

Clinical Concepts for Clinical Research: Terminology, Models & Information

Christopher G. Chute, M.D., Dr.P.H.
Professor and Head,
Medical Informatics
Mayo Clinic
Rochester, MN USA

Of Word, Terms, and Conceptual Things

- Role of Terminology in Clinical Practice
- Concepts and Categories Beyond Practice
- The Nature of Terminologies
- Integration of Terminologies in EMRs
- Notion of Terminology Services
- Example of Problem List Application
- Terminologies I Have Known

Health Care Is An Information Intensive Industry

- Control of Health Care Costs ...
- Improved Quality of Care ...
- Improved Health Outcomes ...
- Appropriate Use of Health Technology...
- Compassionate Resource Management...
- ... *depend upon information*
- ... *Ultimately Patient Data*

Aggregate Data Uses

Core of Improvement and Excellence

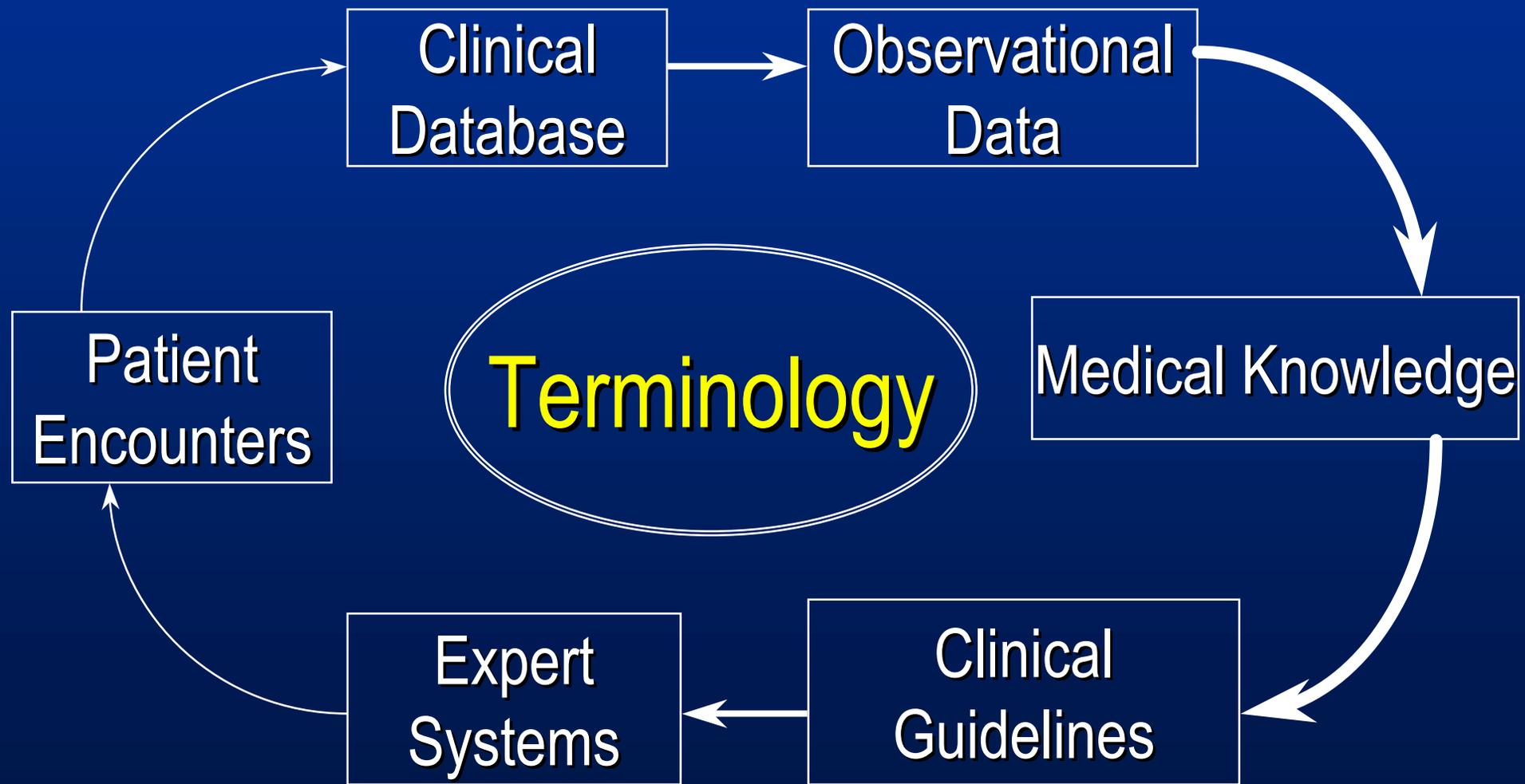
- Disease Natural History
- Treatment Response (non-RCT)
- Basis for Guidelines, Clinical Paths, Best Practice
- “Just in Time” Source for Decision Support
 - ◆ Have we seen a patient just like this...
- **Efficient and Effective Care Delivery**

Medical Concepts/Events

- How should we represent it? Language:
 - ◆ Nuance, detail, unfettered combination
 - ◆ Timely, current, never obsolete
 - ◆ Natural, friendly, established
 - ◆ [Ambiguous, imprecise, unpredictable]
- Codes:
 - ◆ Concise, precise
 - ◆ Structured, consistent, well formed
 - ◆ Analyzable, manipulable
 - ◆ [Rigid, tedious, high maintenance]

Heritage of Continuous Improvement

Central Role of Terminology



Clinical Process and Outcome Well-Established Traditions

- Practice Description
- Outcomes Research
- Continuous Quality Improvement
- Health Services Research
- Clinical Epidemiology
- Resource Management
- Guideline Development

Shifts in Emphasis: Early Perspective (the rest of Medicine)

