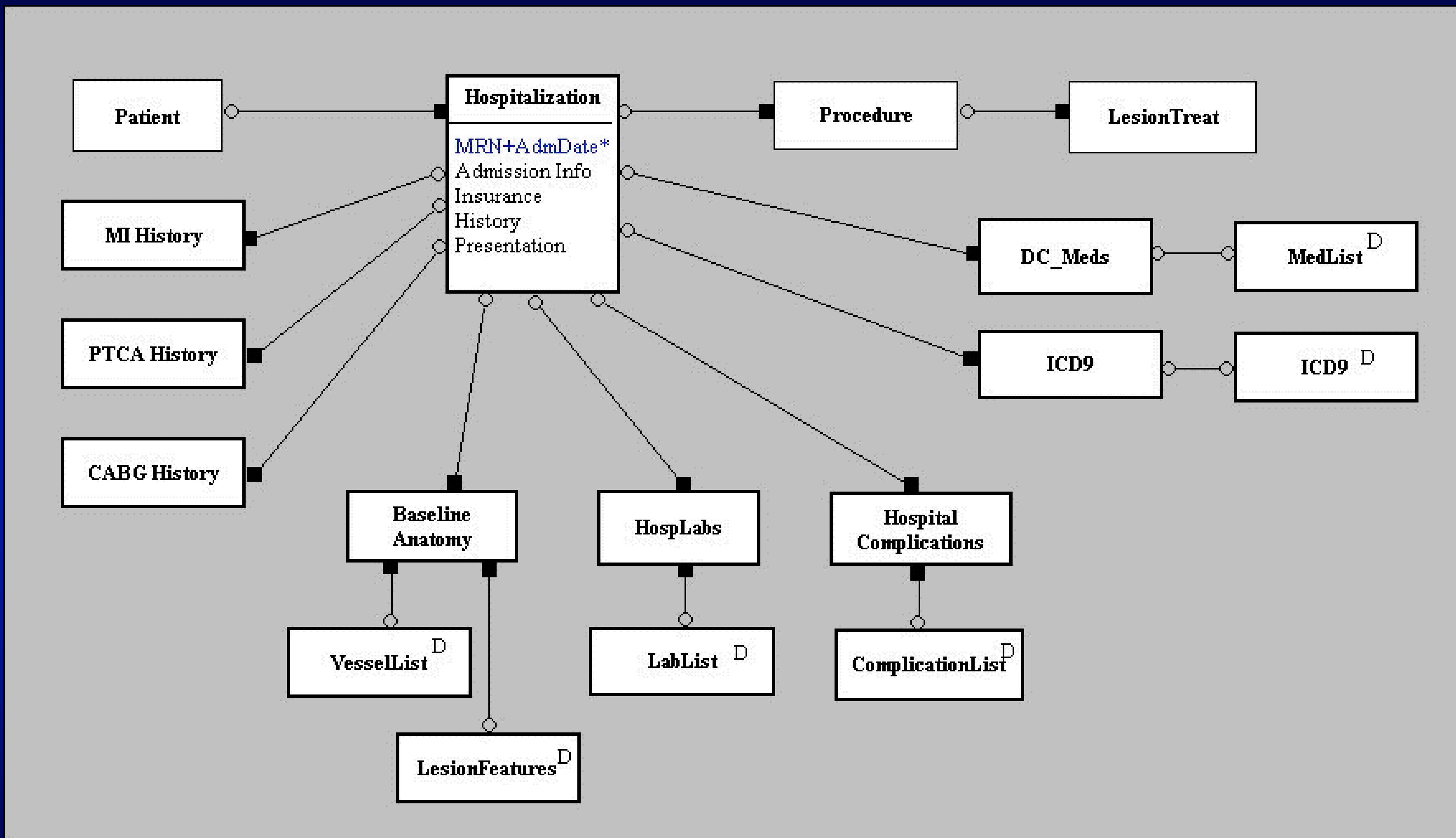
Relational DB Schema: Overview



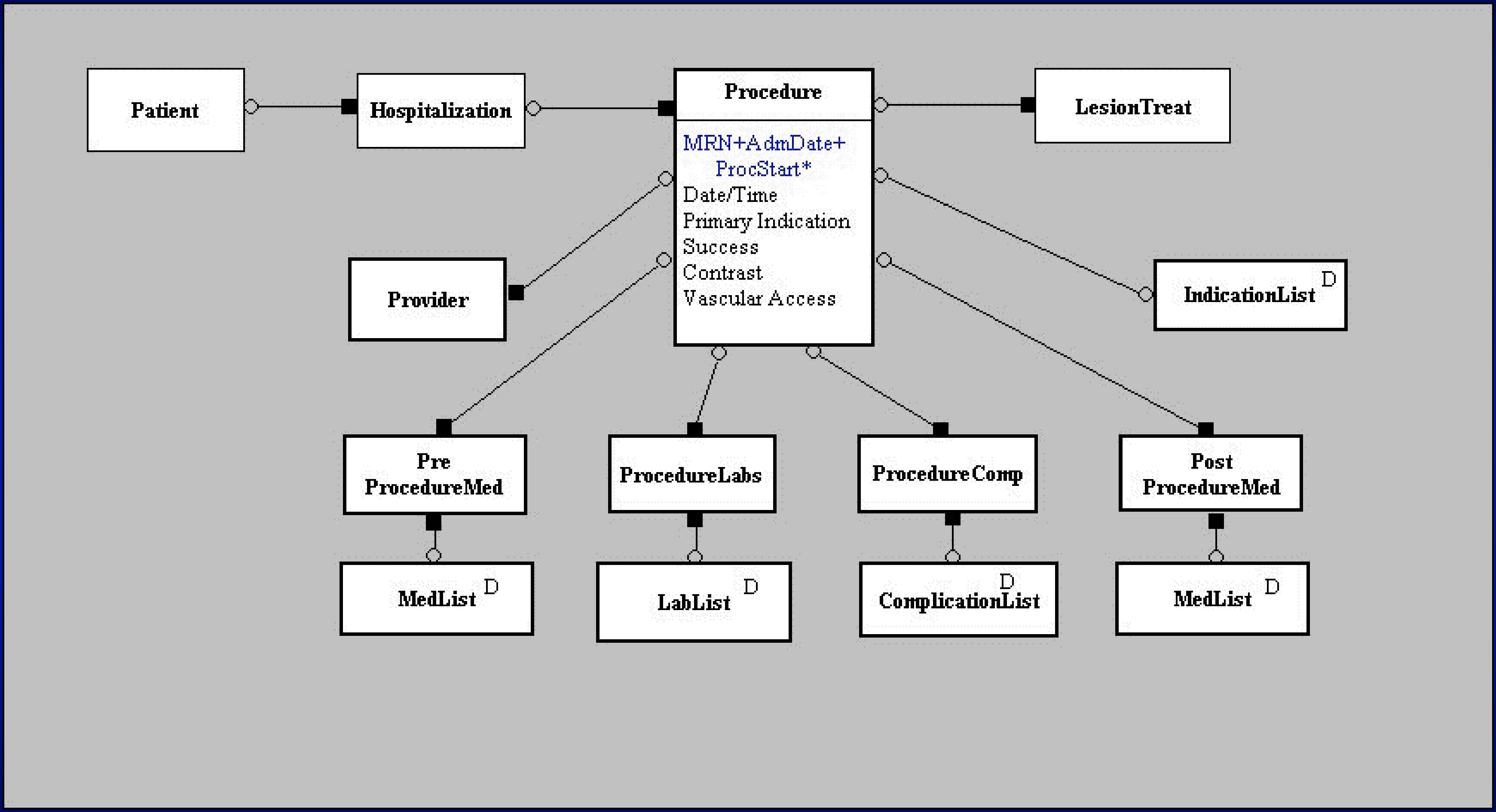
Relational DB Schema: Patient



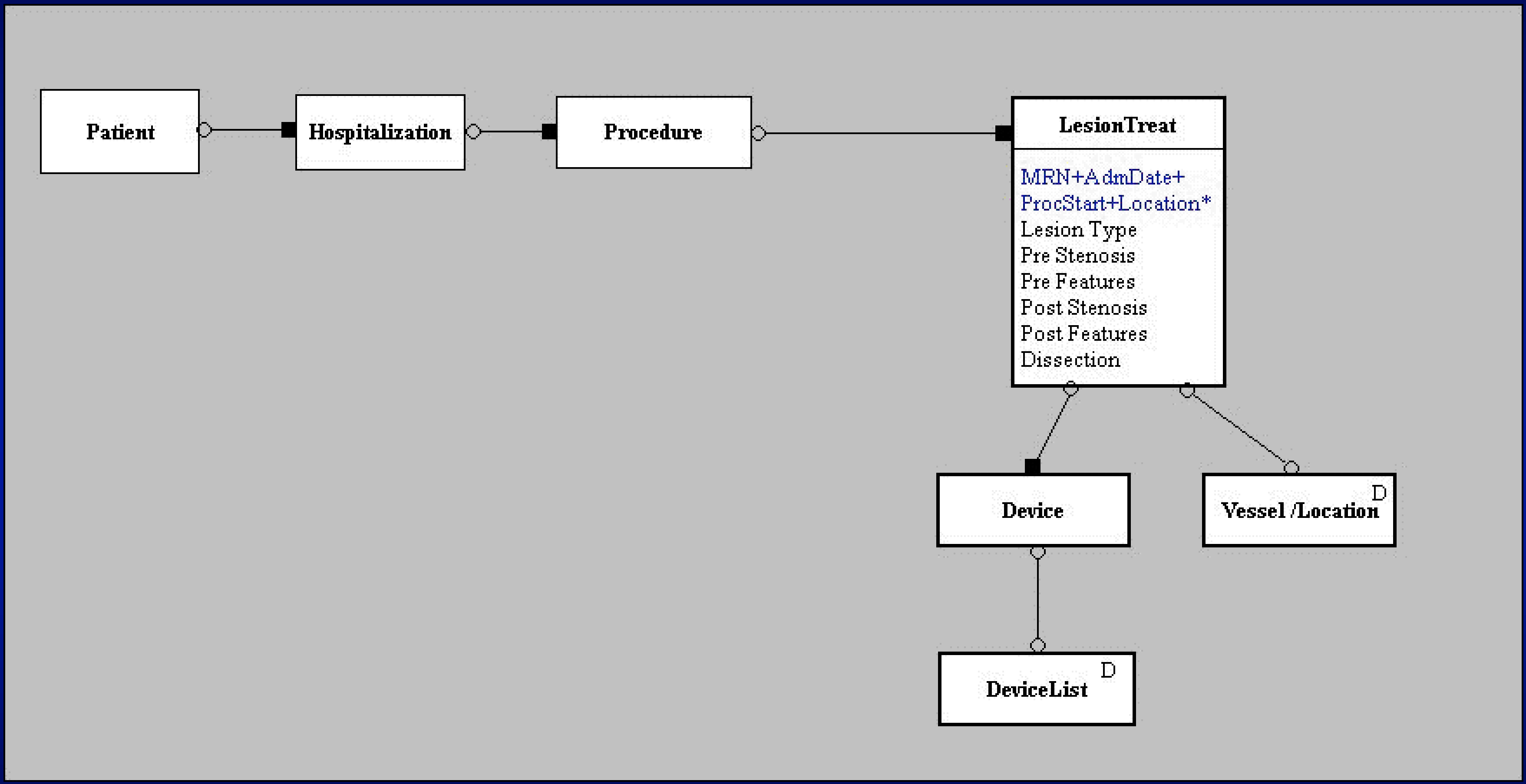
Relational DB Schema: Hospitalization



Relational DB Schema: Procedure



Relational DB Schema: Lesion Treated



Overview

- Background: History and utility of clinical data repositories
- Strategies: Integrating the outcomes tracking database into clinical workflow
- Brigham and Women's Catheterization Laboratory Database: Architecture, Advances, Issues
- Examples of Data Exploration: Risk models, "drilling down", and device safety monitoring

Applications of BWH CCL Database

Risk Prediction Model Development

Drilling Down – Novel Risk Factors

Retrospective Device Safety Analysis

