Speaker

Christopher Chute, MD, DrPH
Johns Hopkins University
Baltimore, MD, USA

JOHNS HOPKINS MEDICINE

Comparability and Consistency of Clinical Data: The ICD11 initiative

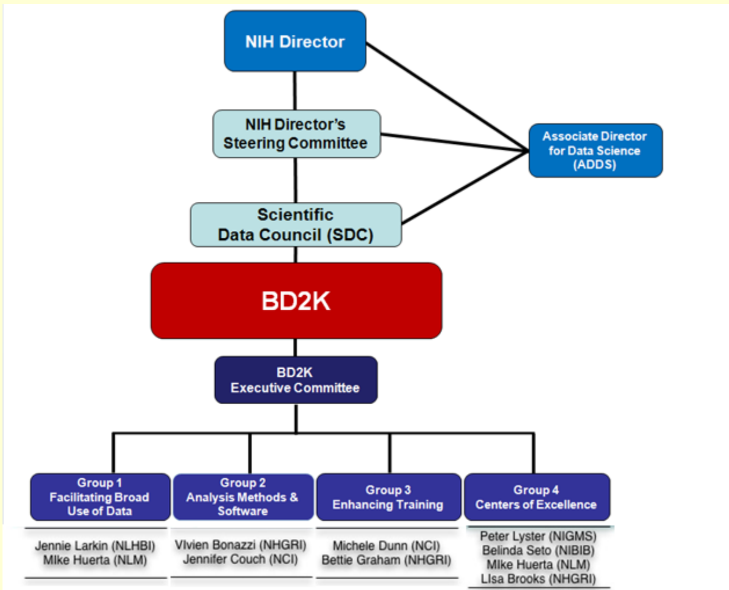
Christopher G. Chute, MD DrPH
Bloomberg Distinguished Professor of Health Informatics
Professor of Medicine, Public Health, and Nursing
Chief Research Information Officer
Johns Hopkins Medicine, Baltimore, MD, USA

Chair, International Classification of Disease Revision, WHO
Former Chair, ISO TC215 Health Informatics

ISPOR
Singapore, 5 Sept 2016

JOHNS HOPKINS MEDICINE

Big Data to Knowledge: BD2K



JOHNS HOPKINS MEDICINE

NIH Precision Medicine Initiative


The NEW ENGLAND JOURNAL of MEDICINE

Perspective
FEBRUARY 26, 2015

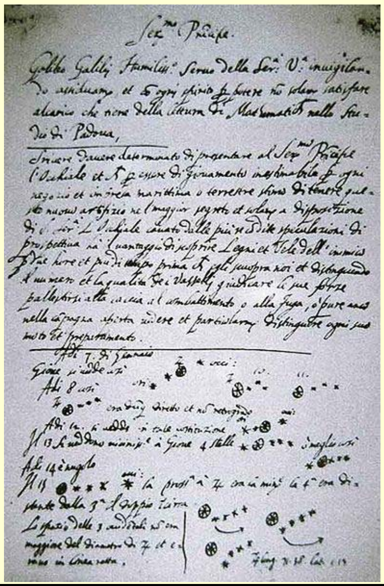
A New Initiative on Precision Medicine

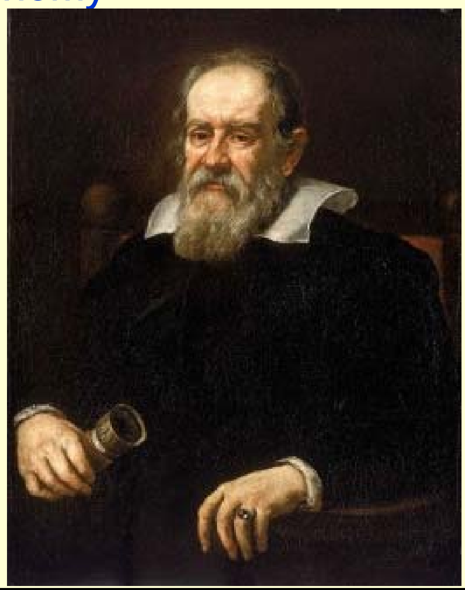
Francis S. Collins, M.D., Ph.D., and Harold Varmus, M.D.

