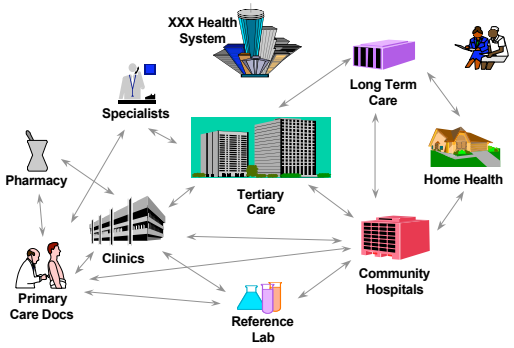
Health Enterprise Computing and Patient Identification

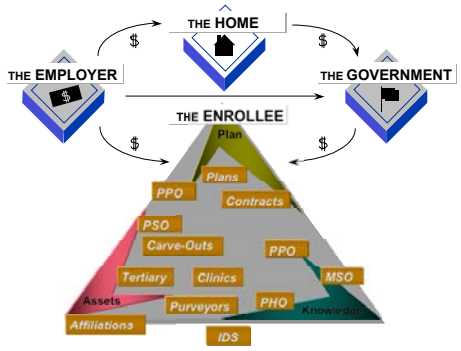
6.872/HST.950

Peter Szolovits
(some content from David Margulies)

Integrated Delivery System...



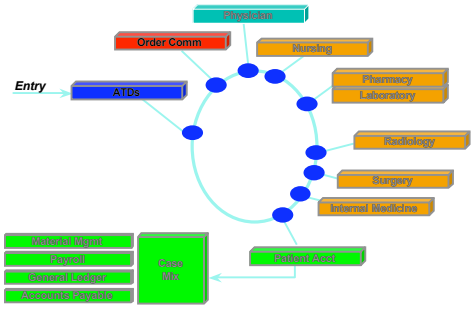
Complexity now within...



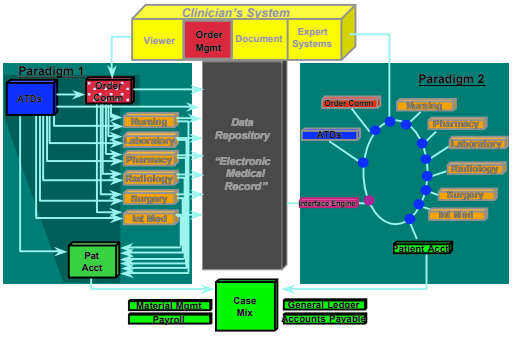
Context: Evolution of Clinical Computing...

- Focus evolved from automating AR process in hospitals, to LAN-based SI of departmental systems, to enterprise clinical computing, toward region- and system-level process automation
- Changing healthcare organizational models and economic models ---> changing clinical computing priorities and designs

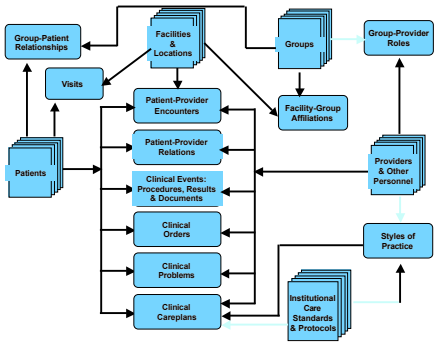
Phase 2: LAN-based SI -- Automate Data Distribution to MDs...



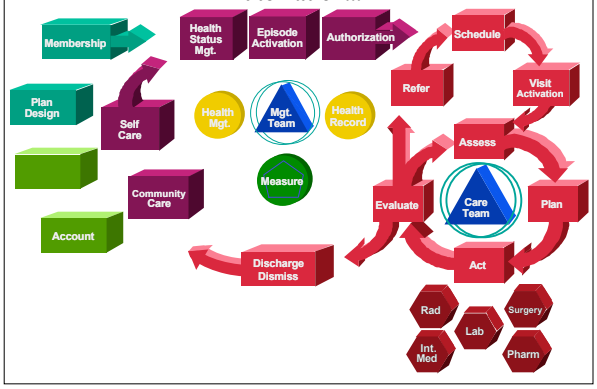
Phase 3: Data-based SI to Create an EMR...



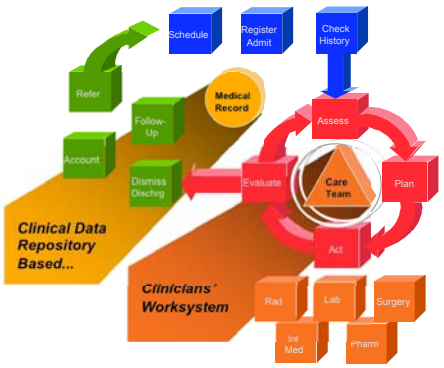
DATA REPOSITORY



Phase 4: Enterprise-level Clinical Process Automation...