Monthly QA – Cath Lab M&M

Operations Management

Risk Models: Background

- Unadjusted Mortality Rates – Published 1999-2000

		Shock
- NY State PTCA Registry Model:	0.9%	
- NNE Cooperative Model:	1.1%	0.6%
- Holmes et al (Mayo Clinic):	1.6%	
- Moscucci et al (Univ. Michigan):	3.3%	3.4%

- Risk prediction models help adjust for severity of illness
  - providers assess quality of care – improve process
  - State / public compare institutions and providers
  - researchers assess effect of changes in care

See Hannan JAMA 277(11); Holmes Circ 102:517;
Moscucci JACC 34(3); O'Conner JACC 34(3)

Logistic and Score Models for Death

Logistic Regression Model

	Odds Ratio	p-value
Age > 74yrs	2.53	0.01
B2/C Lesion	1.93	0.08
Acute MI	1.83	0.20
Class 3/4 CHF	8.14	0.00
Left main PCI	6.59	0.02
Stent Use	0.50	0.08
Cardiogenic Shock	8.33	0.00
Unstable Angina	1.69	0.17
Tachycardic	2.77	0.04
Chronic Renal Insuf.	2.71	0.05

Logistic and Score Models for Death

Logistic Regression Model

	Odds Ratio	p-value
Age > 74yrs	2.53	0.01
B2/C Lesion	1.93	0.08
Acute MI	1.83	0.20
Class 3/4 CHF	8.14	0.00
Left main PCI	6.59	0.02
Stent Use	0.50	0.08
Cardiogenic Shock	8.33	0.00
Unstable Angina	1.69	0.17
Tachycardic	2.77	0.04
Chronic Renal Insuf.	2.71	0.05

Logistic and Score Models for Death

Logistic Regression Model

	Odds Ratio	p-value
Age > 74yrs	2.53	0.01
B2/C Lesion	1.93	0.08
Acute MI	1.83	0.20
Class 3/4 CHF	8.14	0.00
Left main PCI	6.59	0.02
Stent Use	0.50	0.08
Cardiogenic Shock	8.33	0.00
Unstable Angina	1.69	0.17
Tachycardic	2.77	0.04
Chronic Renal Insuf.	2.71	0.05