"Tonight, I'm launching a new Precision Medicine Initiative to bring us closer to curing diseases like cancer and diabetes — and to give all of us access to the personalized information we need to keep ourselves and our families healthier."

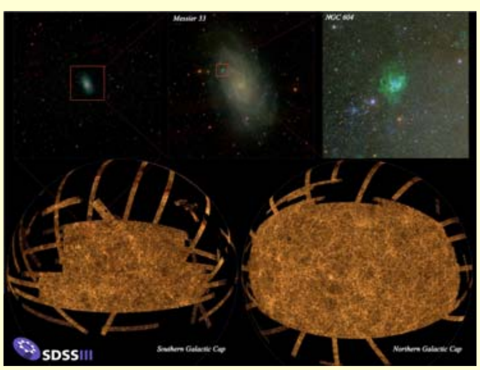
— President Barack Obama, State of the Union Address, January 20, 2015

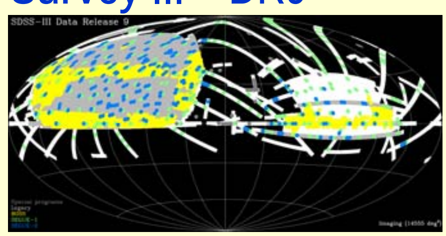
Origins of Big Science
Astronomy





Sloan Digital Sky Survey III – DR9

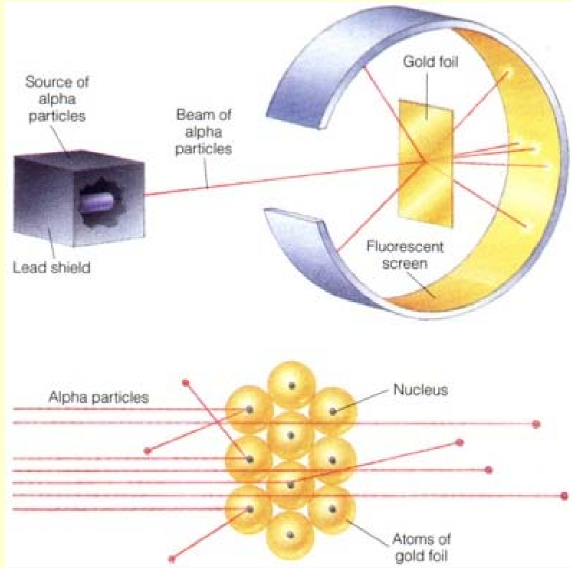


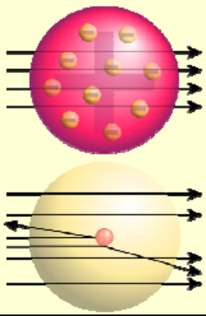


- Images
- Spectra
- Object catalog
- Metadata

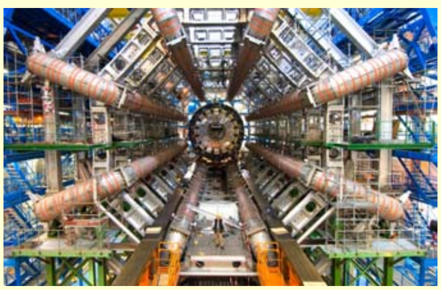
Total area of imaging	31,637 square degrees
Image field size	1361x2048 pixels
Number fields	938,046 (excluding supernovae runs)
Catalog objects	1,231,051,050
Number of unique, primary sources	
Total	469,053,874
Stars	260,562,744
Galaxies	208,478,448
Unknown	12,682

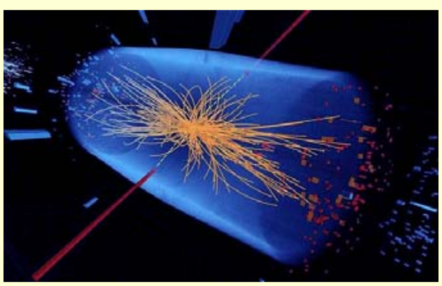
Rutherford "Table-top" Experiment





That Higgs Boson





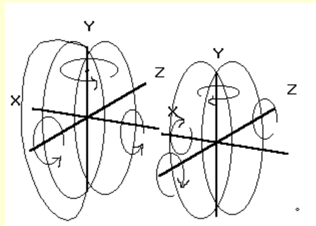
- 600 institutions
- 10,000 scientists
  - 2 Nobel prizes
- 800 trillion collisions
- 200PB of data =
  - 2×10^{17} bytes of data

Boarding on an astronomical number in its own right!