Most Installations...



Mayo
Genesys
SOM
EMMC
HMW
…...
Samaritan

Different Imperatives, Different IT Strategies

Computer-Based Patient Record

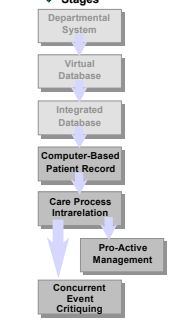
- World-Wide
- Community-Wide
- Multiple Facilities

Care Team Support

- Critiqueing
- Scheduling
- Common Clinical Ordering

- Electronic Chart Viewer
- Clinical Data Repository
- MPI/ADT Synchronization
- Independent Departmental Systems

Evolutionary Stages



Stages	Characteristics	Examples
	Access Directly Via Program	PathNet Laboratory Information System
	• highly structured • powerful	
	Workstation Access Via Network • distribution data • disparate databases and user interface	HL7 Query Tools
	Copy Data into Central Depository • relationships possible • common user interface • multimedia	ORACLE
	Control Access and Encode • nomenclature support • security and audit control • data succession • date authenticity	Open Clinical Foundation Data Repository
	Diagnoses, Care Plans and Orders • context for results • expectations possible • requirements for records	ProNet Orders CareNet Plans MRNet Operations
	Operations Data • care management • fiscal management • EIS	Open Management Foundation Data Repository
	Normative Data and Rules • extrapolations • management by protocol	Discern Rules APACHE MEDIQUAL

Beyond the Enterprise: CHINs Valid needs, No Buyer...

Network Applications
- *E-Mail*

Database Application
- *Membership*
- *Eligibility*
- *Health Record Access*
- *Outcomes Measurement*

Transaction Applications
- *Scheduling/Referrals*
- *Electronic Signature*
- *Claims Processing*

Clinical Systems
- *Case Management*
- *AHP Business Systems*

Providers

Utility

External Organizations

Core Services
- *E-Mail*

Database Systems
- *Membership Server*
- *Health Plan Server*
- *Health Record Index*
- *Outcomes Management*

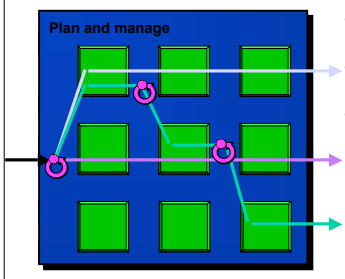
Transaction Systems
- *Scheduling Services*
- *Report Management*
- *Claims Consolidation*
- *Utilization Review*

Business Systems
- *Client Billing*
- *Financial Accounting*

Community Health Information Needs

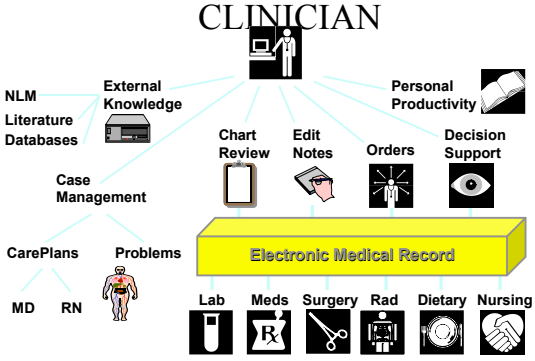
- Person Identification Server
- Membership Server
- Plan/Eligibility Server
- Health Record Server
- Resource Scheduling Server
- Care Protocol Server
- Commerce Server

Care Maps

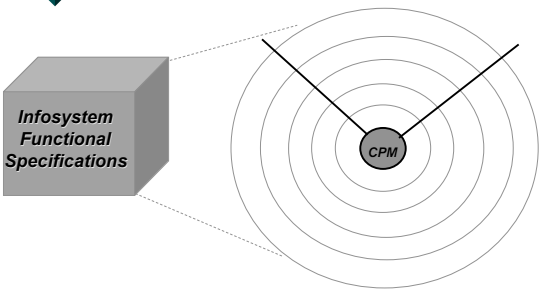


- Defined and tested clinical methods
- Specific paths through map individualized
- Multiple opportunities for triage and adaptation
  - by planned choice
  - by risk
  - by availability
  - by length of trajectory
  - by cost
  - by new information