Logistic and Score Models for Death

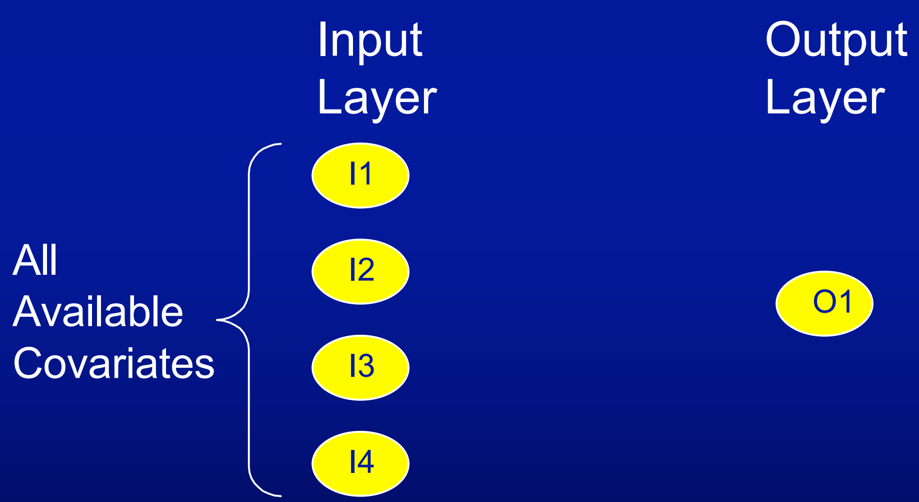
	Logistic Regression Model		Risk Score Model	
	Odds Ratio	p-value	Beta coeff	Risk value
Age > 74yrs	2.53	0.01	0.927	**2**
B2/C Lesion	1.93	0.08	0.659	**1**
Acute MI	1.83	0.20	0.601	**1**
Class 3/4 CHF	8.14	0.00	2.097	**4**
Left main PCI	6.59	0.02	1.886	**3**
Stent Use	0.50	0.08	-0.683	**-1**
Cardiogenic Shock	8.33	0.00	2.120	**4**
Unstable Angina	1.69	0.17	0.522	**1**
Tachycardic	2.77	0.04	1.020	**2**
Chronic Renal Insuf.	2.71	0.05	0.996	**2**

ROC Curves: Death Models
Validation Set: 1460 Cases

See Resnic et al. *Am J. Cardiol* 2001 Jul 1:88(1):5-9

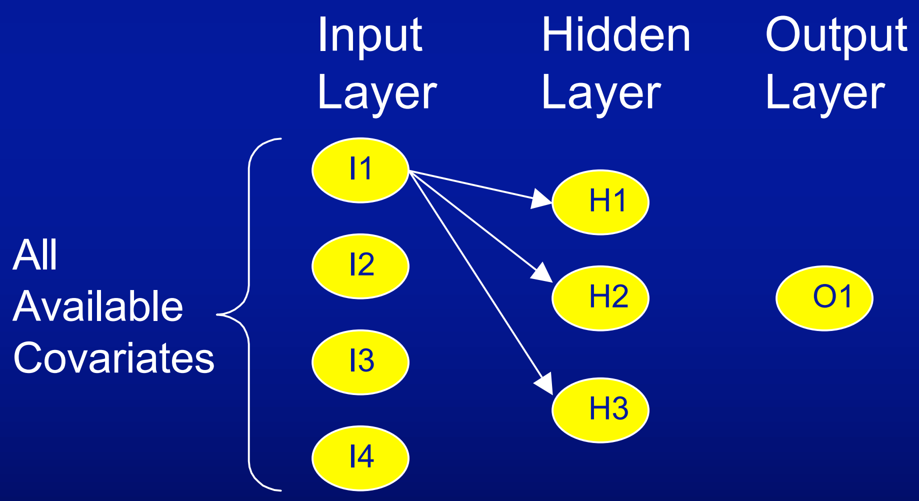
Model Building: Artificial Neural Networks

- Artificial Neural Networks are non-linear mathematical models which incorporate a layer of hidden "nodes" connected to the input layer (covariates) and the output.



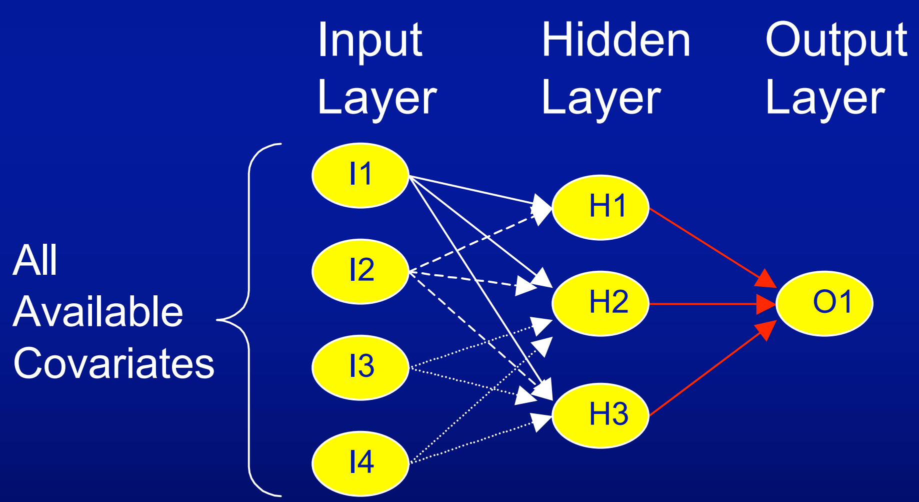
Model Building: Artificial Neural Networks

- Artificial Neural Networks are non-linear mathematical models which incorporate a layer of hidden “nodes” connected to the input layer (covariates) and the output.



Model Building: Artificial Neural Networks

- Artificial Neural Networks are non-linear mathematical models which incorporate a layer of hidden "nodes" connected to the input layer (covariates) and the output.