- $\$13.25 \cdot 10^9$ USD

Dimensionality of Higgs "Big Data"

- Mass/Energy
- Direction
- Charge



- Medicine is more complicated than that

Dimensionality of Big Data

- Broad
  - Small amounts of data; Huge number observations
  - National Claims data
- Deep
  - Large amounts of data; Few observation
  - NGS Complete Genome
- Rich
  - Broad and Deep
  - Clinical Phenotyping data (EMRs)
    - Labs, Vitals, Exam, Waveform, Images, Omics, …
    - Social, environmental, diet,

JOHNS HOPKINS MEDICINE

Will Big Data Save Us?

Genetics inMedicine | **REVIEW**

Genet Med 15: 802-809;
© American College of Medical Genetics and Genomics
Oct, 2013

Some experiences and opportunities for big data in translational research

Christopher G. Chute, MD, DrPH[1], Mollie Ullman-Cullere, MS, MSE[2], Grant M. Wood, BS[3], Simon M. Lin, MD[4], Min He, PhD[4] and Jyotishman Pathak, PhD[1]

Health care has become increasingly information intensive. The advent of genomic data, integrated into patient care, significantly accelerates the complexity and amount of clinical data. Translational research in the present day increasingly embraces new biomedical discovery in this data-intensive world, thus entering the domain of "big data." The Electronic Medical Records and Genomics consortium has taught us many lessons, while simultaneously advances in commodity computing methods enable the academic community to affordably manage and process big data. Although great promise can emerge from the adoption of big data methods and philosophy, the heterogeneity and complexity of clinical data, in particular, pose additional challenges for big data inferencing and clinical application. However, the ultimate comparability and consistency of heterogeneous clinical information sources can be enhanced by existing and emerging data standards, which promise to bring order to clinical data chaos. Meaningful Use data standards in particular have already simplified the task of identifying clinical phenotyping patterns in electronic health records.

Genet Med advance online publication 5 September 2013

Key Words: clinical data representation; big data; genomics; health information technology standards

JOHNS HOPKINS MEDICINE

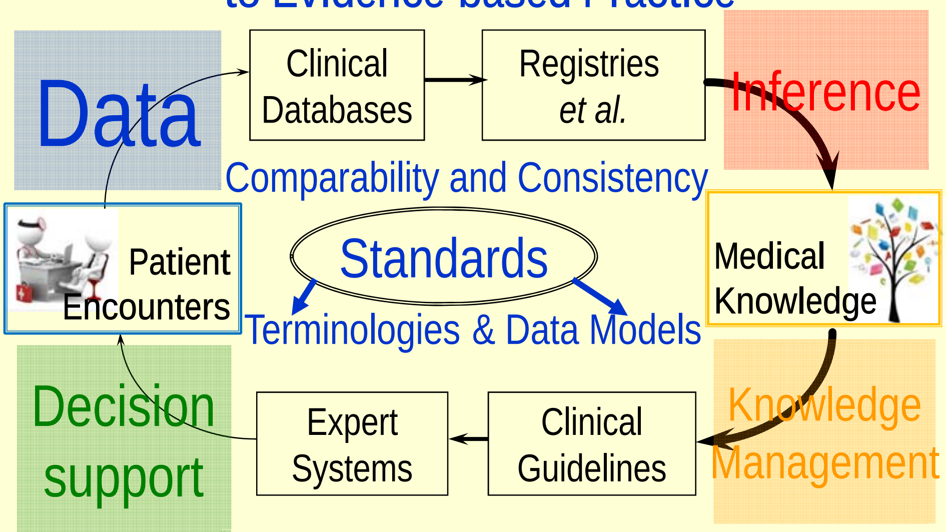
Comparable and Consistent Semantics as Crucial Requirement

Without Terminology/Ontology Standards...