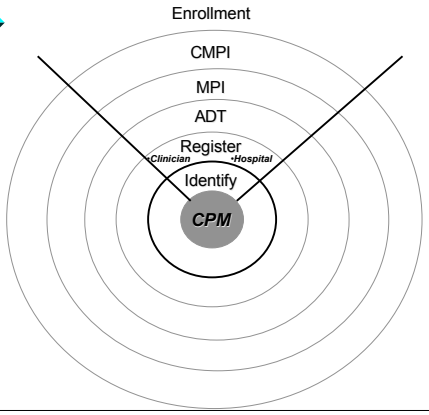
THE EYE OF THE CLINICIAN



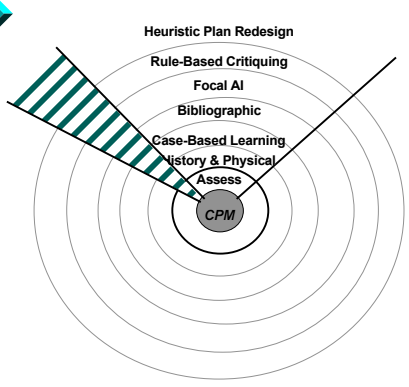
"EMERGING FROM THE COMMON PROCESS MODEL"

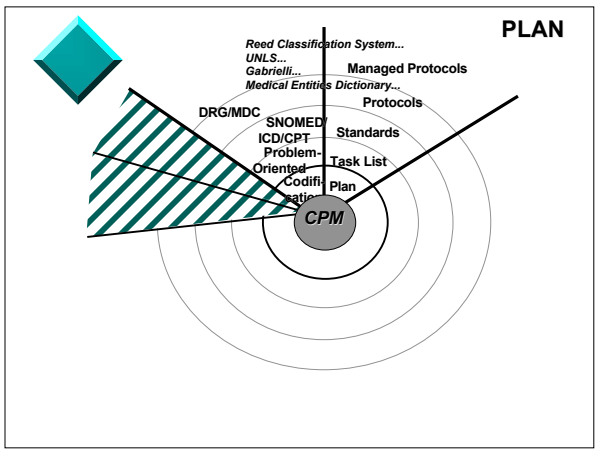


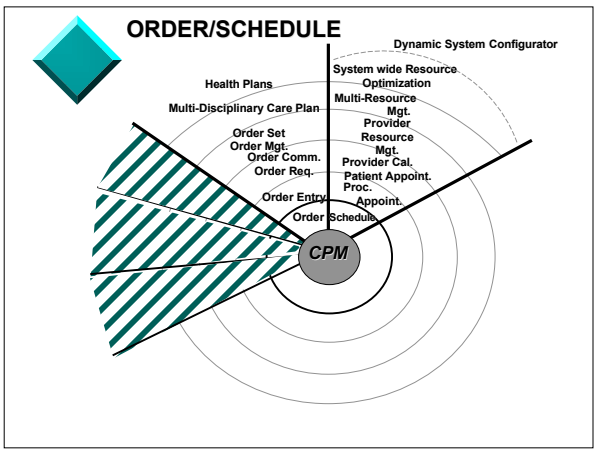
IDENTIFY

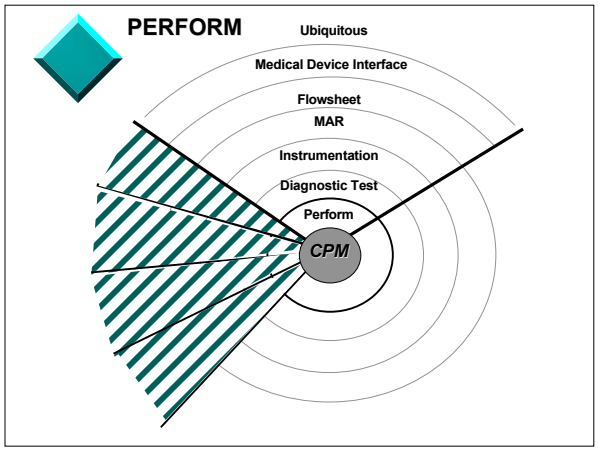


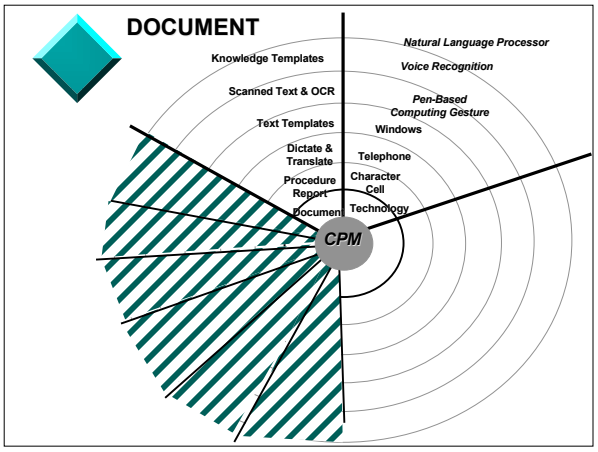
ASSESS

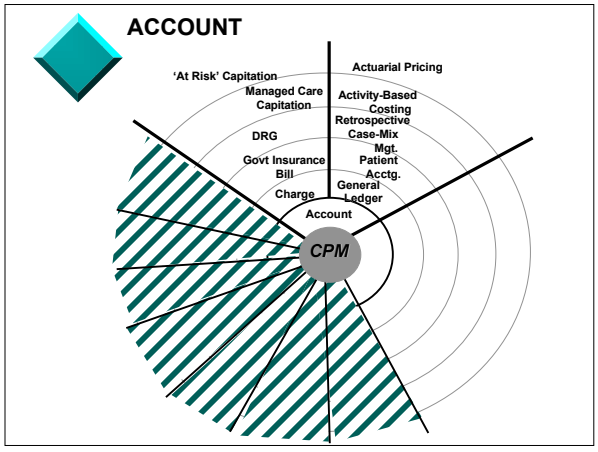


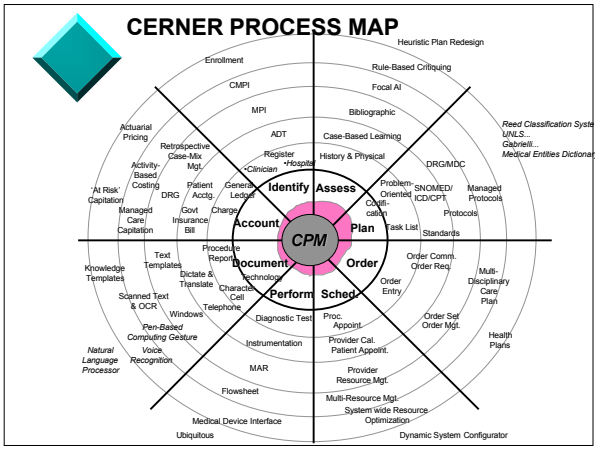

PLAN
Reed Classification System...
UNLS...
Gabrielli...
Medical Entities Dictionary...
Managed Protocols
Protocols
DRG/MDC
SNOMED/
ICD/CPT
Standards
Problem-
Oriented
Task List
Codifi-
Plan
CPM

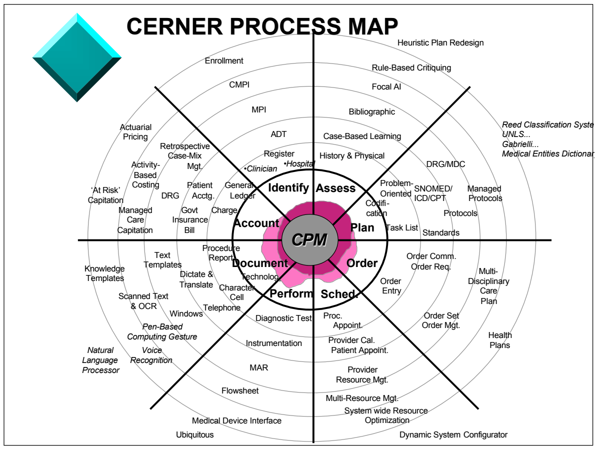

ORDER/SCHEDULE