ROC Curves: Death Models
Validation Set: 1460 Cases

See Resnic et al. *Am J. Cardiol* 2001 Jul 1:88(1):5-9

Risk Score of Death: BWH Experience

Unadjusted Overall Mortality Rate = 2.1%

See Resnic et al. *Am J. Cardiol* 2001 Jul 1:88(1):5-9

Applications of BWH CCL Database

Risk Prediction Model Development

Drilling Down – Novel Risk Factors

Retrospective Device Safety Analysis

Monthly QA – Cath Lab M&M

Operations Management

MACE Models: Impact of No-Reflow

	Logistic Regression Model		Risk Score Model	
	Odds Ratio	p-value	beta coefficient	Risk Value
Age > 74yrs	1.40	0.16	0.337	0
B2/C Lesion	2.56	0.00	0.939	2
Acute MI	2.99	0.00	1.096	2
Class 3/4 CHF	3.61	0.00	1.283	3
Left main PCI	2.30	0.28	0.831	2
Stent Use	0.58	0.03	-0.539	-1
Cardiogenic Shock	3.33	0.01	1.202	3
USA	2.69	0.00	0.989	2
Tachycardic	1.36	0.44	0.311	0
No Reflow	2.90	0.01	1.044	2
Unscheduled	1.49	0.08	0.396	0
Chronic Renal Insuff.	1.58	0.23	0.457	1

No-Reflow: Angiographic Case Study

63yo male 4yrs s/p 4v CABG.

Presents with NQWMI W/ lateral ST depress

Posis Angiojet: Rheolytic Thrombectomy

Direct Stenting of Culprit Lesion

No Reflow: BWH Experience 1997-2000

Risk of In-Hospital Complication

See Resnic et al. *Am Heart J.* In press.

TIMI Flow Rates Improved Significantly

See Resnic et al. *Am Heart J.* In press.

Lack of Effective Treatment: BWH Experience

Risk of Death or Myocardial Infarction

See Resnic et al. ***Am Heart J.*** **In press.**

Applications of BWH CCL Database

Risk Prediction Model Development

Drilling Down – Novel Risk Factors

Retrospective Device Safety Analysis

Monthly QA – Cath Lab M&M

Operations Management

Patients receiving a closure device experienced a 44% reduction in vascular complications.

See Resnic et al. *Am J. Cardiol.* 2001 Sep 1:88(5):493-496.

This effect was preserved in those patients receiving gp 2b3a inhibitors.