Billable
Diagnoses

Clinical Data



Copernican Healthcare

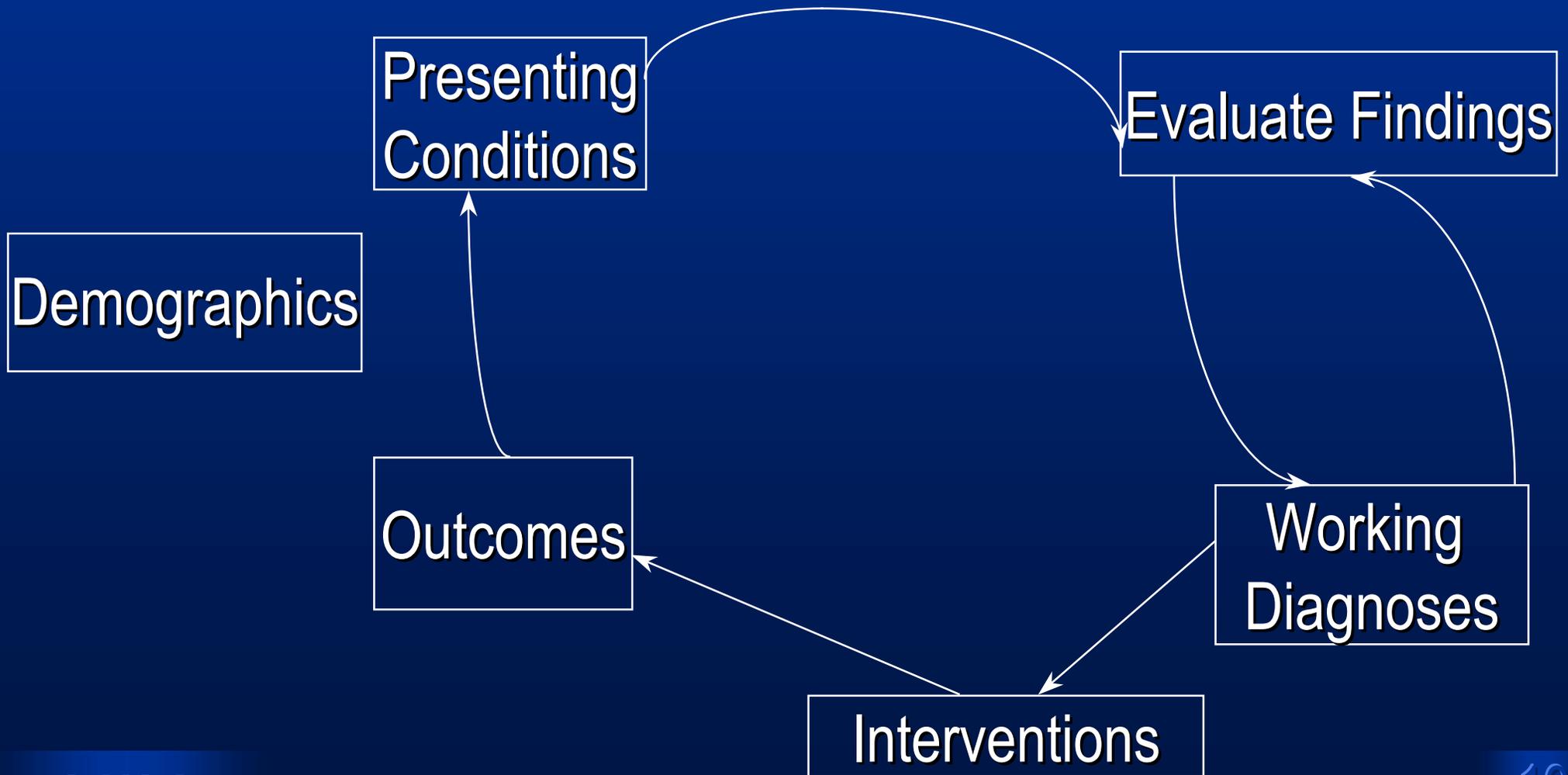
(Niklas Koppnigk)

Clinical
Data

Billable Data



Cycle of Longitudinal Care



Weights and Measures

“The nomenclature is of as much importance in this department of inquiry, as weights and measures in the physical sciences, and should be settled without delay.”

- ◆ William Farr, about Cullenian system
- ◆ First Annual Report of the Registrar-General of Births, Deaths, and Marriages in England. London: 1839 p. 99.

The Birth of a Formal Data Model for Cause of Death Classification

- First Statistical Congress, 1853 (Brussels)
 - ◆ William Farr (London)
 - ◆ Marc d'Espine (Geneva)
 - ◆ Achille Guillard (Paris) - Sponsor
- First International List (139 rubrics)
 - ◆ *Il y a lieu de former une nomenclature uniforme des causes de décès applicable a tous les pays.*
 - Sixteenth Annual Report of Registrar-General of England, 1853, Appendix, p. 73.

ICD *not* Catalog of Disease

“The International List of Causes of Death makes no pretension of being a proper nomenclature of diseases or of including a scientific classification of diseases.”

Introduction ICD-2, 1909

ICD Adaptations to Index Patients

- VA, Columbia, USPHS
 - ◆ coordinated effort, 1950-54
- Professional Activity Study (PAS) of
 - ◆ Commission on Professional and Hospital Activities (CPHA)
 - ◆ Completed 5 revisions of ICD (7), 1954-59.
 - ◆ ICDA #719 published 1959, 1962 (USPH)
 - included operations
 - ◆ *Adapted* for indexing hospital records.

Consider Confusion

Hospital patients coded in?

- ? VA/Columbia/USPHS ICD-7 revision
- ? PAS ICD-7 adaptation
- ? 1 of 5 CPHA ICD 7 revisions
- ? ICD-7 based ICDA #719; USPHS
- ? ICD-8 based ICDA #1693 ; USPHS
- ? H-ICDA on fragmented ICD-8
- ? H-ICDA-2, different still

ICD-9-CM

Grand Unified Effort

- Faithful template ICD-9
- Replaces ICDAs, H-ICDAs
- Contracted by HCFA to CPHA
- Professional Societies and Colleges

“...jointly restored and reformed the partnership between [*sic*] clinicians, statisticians, epidemiologists, and nosologists that originally existed when Dr. William Farr initiated work on the first international classification in the latter part of the 19th century.”

- Cost: \$2,700,000 (1977)

ICD-9-CM Resolution?

185 Malignant neoplasm of prostate
Excludes: seminal vesicles (187.8)

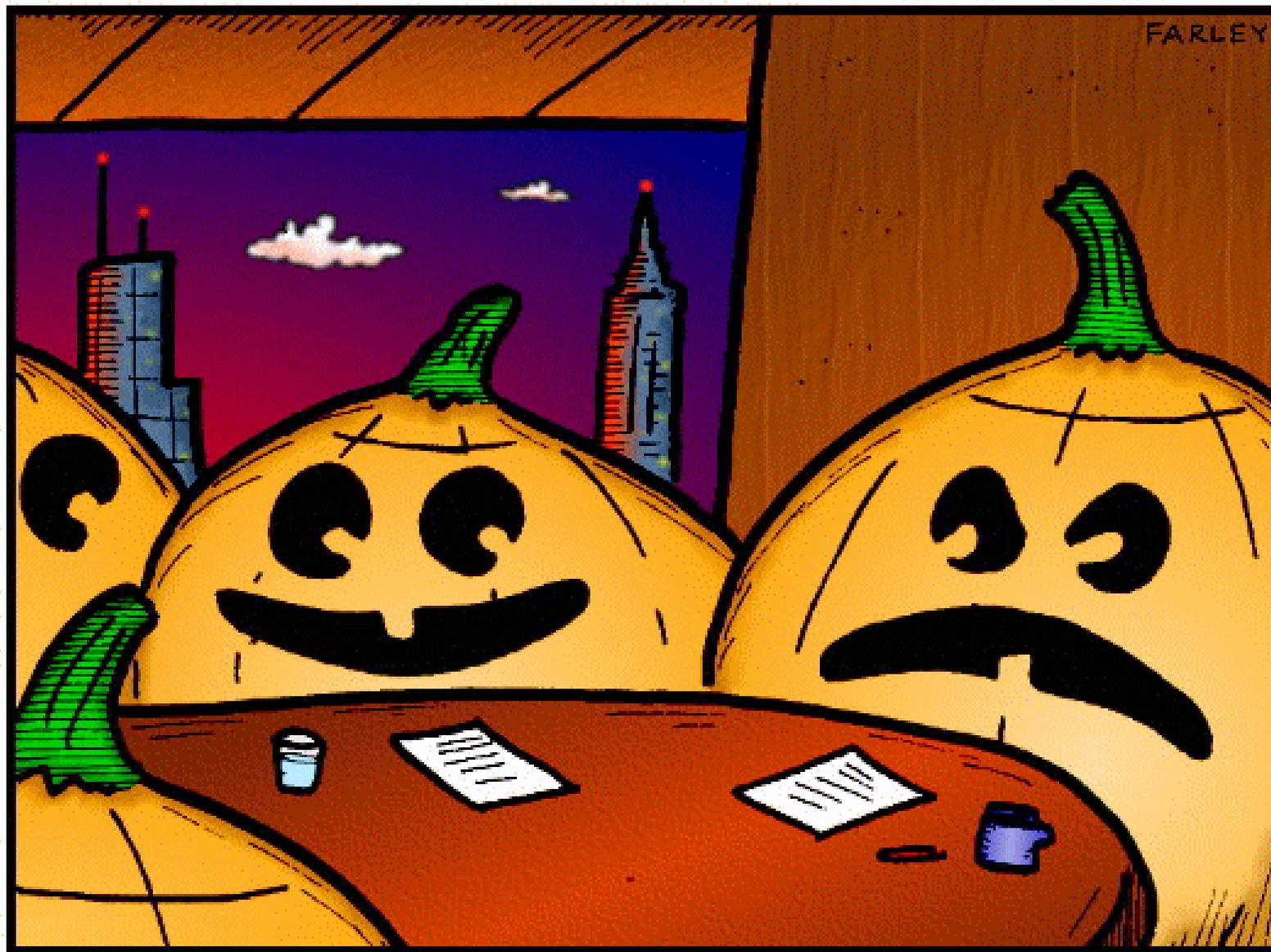
"Look! We can't all be named Jack!"