Dynamic System Configurator
System wide Resource
Optimization
Health Plans
Multi-Resource
Mgt.
Multi-Disciplinary Care Plan
Provider
Resource
Mgt.
Order Set
Order Mgt.
Provider Cal.
Order Comm.
Order Req.
Patient Appoint.
Proc.
Order Entry
Appoint.
Order
Schedule
CPM

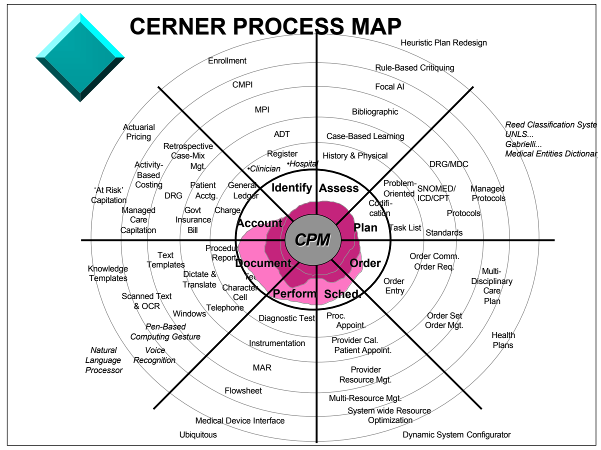

PERFORM
Ubiquitous
Medical Device Interface
Flowsheet
MAR
Instrumentation
Diagnostic Test
Perform
CPM

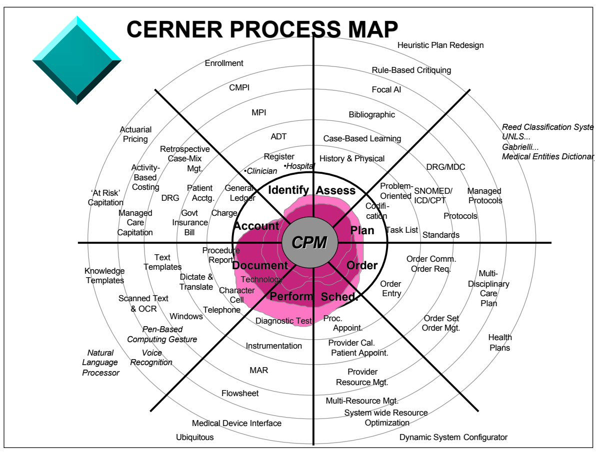

DOCUMENT
Natural Language Processor
Knowledge Templates
Voice Recognition
Scanned Text & OCR
Pen-Based
Computing Gesture
Text Templates
Windows
Dictate &
Translate
Telephone
Procedure
Report
Character
Cell
Technology
CPM