See Resnic et al. ***Am J. Cardiol.*** **2001 Sep 1:88(5):493-496.**

Applications of BWH CCL Database

Risk Prediction Model Development

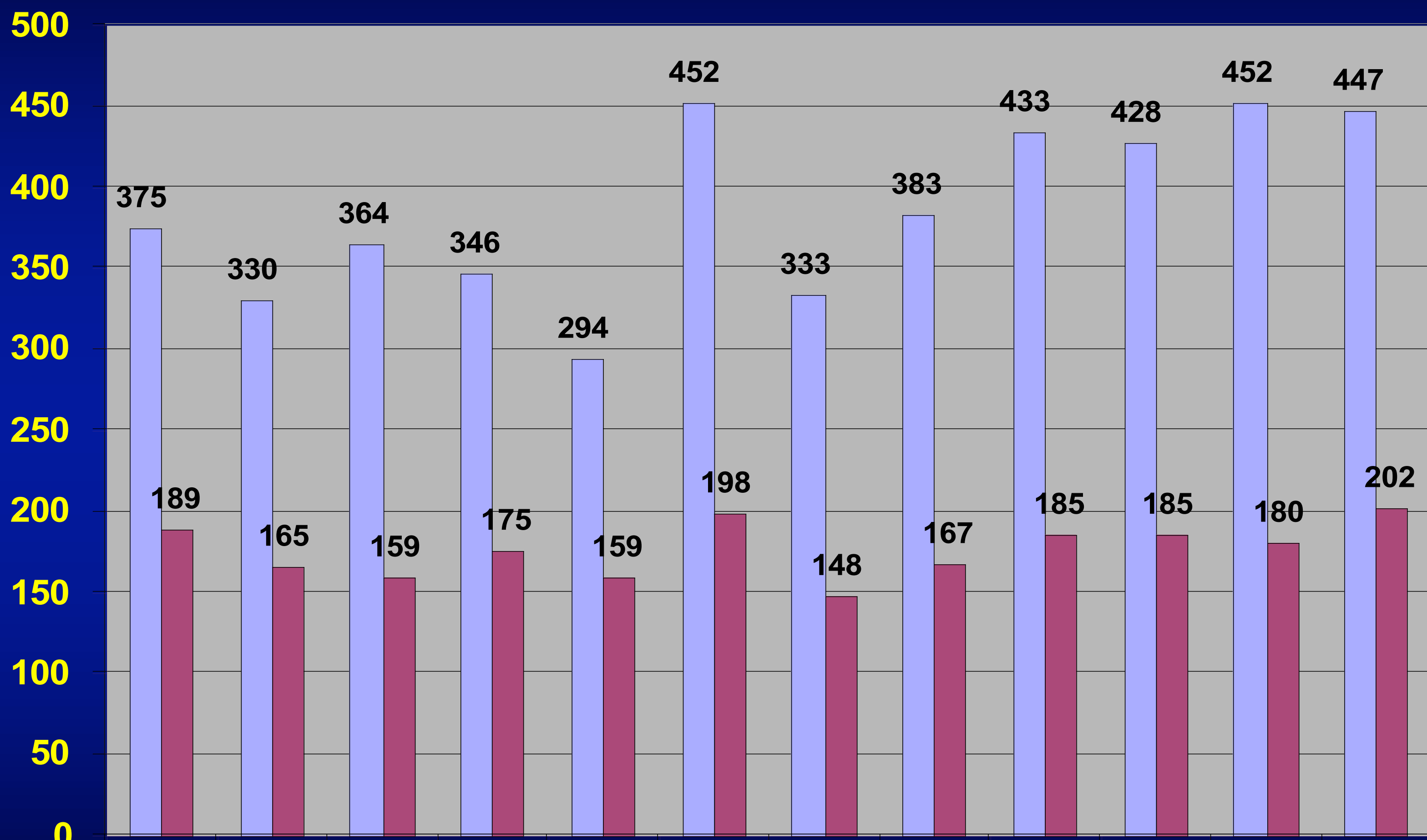
Drilling Down – Novel Risk Factors

Retrospective Device Safety Analysis

Monthly QA – Cath Lab M&M

Operations Management

Coronary Procedures by Month



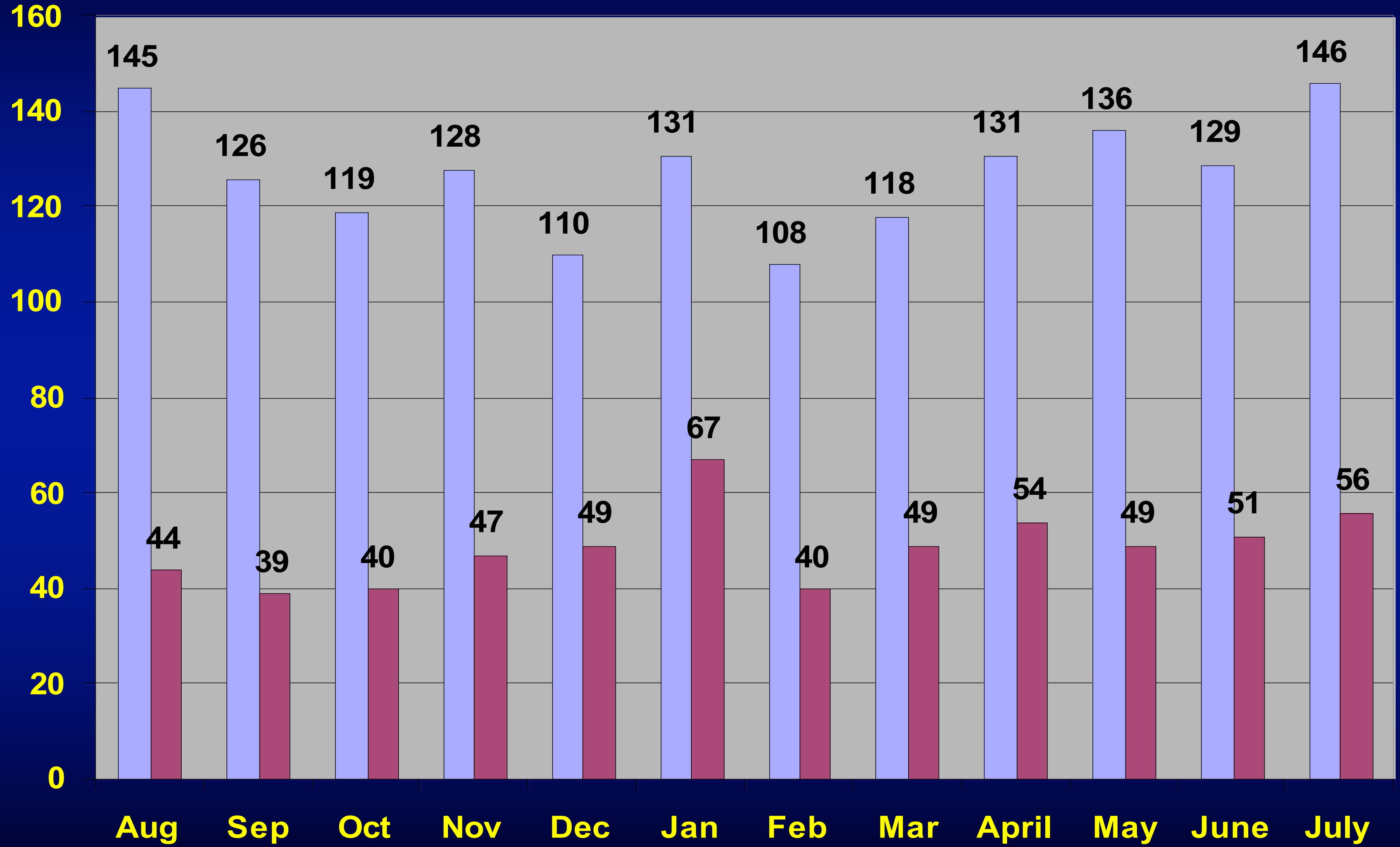
	Aug	Sept	Oct	Nov	Dec	Jan	Feb	Mar	April	May	June	July
Diagnostic	375	330	364	346	294	452	333	383	433	428	452	447
Intervention	189	165	159	175	159	198	148	167	185	185	180	202
Total Cases	384	342	371	356	296	458	337	388	437	436	458	456