[Reproduced by Mayo with Permission]



10/29/93

DOCTOR FUN



© Copyright 1993 David Farley. World rights reserved.
This cartoon is made available on Usenet for personal viewing only.
dgfl@midway.uchicago.edu
Opinions expressed herein are not those of the University of Chicago.
This cartoon originally appeared in Campus Life.

"Look! We can't all be named Jack!"

Strategic Role of Vocabulary

- Interface to Knowledge Resources
 - ◆ Guidelines, Critical Paths, Reminders
 - ◆ Decision Support, Reference
- Support Practice Analysis
 - ◆ Quality Improvement
 - ◆ Clinical Epidemiology
 - ◆ Outcomes Analyses

Terminology as Crucial Requirement

Without Terminology Standards...

- Health Data is *non-comparable*
- Health Systems *cannot* Interchange Data
- Secondary Uses (Research, Efficiency) are *not possible*
- Linkage to Decision Support Resources *not Possible*

Information Beyond Practice

Secondary Re-use as Primary A Interest

- Data Collected for Clinical Care Forms the Basis for Patient Experience Repositories
- The Importance of a Well Characterized, High Quality Patient Experience Repository May Exceed the Value of the Primary Information Many Fold

Repositories of Patient Information

- Disease Natural History
- Treatment Response (non-RCT)
- Basis for Guidelines, Clinical Paths, Best Practice
- “Just in Time” Source for Decision Support
 - ◆ Have we seen a patient just like this...
- **Efficient and Effective Care Delivery**

Prospective Clinical Trials vs Patient Data Repository

- Still Require Consistent and Comparable Information
- Most Reliable if Captured *in and for* the Process of Care
 - ◆ Requires Terminology Services Integrated into Record System
- Efficiencies from *Non* ~~Boutique~~ Approach

How Well do Clinical Classifications Work?

- Computerized Patient Record Institute (CPRI)
- Study of Coding efficacy to measure content capture
- Volunteer effort among national working group on Codes and Vocabularies

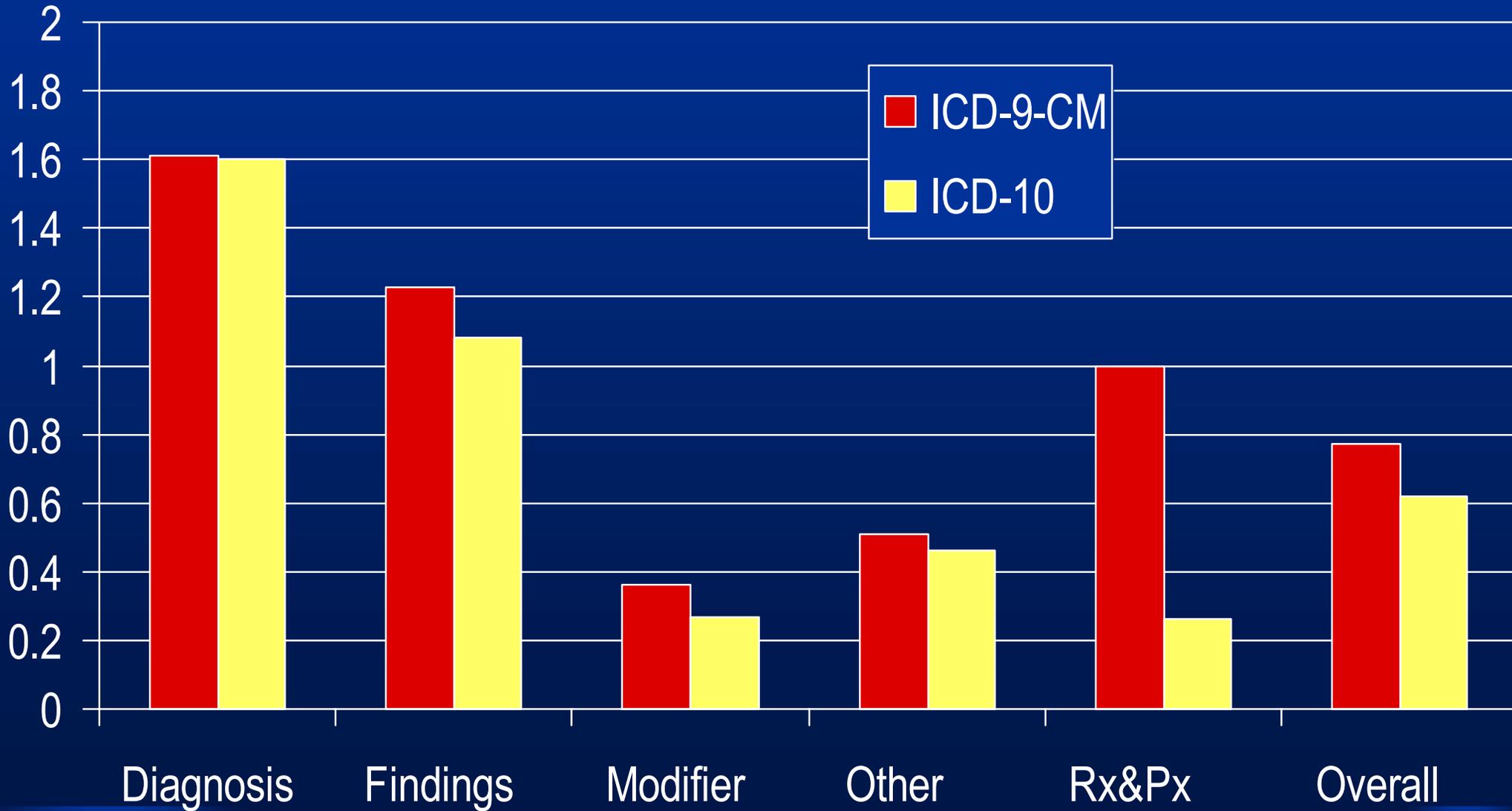
CPRI

Clinical Classifications

- Clinical text collected from:
 - ◆ Kaiser, Mayo Clinic, Park Nicollet, Stanford, University of Nebraska
- Clinical Domains:

<p>Inpatient notes</p> <p>Outpatient notes</p> <p>History and Physical Exam</p> <p>Pathology Report</p>	<p>Discharge Summary</p> <p>X-Ray Report</p> <p>Operative Report</p> <p>Nursing Notes</p> <p>Consultation Notes</p>
---	---

ICD-9-CM vs ICD-10 Composite Scores



Temporal Drift

Re-use of Codes

ICD-9-CM Procedure Codes

External Fixation example

1990 93.59 Other immobilization...