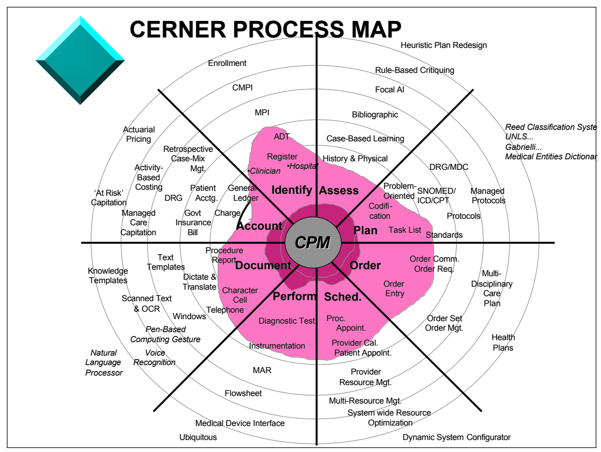

ACCOUNT
Actuarial Pricing
'At Risk' Capitation
Managed Care
Capitation
Activity-Based
Costing
Retrospective
Case-Mix
DRG
Mgt.
Govt Insurance
Bill
Patient
Acctg.
General
Charge
Ledger
Account
CPM

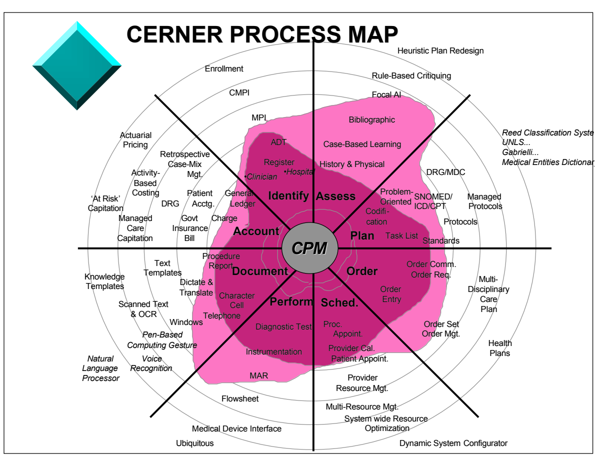

CERNER PROCESS MAP
Heuristic Plan Redesign
Enrollment
Rule-Based Critiquing
CMPI
Focal AI
MPI
Bibliographic
Actuarial
Pricing
ADT
Case-Based Learning
Retrospective
Case-Mix
Mgt.
Register
Hospital
History & Physical
Clinician
Activity-
Based
Costing
DRG/MDC
'At Risk'
Capitation
DRG
Patient
Acctg.
General
Ledger
Identify
Assess
Problem-
Oriented
SNOMED/
ICD/CPT
Managed
Protocols
Codifi-
cation
Managed
Care/
Capitation
Govt
Insurance/
Bill
Charge
Protocols
Account
Plan
Task List
CPM
Standards
Procedure
Report
Text
Templates
Order Comm.
Order Req.
Knowledge
Templates
Document
Order
Dictate &
Translate
Technology
Order
Entry
Multi-
Disciplinary
Care
Plan
Character
Cell
Scanned Text
& OCR
Perform
Sched
Telephone
Windows
Diagnostic Test
Proc.
Appoint.
Order Set
Order Mgt.
Pen-Based
Computing Gesture
Health
Plans
Instrumentation
Provider Cal.
Patient Appoint.
Natural
Language
Processor
Voice
Recognition
MAR
Provider
Resource Mgt.
Flowsheet
Multi-Resource Mgt.
System wide Resource
Optimization
Medical Device Interface
Ubiquitous
Dynamic System Configurator