Planned vs. Ad Hoc PCI

	Jan	Feb	Mar	April	May	June	July
Ad Hoc PCI	125	103	118	115	106	121	132
Planned PCI	70	78	52	67	78	50	66

Ad Hoc PCI rate	33%	40%	36%	31%	30%	30%	35%

Internal vs. External MD Volume



Internal	77%	76%	75%	73%	69%	66%	73%	71%	71%	74%	72%	72%
External	23%	24%	25%	27%	31%	34%	27%	29%	29%	26%	28%	28%

Post-Procedural Events for July, 2002

•Significant events *reported* during July, 2002:

Death	**3**	
Stroke	**1**	
CABG	**1**	**(perforation of LCx)**
MI*	**6**	**(1 SAT)**
TVR	**2**	
Vascular	**7**	**(1 transfusion reported, 1 PSA req. surg)**
Renal	**3**	
CHF	**1**	

*** MI defined as total CK-MB>3x ULN in patient w/o index AMI.**

Clinical Event Listing by Physician 2001

Applications of BWH CCL Database

Risk Prediction Model Development

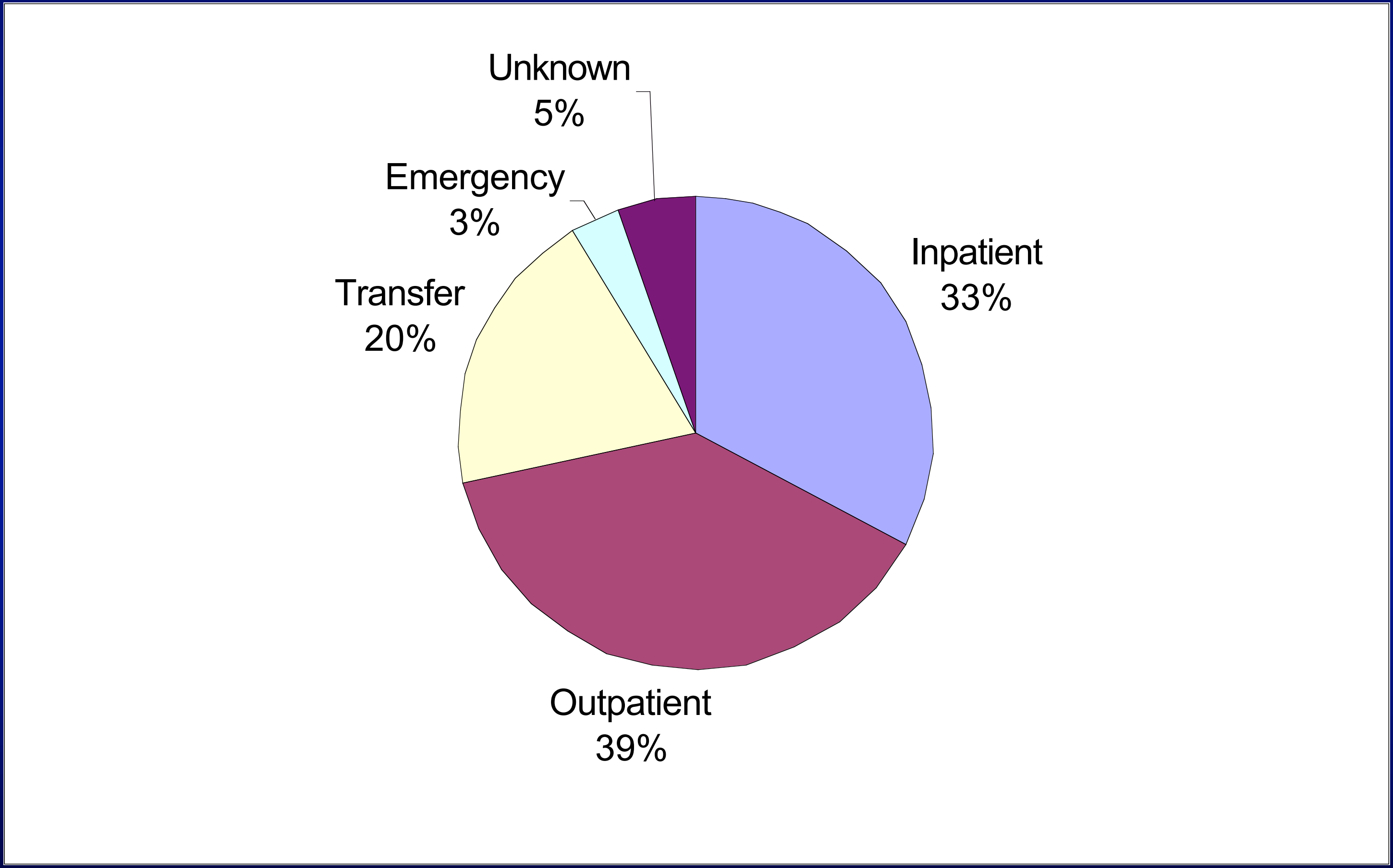
Drilling Down – Novel Risk Factors

Retrospective Device Safety Analysis

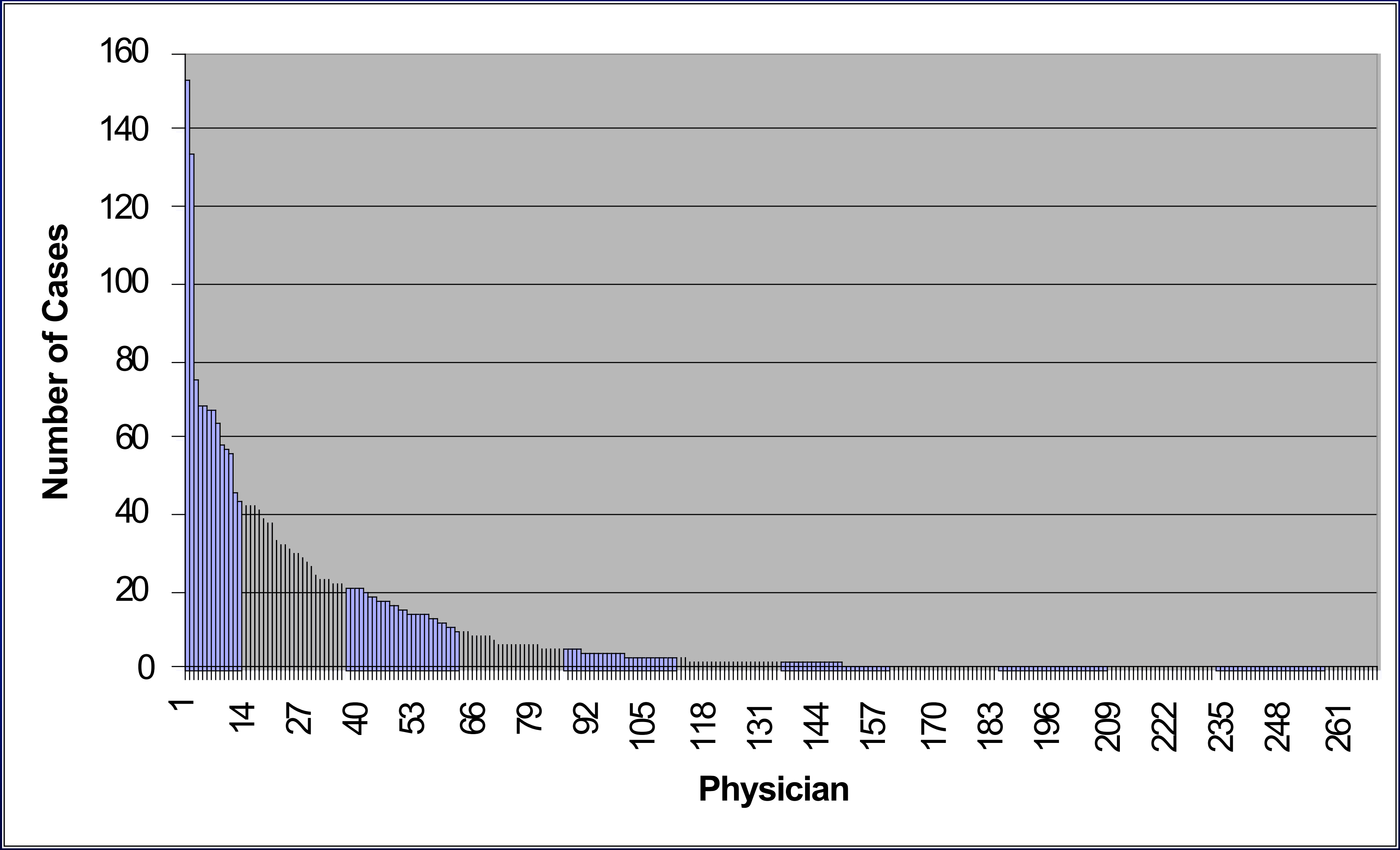
Monthly QA – Cath Lab M&M

Operations Management

One-third of total case volume is noted as inpatient source in WITT.



The case volume is distributed according the usual 80/20 rule. Nearly 80% of cases is referred from 20% of the MDs.



Rules for Designing an Outcomes Database

- Understand workflow in detail. Identify immutable points (most of these depend on perspective).
- Incremental design – identify successful milestones
- Open architecture – use ODBC compliant relational databases as backbone

  Systems integration is most complex challenge

- Goal of distributed information availability.
- Identify implementation team. Responsiblities, project plan, regular operational meeting.

Acknowledgements

Cardiovascular Division

Jeff Popma, MD

Andrew Selwyn, MD

Campbell Rogers, MD

Charles Lin, MBA

Benjamin Paul

Decision Systems Group

Lucila Ohno-Machado, MD PhD

Robert Greenes, MD PhD

Aziz Boxwala, MD PhD

Partners Information Systems

Mark Nightingale

Thank You!