1991 78.4 Other repair or plastic op. bone

1992 78.1 Application of external fixation device

Note: 78.1 retitled from “periosteal suture”

=> 78.2 “limb shortening procedures”

which used to be in 78.3, now just “lengthening”

Standard Nomenclature of Diseases and Operations

- Initiated NY Academy of Medicine, 1928
 - ◆ virtually all medical societies
 - ◆ insurance companies
 - ◆ private industry
 - ◆ Commonwealth Fund
- National Conf. on Disease Nomenclature
 - ◆ Published 1933, 1935
- Responsibility assumed by AMA
 - ◆ Editions 1942, 1952, 1961

SNDO Example

Prostate Cancer

- Topographic Classification
 - ◆ 7 Urogenital system
- Etiologic Classification
 - ◆ 8 New growths

764-8091 Prostate Adenocarcinoma

- Unanticipated composition

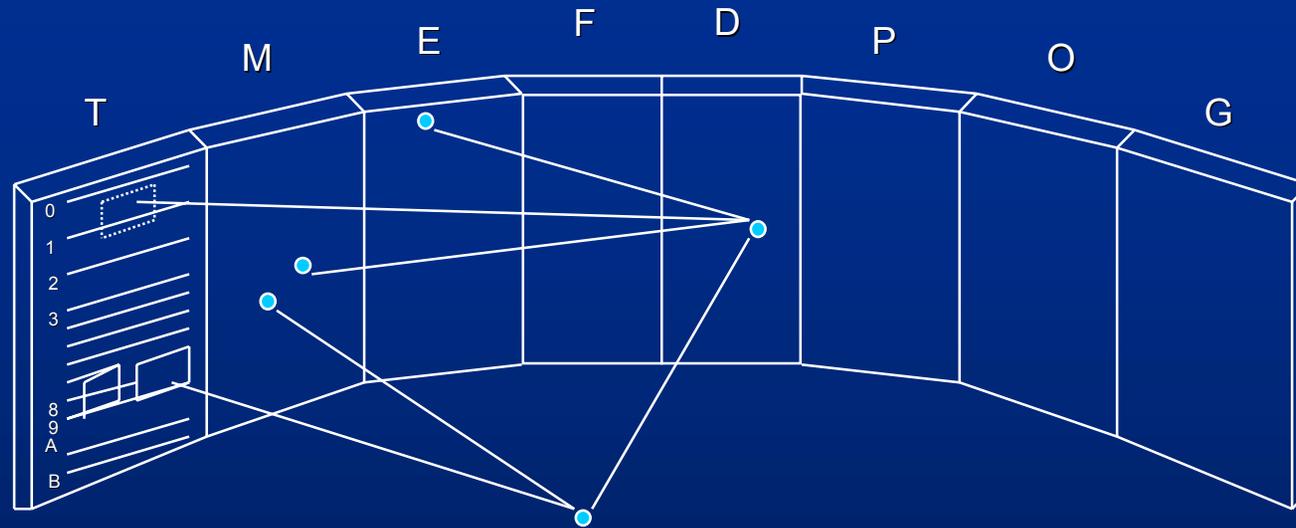
Systematized Nomenclature of Pathology (SNOP)

- College of American Pathologists, 1965
- Four Axes:
 - ◆ Function
 - ◆ Etiology
 - ◆ Morphology
 - ◆ Topography
- MONTAC cancer morphologies.

Systematized **NO**menclature of **MED**icine SNOMED II

- College of American Pathologists 1979, 1982
 - ◆ Topology*
 - ◆ Morphology*
 - ◆ Etiology*
 - ◆ Function*
 - ◆ Procedure, Occupation (ILO codes)
- } Disease
* *from SNOP*
- Information Qualifiers
 - ◆ Negation, Dx status, probability, special, admin.
 - Syntactic Linkage
 - ◆ 18 semantic relators

Information Space --- SNOMED



PEUTZ JAEGER'S POLYP/SYNDROME

T64, M7563

(CR.

T67 - Large Bowel;

T01 - skin M57 - pigmentation

T5103 - mouth

E-0103 - autosomal dominant

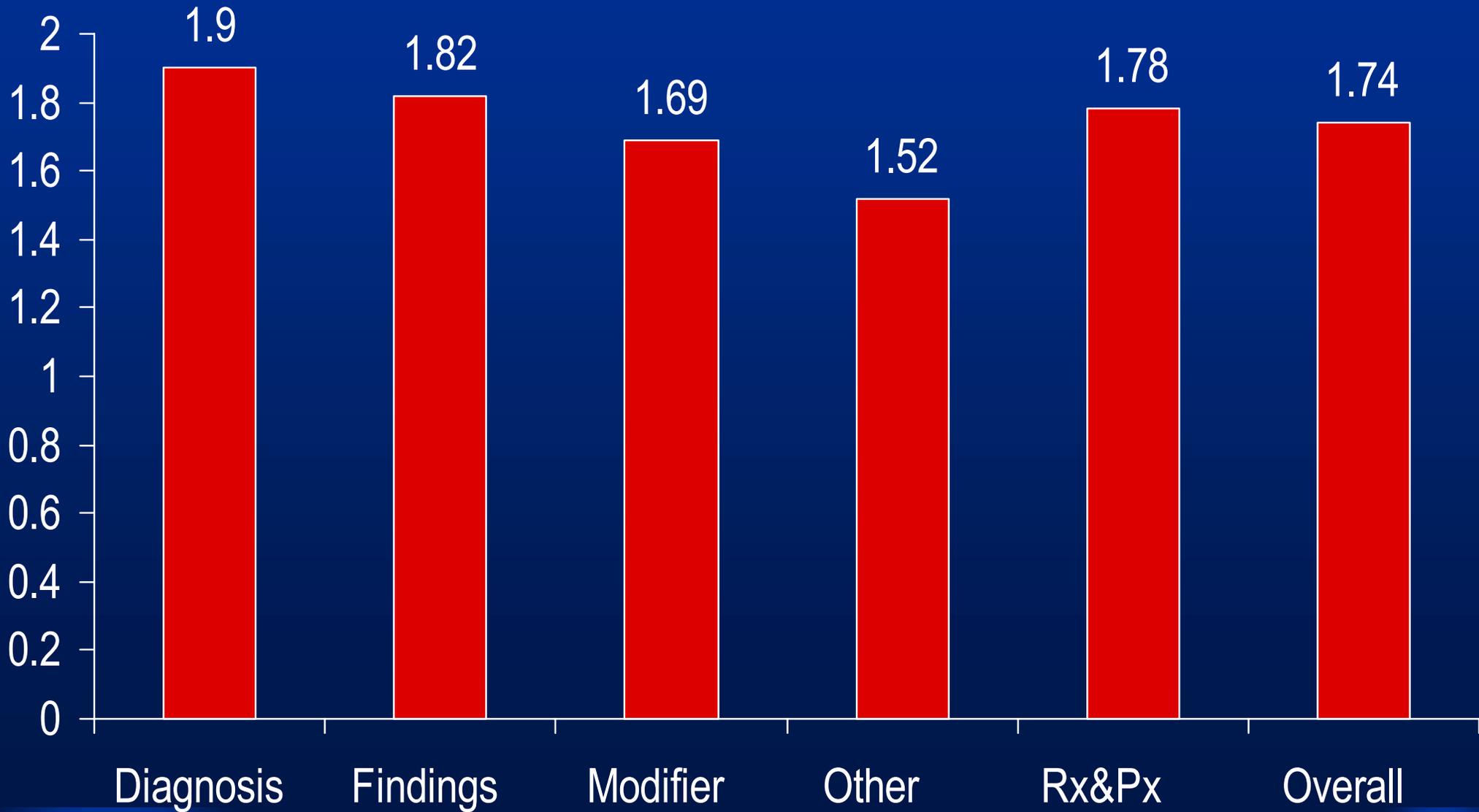
D 5432 - P.J. Syndrome)