- Health Data is *non*-comparable
- Health Systems *cannot* Interchange "Data"
- Secondary Uses (Research) are *not* practical
- Big Data methodologies cannot leverage epidemiologic principles for observational data
  - Adjustment for confounding
    - Stratification
    - Multivariate models
    - Machine learning features

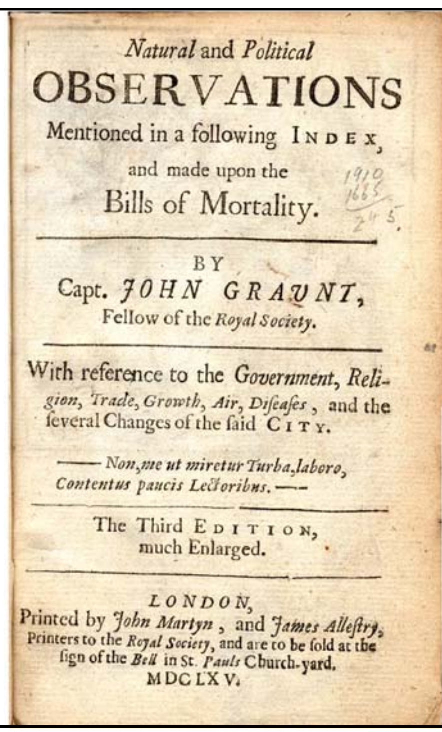
JOHNS HOPKINS MEDICINE

From Practice-based Evidence to Evidence-based Practice



Foundations for Learning Health System

JOHNS HOPKINS MEDICINE



Natural and *Political*

OBSERVATIONS

Mentioned in a following INDEX,
and made upon the
Bills of Mortality.

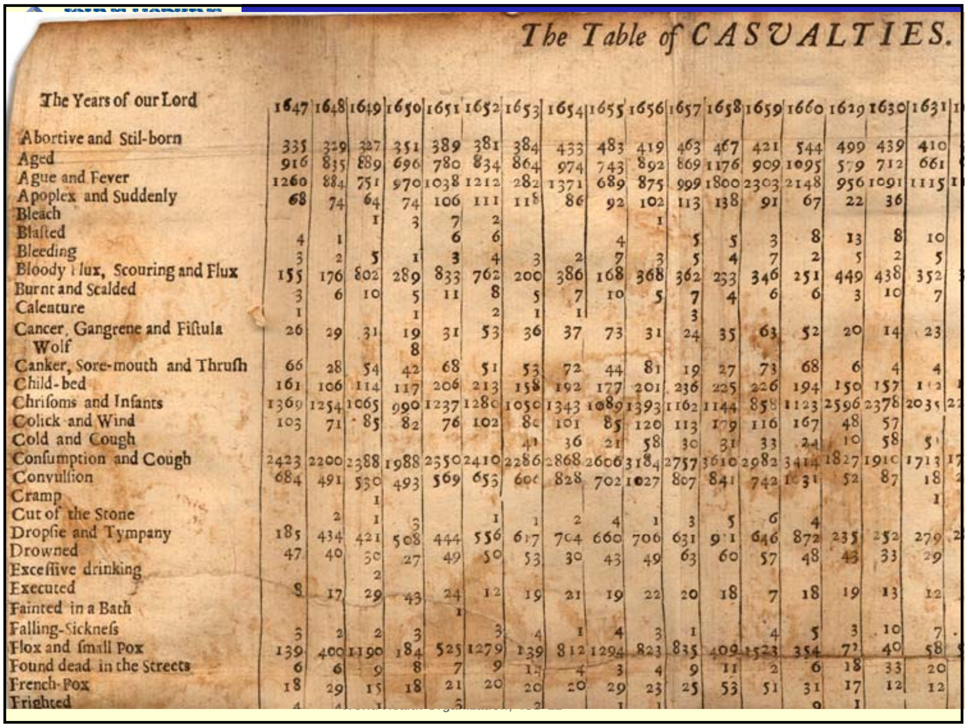
BY
Capt. JOHN GRAUNT,
Fellow of the *Royal Society*.