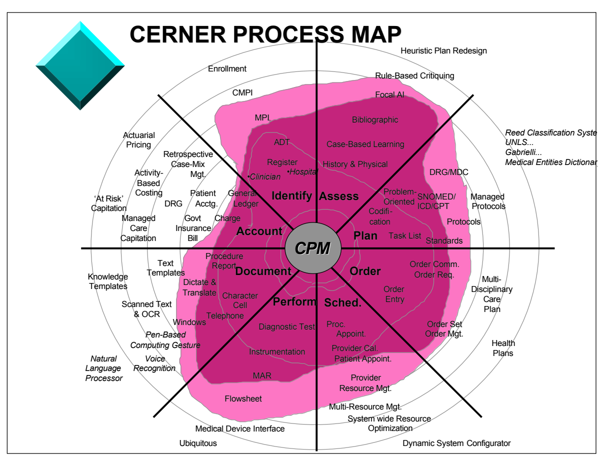

CERNER PROCESS MAP
CPM
Identify
Assess
Plan
Order
Sched.
Perform
Document
Account
Heuristic Plan Redesign
Rule-Based Critiquing
Focal AI
Bibliographic
Case-Based Learning
History & Physical
Reed Classification System
UNLS...
Gabrielli...
Medical Entities Dictionary
DRG/MDC
Problem-Oriented
SNOMED/ ICD/CPT
Managed Protocols
Codifi-cation
Protocols
Task List
Standards
Order Comm. Order Req.
Order Entry
Multi-Disciplinary Care Plan
Order Set Order Mgt.
Health Plans
Proc. Appoint.
Provider Cal. Patient Appoint.
Provider Resource Mgt.
Multi-Resource Mgt.
System wide Resource Optimization
Dynamic System Configurator
Diagnostic Test
Instrumentation
MAR
Flowsheet
Medical Device Interface
Ubiquitous
Natural Language Processor
Voice Recognition
Pen-Based Computing Gesture
Windows
Telephone
Scanned Text & OCR
Character Cell
Dictate & Translate
Technology
Text Templates
Knowledge Templates
Procedure Report
'At Risk' Capitation
Managed Care Capitation
Govt Insurance Bill
Charge
DRG
Patient Acctg.
General Ledger
Activity-Based Costing
Retrospective Case-Mix Mgt.
Actuarial Pricing
Enrollment
CMPI
MPI
ADT
Register
•Clinician
•Hospital

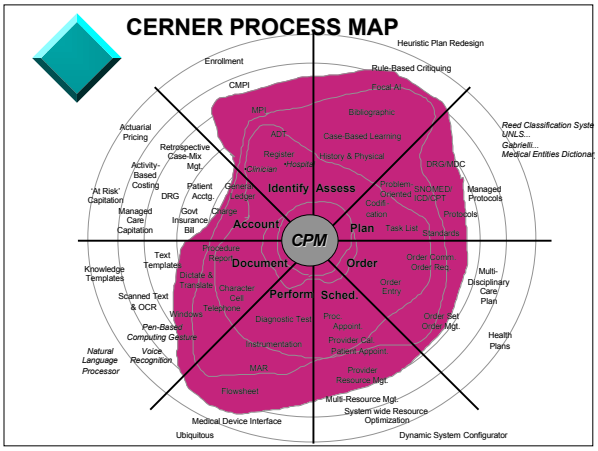

CERNER PROCESS MAP

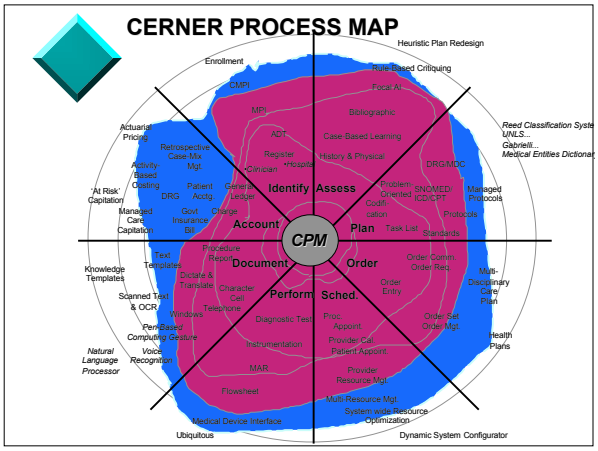

CERNER PROCESS MAP


CERNER PROCESS MAP


CERNER PROCESS MAP


CERNER PROCESS MAP





Summary...

- Diverse organizational models
- Healthcare IT emphasis moving from enterprise to health system focus
- Increased emphasis on patient's role and MD risk
- An open, net-based commerce model fits in all organizational models

Definition...

- ***Definition:***
  - ***"Net-based clinical commerce"*** is the ability for patients, physicians, and other buyers of healthcare products and services to arrange, track, and pay for services, purchase healthcare commodity products from an electronic exchange, and manage health records on the Internet.

Context: Evolution of Commerce on the Internet...

- Commerce-enabled Web
  - banking
  - retail
- Extensive use in healthcare
  - research
  - communication
- Toward Clinical Commerce!!