SNOMED International (SNOMED III)



- Topology (12,385)
- Morphology (4,991)
- Function (16,352)
- Living Organisms (24,265)
- Chemicals, Drugs, and Biologicals (14,075)
- Physical Agents, Forces, and Addictives (1,355)
- Occupations (1,886)
- Social Context (433)
- Disease/Diagnoses (28,623)
- Procedures (27,033)
- General Linkage/Modifiers (1,176)

SNOMED 3 Composite Scores



NHS - CAP Agreement SNOMED CT (Clinical Terms)

- SNOMED RT and Clinical Terms version 3 (Read codes) is being merged to create a new terminology of health
- NHS Information Authority has appointed representatives to the SNOMED Authority and SNOMED International Editorial Board
- Release: 10 Dec 2001

What Constitutes A “Good” Terminology Desiderata

- Cimino Desiderata
- ANSI HISB – CPRI Framework
- ASTM Standard

CPRI ANSI/HISB Terminology Framework

General Characteristics

- Completeness
 - ◆ Explicit In-depth Coverage
- Non-redundant
- Mapping
- Comprehensiveness -
 - ◆ Addressed All Segments of Healthcare
- Characteristics of Integration
 - ◆ Non-overlapping
 - ◆ Integrated/interlocking

Terminology Framework

Structure of the Terminology Model

- Atomic Base
- Compositional
- Synonyms
- Attributes
 - ◆ Inheritance
- Multiple Hierarchies
 - ◆ Consistency of View
- Explicit Uncertainty
- Lexical Rules
- Representation
 - ◆ must not have arbitrary restrictions, such as numbers of digits

Terminology Framework *Maintenance*

- Context Free Identifiers
- Unique Identifiers
- Version Control
 - ◆ Dated
 - ◆ Obsolete Marking
- Definitions
- Language Independence
- Responsiveness

Terminology Framework *Administration*

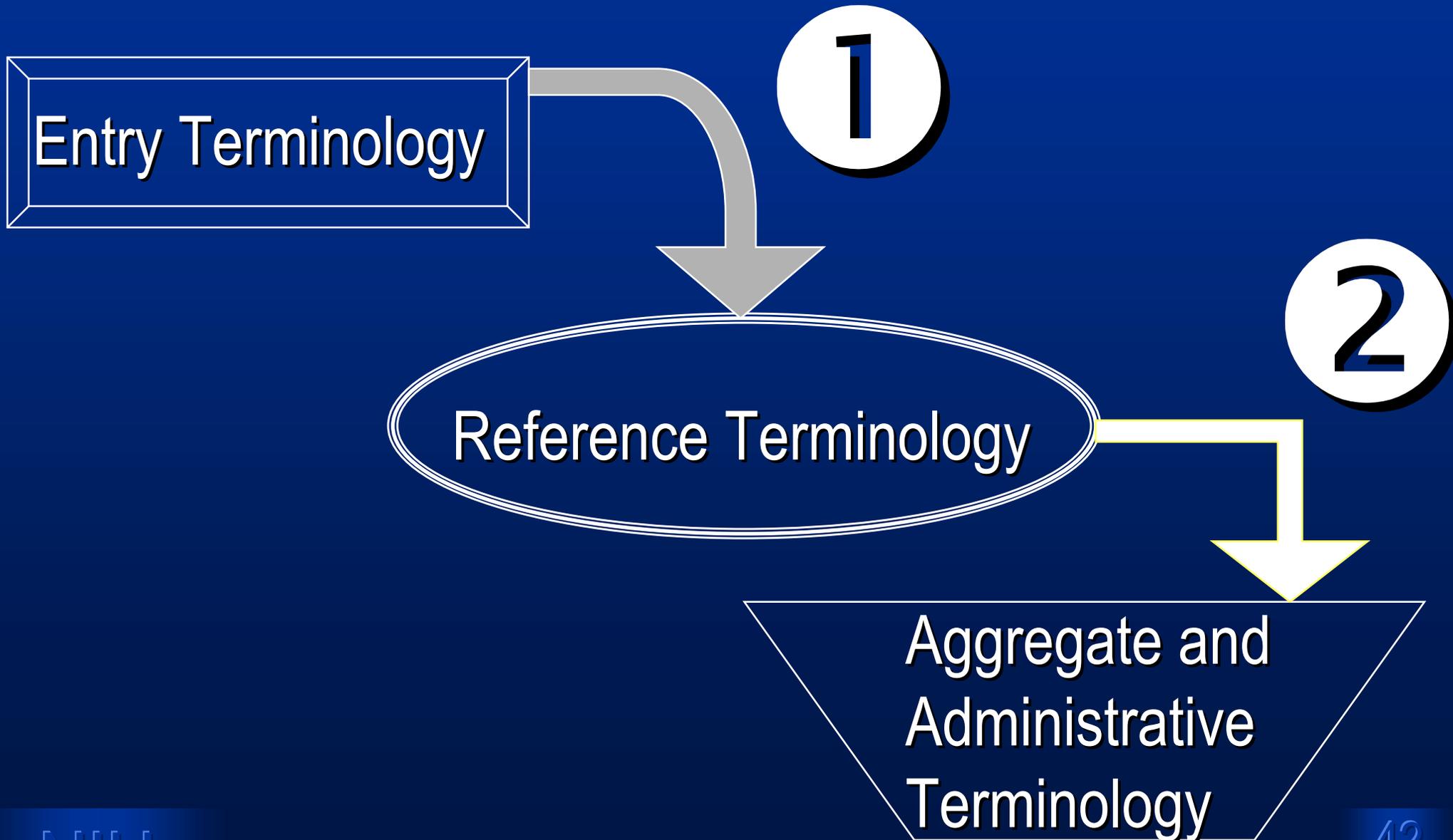
- Coordination
- Distribution
 - ◆ Negligible Cost, Internet
- Usage - public domain is desirable.
- Pricing
 - ◆ Distribution costs must not impose a barrier
- Funding
 - ◆ from the public and private sectors.