With reference to the *Government*, *Religion*, *Trade*, *Growth*, *Air*, *Diseases*, and the several Changes of the said CITY.

—— *Non, me ut miretur Turba, laboro, Contentus paucis Lectoribus.* ——

The Third EDITION, much Enlarged.

LONDON,
Printed by *John Martyn*, and *James Allestry*, Printers to the *Royal Society*, and are to be sold at the sign of the *Bell* in St. *Pauls* Church-yard.
MDCLXV.



The Table of CASUALTIES.

The Years of our Lord	1647	1648	1649	1650	1651	1652	1653	1654	1655	1656	1657	1658	1659	1660	1629	1630	1631
Abortive and Stil-born	335	329	327	351	389	381	384	433	483	419	463	467	421	544	499	439	410
Aged	916	835	889	696	780	834	864	974	743	892	869	1176	909	1095	579	712	661
Ague and Fever	1260	884	751	970	1038	1212	282	1371	689	875	999	1800	2303	2148	956	1091	1115
Apoplex and Suddenly	68	74	64	74	106	111	118	86	92	102	113	138	91	67	22	36	
Bleach			1	3	7	2				1							
Blasted	4	1			6	6			4		5	5	3	8	13	8	10
Bleeding	3	2	5	1	3	4	3	2	7	3	5	4	7	2	5	2	5
Bloody Flux, Scouring and Flux	155	176	802	289	833	762	200	386	168	368	362	233	346	251	449	438	352
Burnt and Scalded	3	6	10	5	11	8	5	7	10	5	7	4	6	6	3	10	7
Calenture	1			1		2	1	1			3						
Cancer, Gangrene and Fistula	26	29	31	19	31	53	36	37	73	31	24	35	63	52	20	14	23
Wolf				8													
Canker, Sore-mouth and Thrush	66	28	54	42	68	51	53	72	44	81	19	27	73	68	6	4	4
Child-bed	161	106	114	117	206	213	158	192	177	201	236	225	226	194	150	157	112
Chrisoms and Infants	1369	1254	1065	990	1237	1280	1050	1343	1089	1393	1162	1144	858	1123	2596	2378	2035
Colick and Wind	103	71	85	82	76	102	80	101	85	120	113	179	116	167	48	57	
Cold and Cough							41	36	21	58	30	31	33	24	10	58	5
Consumption and Cough	2423	2200	2388	1988	2350	2410	2286	2868	2606	3184	2757	3610	2982	3414	1827	1910	1713
Convulsion	684	491	530	493	569	653	606	828	702	1027	807	841	742	1031	52	87	18
Cramp			1														1
Cut of the Stone		2	1	3		1	1	2	4	1	3	5	6	4			
Dropsie and Tympany	185	434	421	508	444	556	617	704	660	706	631	931	646	872	235	252	279
Drowned	47	40	30	27	49	50	53	30	43	49	63	60	57	48	43	33	29
Excessive drinking			2														
Executed	8	17	29	43	24	12	19	21	19	22	20	18	7	18	19	13	12
Fainted in a Bath					1												
Falling-Sickness	3	2	2	3		3	4	1	4	3	1		4	5	3	10	7
Flox and small Pox	139	400	1190	184	525	1279	139	812	1294	823	835	409	1523	354	72	40	58
Found dead in the Streets	6	6	9	8	7	9	14	4	3	4	9	11	2	6	18	33	20
French-Pox	18	29	15	18	21	20	20	20	29	23	25	53	51	31	17	12	12
Frighted	[illegible]	[illegible]	[illegible]		[illegible]		[illegible]		[illegible]	[illegible]				[illegible]	[illegible]		

JOHNS HOPKINS MEDICINE

Weights and Measures
William Farr

"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."