Healthcare Commerce Entry Barriers

- Consumers purchase health *processes*
- Processes are ongoing with multiple points of contact with the provider
- Payment is by assignment of entitlement
- Payer exerts significant control over choice
- Payer exerts significant control over care
- Some goods and services require license

Patient Identification

Patient Identification

- Example: Partners Health Care EMPI (Enterprise Master Patient Index)
  - Lisa Adragna, J AHIMA 1998 Oct;69(9):46-8, 50, 52
- Probabilistic Models for Matching

Partners EMPI Goals

- (Context) Partners Healthcare is affiliation of Brigham and Women's, Mass. General, North-Shore Medical Center, Dana-Farber Cancer Institute, Spaulding Rehab, McLean, Partners Community Healthcare, …, formed in 1994.
- Eliminate need to renumber existing systems by creating Partners-wide unique ID
- Foundation for Clinical Data Repository
- Minimally invasive to current operations
- Rapid installation

Partners EMPI Scope

- Medical record "crosswalk" across Partners
  - Local medical record numbers linked to Partners-wide ID
- Back-end, real-time, one-way interface
  - No modifications to existing systems
- Automated tools for identifying and resolving duplicates
  - No human intervention at "run time"

Partners EMPI Matching

- Three methods of lookup:
  - Last and First names
  - First name, Date of Birth and Gender
  - Social Security Number
- If no matches, create new EMPI number
- If exact match on all known features (above), then link to existing EMPI entry
- If multiple possible matches, create new entry and flag as potential duplicate; *manual intervention*

Partners EMPI Experience (1997)

- 1M records entered from BWH & MGH
- 170,000 pair of records identified as potential duplicates
- 80% of new registrations match a unique EMPI entry; ~1800/day
- Sites made 3308 "within-entity" merges
- EMPI Admin staff (3 FTE) resolved over 200,000 potential duplicates.

Models of Patient Matching

- *Ad hoc* methods are weak, and difficult to improve
- Motivational observation:
  - *Two records that share the last name Kowalczyk are more likely to refer to the same individual than two that share the name Smith.*

Computing posterior odds of a match

- $O(match|obs)=O(match)L(obs|match)$
- $O(match)=1/N$
- $L(obs|match)=P(obs|match)/P(obs|\sim match)$
- P(Smith,Smith|*match*) = P(Smith)
- P(Smith,Smith|~*match*) = P(Smith)^2
- L(Smith,Smith)=.01006/.01006^2=99.4
- L(Kowalczyk,K…)=1/.00001=100,000

Multiple evidence accumulates

- Smith has same Soundex code as Smyth
- L partially matches Liz
- Same street address number, street name
- Ave. vs. Blvd.
- Different phone numbers
- …
- Assume conditional independence

Soundex

- Developed in 1800's to represent phonetic similarity of names
  - E.g., Smith ➔S-530; Smyth ➔S-530
- Rules:
  - 1 B, F, P, V
  - 2 C, G, J, K, Q, S, X, Z
  - 3 D, T
  - 4 L
  - 5 M, N
  - 6 R
  - Disregard the letters A, E, I, O, U, H, W, and Y.
  - Special rules for doubled consonants, prefixes, …

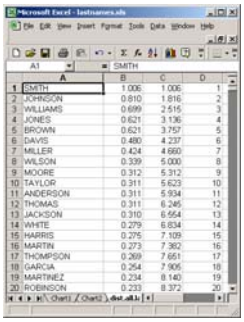
Soundex examples

Szolovits	S-413
Beethoven	B-315
Smyth	S-530
Kowalczyk	K-422

Typing/Transcription Error Model

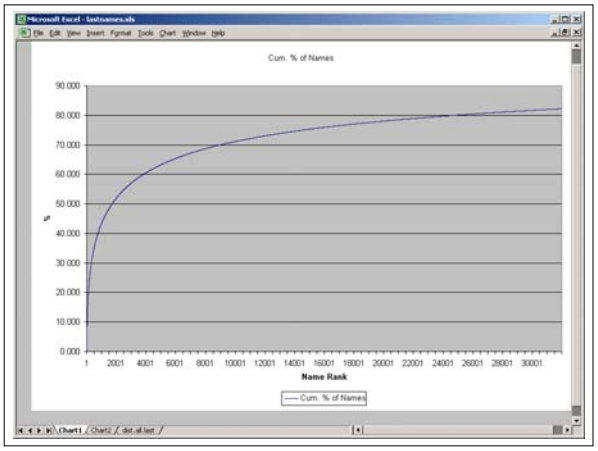
- Consider probabilities of
  - Insertion
  - Deletion
  - Transposition
  - (Mutation) – varies with key layout
- Compute "distance" between mismatched features
- *Same idea in **BLAST***