Challenge

Perfect Vocabulary is not sufficient for the clinically complete and appropriate capture of patient findings, events, or diagnoses.

- Human Navigation
- Computer Interface

Terminology Systems Relationships Engineering Map

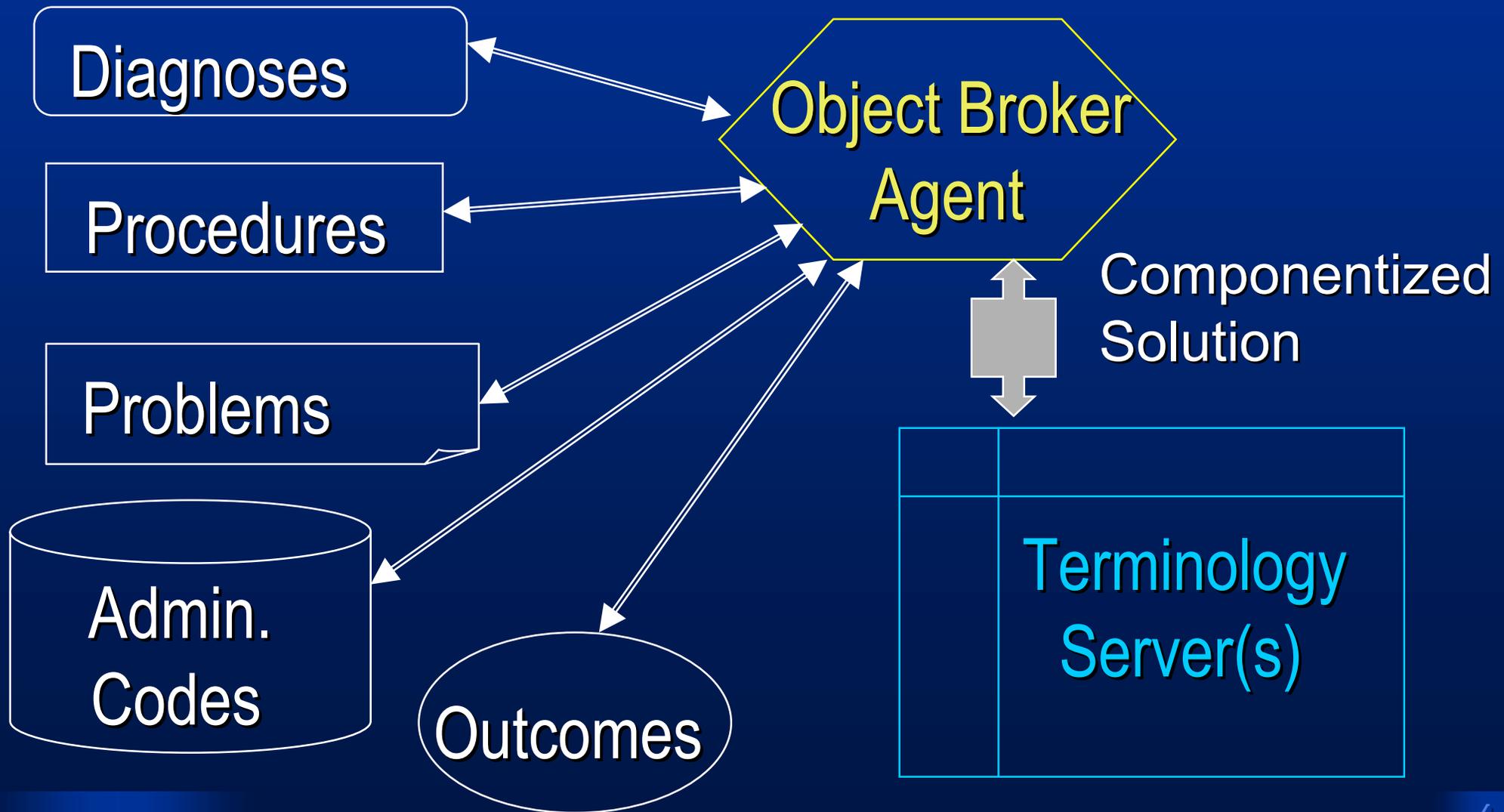


Mayo's Work with Problem List Interface Design



- Premise upon Terminology Server
 - ◆ MetaPhrase Prototypes on the Network
- Iterative Usability Lab Evaluations
 - ◆ Mock-ups in VB, Delphi, Java, ...
- Evolve Toward Subset of Functional Needs
 - ◆ Problem List Specific
 - ◆ Drive Specification and Operation of T Server

Object Terminology Server Middleware Architecture



Terminology Navigator Functionality

- Navigate to well formed clinical terminologies
- Invoke word completion and spelling correction
- Intelligent lexical normalization (Specialist Lexicon tools)
- Invoke Semantic Locality Browsing

Preliminary Terminology Server Desiderata (AMIA '99)

1. Word Normalization
2. Word Completion
3. Target Terminology Specification
4. Spelling Correction
5. Lexical Matching
6. Term Completion
7. Semantic Locality
8. Term Composition
9. Term Decomposition