- First Annual Report of the Registrar-General of Births, Deaths, and Marriages in England. London: 1839 p. 99.

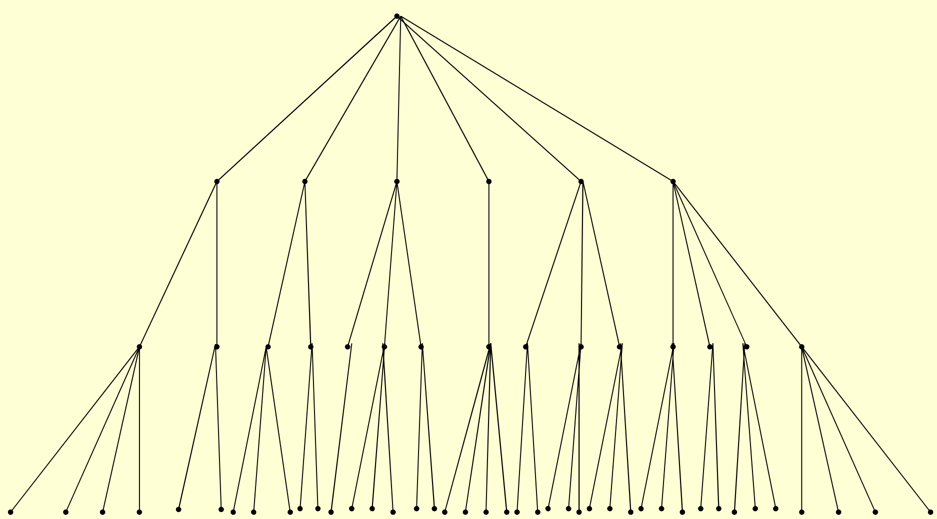
JOHNS HOPKINS MEDICINE

International Classification of Disease ICD11 Use Cases

- Scientific consensus of *clinical phenotype*
- Public Health Surveillance
  - Mortality
  - Public Health Morbidity
- Clinical data aggregation
  - Metrics of clinical activity
  - Quality management
    - Patient Safety
  - Financial administration
    - Case mix
    - Resource allocation

JOHNS HOPKINS MEDICINE

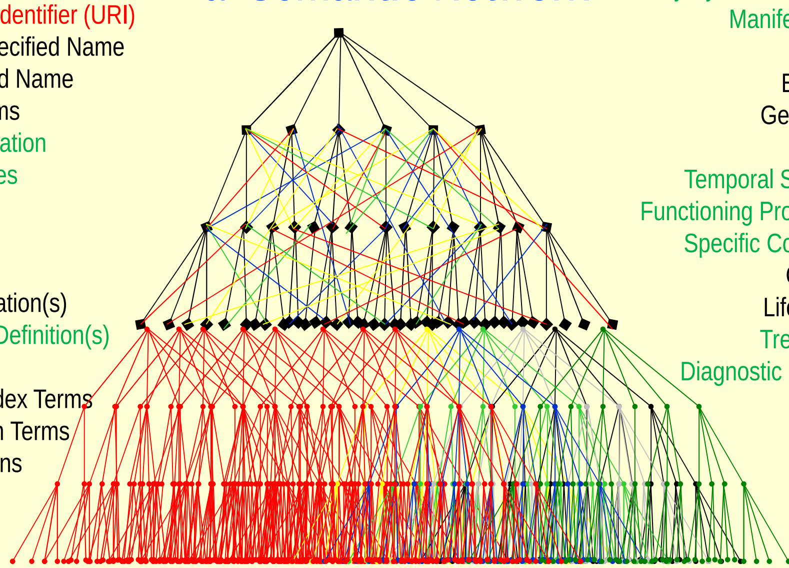
Traditional Hierarchical System ICD-10 and family



JOHNS HOPKINS MEDICINE

The ICD11 Foundation Component a Semantic Network

ICD Concept Title
Unique Identifier (URI)
Fully Specified Name
Preferred Name
Synonyms
Classification
Properties
Parents
Type
Use and
Linearization(s)
Textual Definition(s)
Terms
Base Index Terms
Inclusion Terms
Exclusions