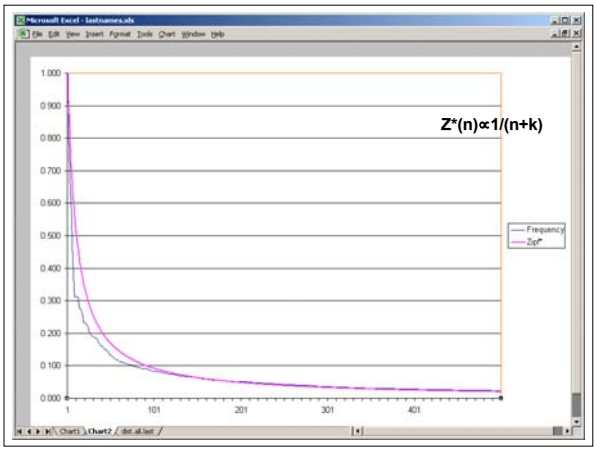
Frequency of U.S. Last Names



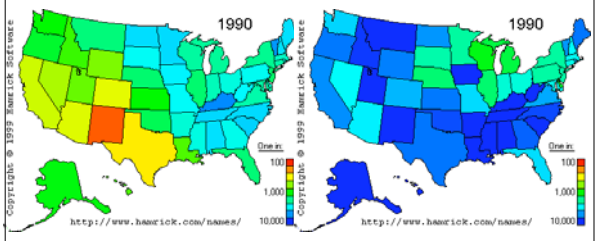
	A	B	C	D
1	SMITH	1.006	1.006	1
2	JOHNSON	0.810	1.816	2
3	WILLIAMS	0.699	2.515	3
4	JONES	0.621	3.136	4
5	BROWN	0.621	3.757	5
6	DAVIS	0.480	4.237	6
7	MILLER	0.424	4.660	7
8	WILSON	0.339	5.000	8
9	MOORE	0.312	5.312	9
10	TAYLOR	0.311	5.623	10
11	ANDERSON	0.311	5.934	11
12	THOMAS	0.311	6.245	12
13	JACKSON	0.310	6.554	13
14	WHITE	0.279	6.834	14
15	HARRIS	0.275	7.109	15
16	MARTIN	0.273	7.382	16
17	THOMPSON	0.269	7.651	17
18	GARCIA	0.254	7.906	18
19	MARTINEZ	0.234	8.140	19
20	ROBINSON	0.233	8.372	20

--U.S. Census Bureau





Geographical Distribution of "Gonzales" vs. "Lewandowski"



A Theory for Record Linkage

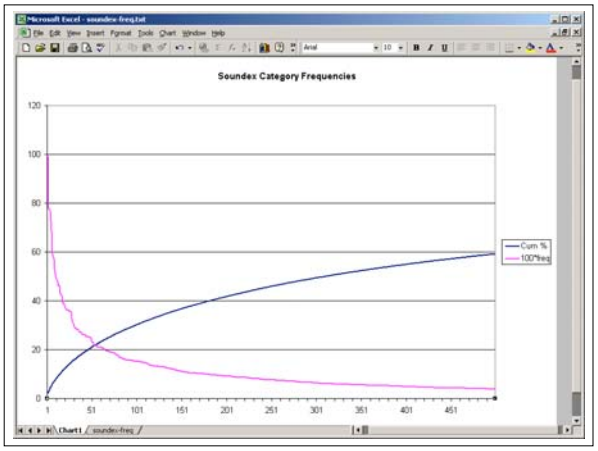
- Fellegi IP, Sunter AB. *J. Am. Stat. Assn.* **64**(328):1183-1210. 1969
- Estimate L(*obs*|*match*) for all possible observations
  - Normally must make independence assumption to factor contributions of different aspects of *obs*
- Define *decision policy*
  - P(A+|*obs*), P(A?|*obs*), P(A-|*obs*), sum=1
  - Based on thresholds on L's
- Yields errors
  - P(A+|~*match*) or P(A-|*match*), minimize P(A?)

Blocking Sets

- Pair-by-pair comparison is expensive
  - E.g., 1M records at Partners → 10^12 comparisons
- Therefore, compare only smaller subgroups
  - E.g., assume first letter of name is never misspelled, as Soundex does
  - E.g., assume Soundex pronunciation code is correct, even if spelling is not
- Increased chance of missing matches, much faster

4234 Soundex Classes /5616 possible

Smith, Schmidt, Schmitt, Sneed, Schmid, Smoot, Snead, Shumate, Smyth, Sandy, Sand, Santo, Santoyo, ...	S-530	1.118%
Johnson, Jenkins, Jensen, Jamison, Jameson, Jansen, Johansen, Janssen, …	J-525	0.998%
Brown, Bryan, Barron, Braun, Byrne, Bruno, Browne, Baron, Barney, Burney…	B-650	0.779%
Jones, James, Johns, Jung, Janes, Jaynes, Jonas, Jenks, Jaimes, Jinks, Jahnke, Johannes, Janik, Janis, Jiang, …	J-520	0.775%



Themes

- Growing levels of integration in health care
- Demand for integration of health IT systems raises challenging problems
- MPI is fundamental for record keeping
- Matching of records critical
  - Good theory, mixed practice
- Flip side: privacy concerns