Mayo “Short List” Breadboard Components

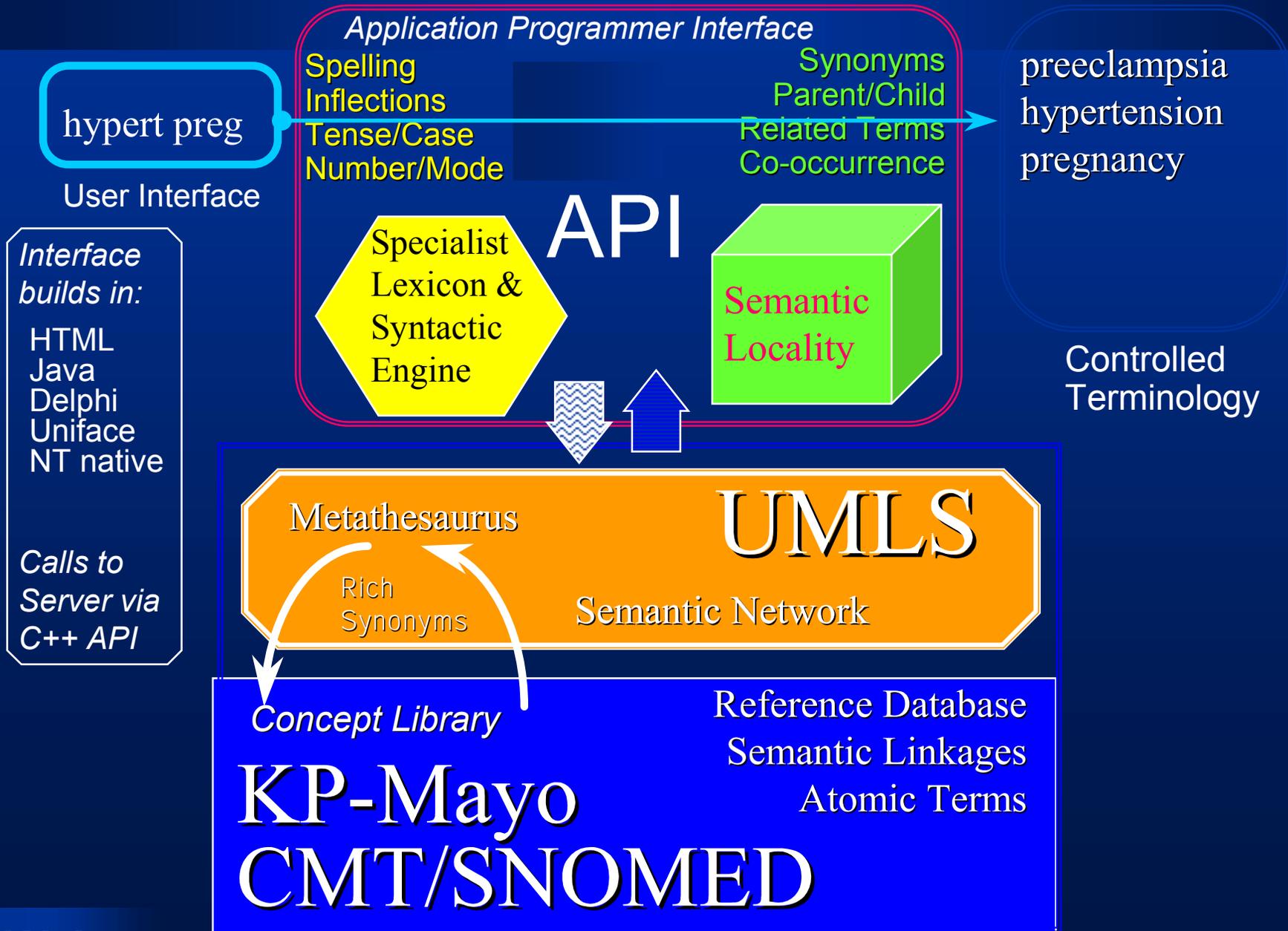
- LVG Interface
- Spelling Correction
- Synonymy
- Term Completion
- Word Completion
- Match Retrieval

Terminology Server Components Goals

- Platform Independent Components
 - ◆ CORBA, Java, COM, etc.
- Enable Independent Optimization
 - ◆ Component Specific Parameter Specification
 - ◆ Inclusion/Exclusion of Functionality
 - ◆ Calling Sequence (Spell Correct before LVG?)
 - ◆ Distribution of Server Resources
- Target Terminology Independent

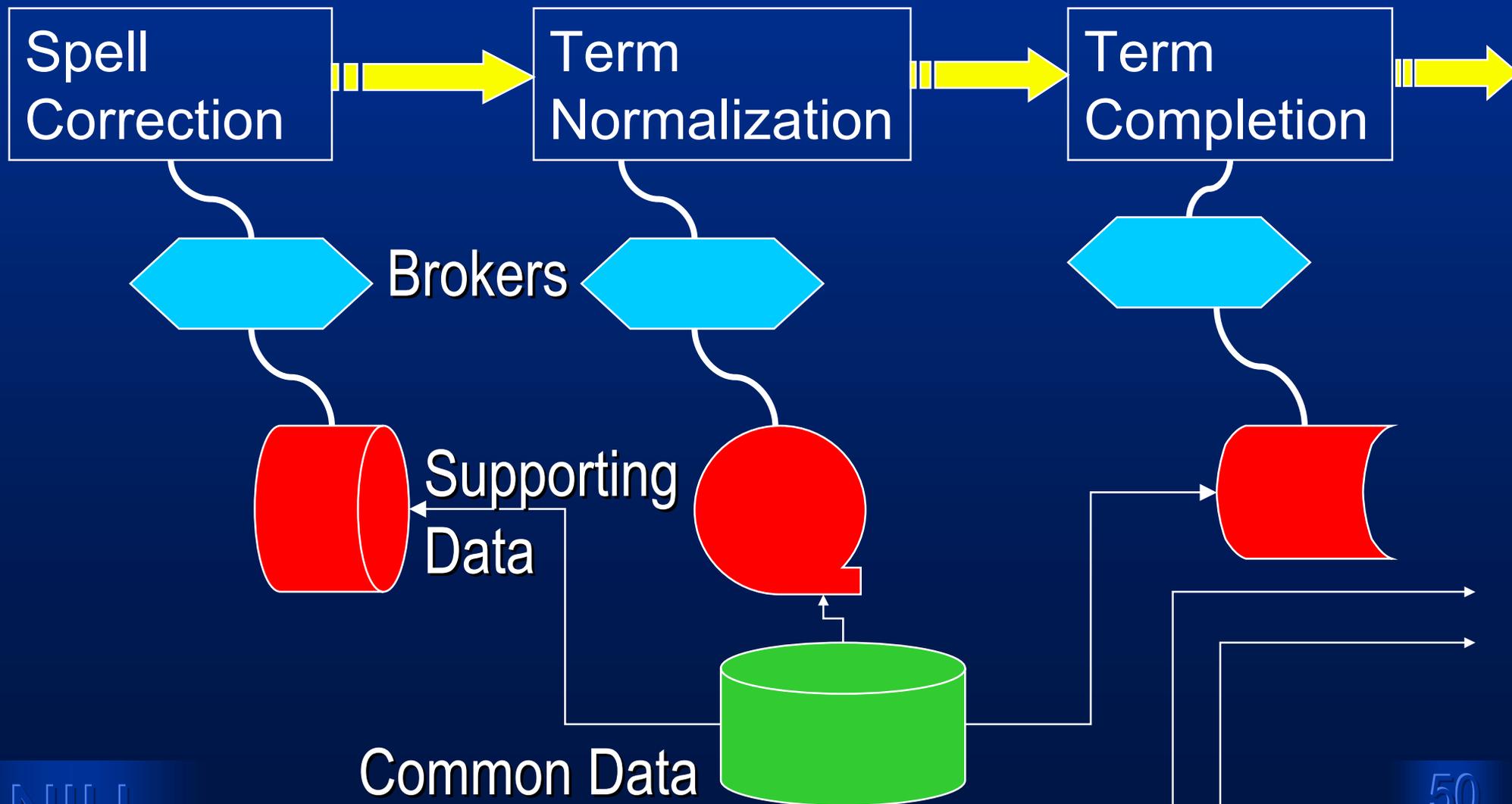
Monolithic Terminology Services

Old Think



Component Object Architecture

New Think



Elements of Enterprise Lexicon

- Abbreviations
- Acronyms
- Synonyms
- Lexical Variants
- Term Expansions
- Word Completions
- Phonemes (Speech Recognition)

Characteristics of Enterprise Lexicon

- Context Sensitive
- Partitioned by Use
- Centrally Maintained and Distributed
- Integrated Across Lexicon Elements
(above)

Compositionality

Next Major Research Arena

- Cannot Possibly Anticipate Coverage
- “Large Scale Vocabulary Test” Conducted by NLM (with Mayo MIR participation)
 - ◆ 58% Coverage Among Major Terminologies
- Mayo Study (AMIA ‘98)
 - ◆ 28% Using Pre-coordinated Terms
 - ◆ 54% Invoking Composition

Attribute/Nominal Nature Medical Concept

- Atomic clinical concepts (existence?)
- (N_, A1, A2, A3, A4, ...) Blois '84
Colon cancer or (Neoplasm _ Colon)
“Dukes B2 Adenosquamous carcinoma of ascending colon with *C-myc* oncogene”
- Pre-coordinated =>combinatorial explosion
- Continuous data (values)
- Attribute levels (stage, grade)

Medical Event Representation

Clinical Concept

Dukes B[d,2] adenosquamous cell carcinoma of ascending colon with *c-myc* oncogene production

- Histology/Morphology
- Anatomy/Laterality
- Probability/Negation
- Extent
- Direction
- Functional impact
- Stage
- Grade
- Severity
- Chronicity
- Etiology
- Characteristics

Human Interface Desiderata

- Fast, Easy, and Intuitive
- Expressive, Compositional, Complete
- Componentized, Configurable, Scaleable
- Maintainable, Centralized, Deployable
- Consistent Backbone

What Have We Done?

- Develop Human Interface to Compositional Tool (4 Years Evolution)
 - ◆ Natural Language Model
 - ◆ Speech Enabled
 - ◆ Invoke Suite Terminology Services
- Componentized Network Services
 - ◆ CORBA or COM Enabled
 - ◆ Synchronized Content

P1: Sir William Osler, MD 10:30 AM
Cellulitis: M-41650
Right foot: T-D9710
Osteomyelitis: D1-60210
Third metatarsal bone: T-12870
Without: G-C009
Lymphangitis: D3-A0040

- Suggested Terms
Right foot: T-D9710
Cellulitis: M-41650
Osteomyelitis: D1-60210
Without: G-C009
Lymphangitis: D3-A0040
Third metatarsal bone: T-12870
Right: G-A100
Foot: T-D9700
Hindfoot: T-D9701
Third: G-A703
Metatarsal: T-12840

- Related Terms for: Cellulitis
Children
Synonyms
Cellulitis
Phlegmonous cellulitis
Phlegmon
Parents
Siblings
Abscess cavity: M-41612
Acute abscess: M-41614
Acute empyema: M-41620
Acute exudative inflammation: M-41100
Acute fibrinous inflammation: M-41300
Acute hemorrhagic inflammation: M-41700

Application Opportunities Controlled Terminology Services

- Physician Input Tool
 - ◆ Problem Lists, Indication for Order, Reason for Visit, Summary Diagnoses, etc
- Parse and Code
 - ◆ Clinical Notes, Procedure Reports, Discharge Summaries, etc
- Obviate “Master Sheet” Model

Continuum from Nomenclature to Classification

- Patient Data is Highly Detailed
 - ◆ Modifiers: Anatomy, Stage, Severity, Extent
 - ◆ Qualifiers: Probability, Temporal Status
- Aggregate Uses Require Categorization
 - ◆ Granularity of Classifiers
 - Focused Groups and Strata for CQI/Outcomes
 - Broad Statistical/Fiscal Groups

MAPPING:

Three levels of mapping



[from K. Spackman]

1. Code-to-code relationships
2. Rules indicating additional considerations in evaluating a map
3. Software that executes the rules in the context of a particular patient record architecture

The easy part of mapping

[from K. Spackman]

- CPT

- ◆ 47715 Excision of choledochal cyst

- SNOMED

- ◆ P1-5C362 Excision of choledochal cyst

The easy part of mapping???

[from K. Spackman]

- SNOMED
 - ◆ DC-10010 Anemia
- ICD-9-CM
 - ◆ 285.9 Anemia, NOS
- But DC-10010 subsumes chronic blood loss anemia, and 285.9 explicitly excludes it!

A little closer look:

[from K. Spackman]

- 285.9 Anemia, unspecified
 - ◆ Anemia:
 - NOS; essential
 - normocytic, not due to blood loss
 - profound; progressive; secondary
 - ◆ Oligocythemia
 - ◆ **Excludes:** anemia (due to):
 - blood loss:
 - ◆ acute (285.1)
 - ◆ chronic or unspecified (280.0)
 - iron deficiency (280.0-280.9)

SNOMED-to-CPT: Many-to-one

[from K. Spackman]

- CPT:

- ◆ 23130 Acromioplasty or acromionectomy, partial

- SNOMED:

- ◆ P1-16394 Partial acromionectomy
- ◆ P1-1689A Partial acromioplasty

Implications and Opportunities

- Encoding Clinical Data Using a Nomenclature Is More Reproducible Than Coding in Extant Classification Systems
- A Formally Specified Map (Including Rules) Improves Reproducibility of Administrative Encoding and Reduces Cost
- Implications for Fraud/Abuse Provisions, Medical Necessity Documentation, etc.

Mayo's Role in Mapping Development

- Recognize that ICD-9-CM is Poorly Defined
 - ◆ Marginalia, Coding Clinics NA, Case Law
- Negotiating Contract with HCFA
 - ◆ Parse ICD-9-CM Elements (Includes, Excludes)
 - ◆ Parsimonious Subset SNOMED Terms
 - ◆ Compile Elements into Formal Rule Base
- Logical Rule Base may Replace Current Coding Definitions – HCFA Publication

Mapping Practicality

- Classification Requires Formal Rule Base
 - ◆ Does Not Yet Exist
- Recognize Presence of Clinical Elements in Record
- Re-aggregate Clinical Events into “Best Match” Rubric of Classification (ICD)
- **Term to Term Maps *Will* (Do) Fail!**

Terminologies Today

- SNOMED RT/CT
- LOINC
 - ◆ Public Source for Lab Tests and Processes
 - ◆ 6-part Internal Structure
 - ◆ Fits HL7 Message Format (V2 and V3)
- Entry Terms

Future Concerns: The New Medicine Organizationally Play Well or Die

- Genomics and Bioinformatics
 - ◆ Clinical Correlation
 - ◆ Disease Natural History by SNP
 - ◆ Treatment Selection and Outcome
- Repositories of Patient Experience
 - ◆ Drive Delivery of Efficient and Effective Care
- Require Structured, Detailed Clinical Description

Terminology Resources on the Web (All Mayo Sponsored)

- CPRI Terminology Summits
 - ◆ National Conference 1997
 - ◆ Joint HL/7 Overview to G-CPR 1998
 - ◆ Second National Conference 1999
 - ◆ www.cpri.org/terminology
- IMIA WG6 Triennial Forums (94, 97,00)
 - ◆ www.imia.org/wg6

ISO TC 215

WG 3 Health Concept Representation

- Metavocabulary about Terminology
- Foundation Structures of Terminology
- Quality Indicators and Good Development Practices for Terminology
- System of Semantic Links and Concepts for Health Terms
- Unified Terminology Model for Nursing

Critical Path Concern

Confidentiality

Conclusion

- Terminology Is the Second Most Urgent Issue in Healthcare Information Today
- Problem Underlies Virtually All Secondary Uses of Patient Data
- Convergence Toward Common System is Critical and Occuring
- Integration into Clinical Systems is Required

Christopher G. Chute MD, DrPH
Professor of Medical Informatics
Head, Section Medical Information Resources
Mayo Clinic/Foundation
Rochester, MN 55905

Ph: 507 284 5541

Fx: 507 284 1516

Em: chute@Mayo.edu