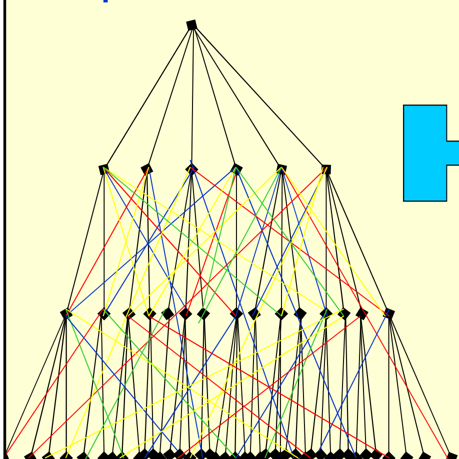


Body System/Structure
Manifestation
Causal
Etiology
Genomics
Agents
Temporal Severity
Functioning Properties
Specific Condition
Gender
Life Cycle
Treatment
Diagnostic Criteria

JOHNS HOPKINS MEDICINE

Algorithmic Serialization of the Foundation Component into a *Linearization*

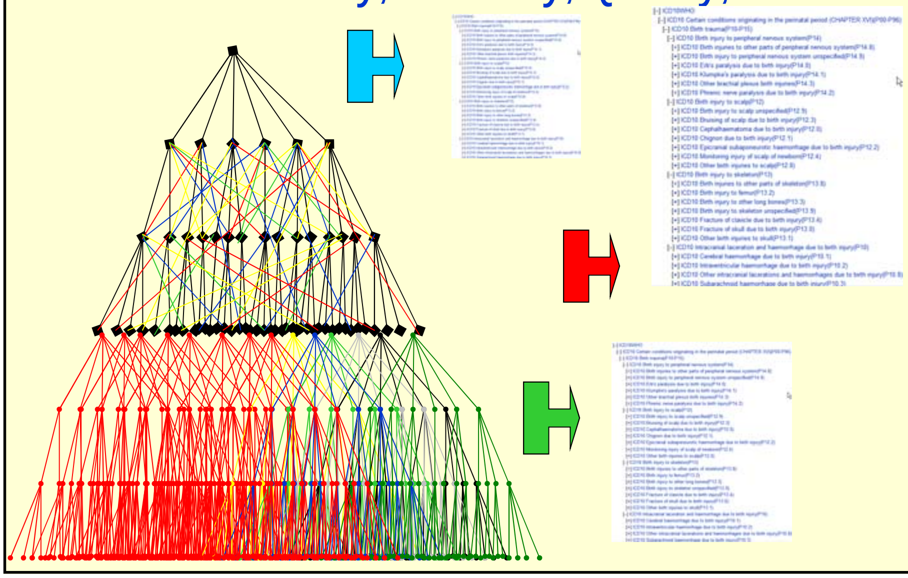
Mutually Exclusive
And Exhaustive



- [–] ICD10WHO
  - [–] ICD10 Certain conditions originating in the perinatal period (CHAPTER XVI)(P00-P96)
    - [–] ICD10 Birth trauma(P10-P15)
      - [–] ICD10 Birth injury to peripheral nervous system(P14)
        - [+] ICD10 Birth injuries to other parts of peripheral nervous system(P14.8)
        - [+] ICD10 Birth injury to peripheral nervous system unspecified(P14.9)
        - [+] ICD10 Erb's paralysis due to birth injury(P14.0)
        - [+] ICD10 Klumpke's paralysis due to birth injury(P14.1)
        - [+] ICD10 Other brachial plexus birth injuries(P14.3)
        - [+] ICD10 Phrenic nerve paralysis due to birth injury(P14.2)
      - [–] ICD10 Birth injury to scalp(P12)
        - [+] ICD10 Birth injury to scalp unspecified(P12.9)
        - [+] ICD10 Bruising of scalp due to birth injury(P12.3)
        - [+] ICD10 Cephalhaematoma due to birth injury(P12.0)
        - [+] ICD10 Chignon due to birth injury(P12.1)
        - [+] ICD10 Epicranial subaponeurotic haemorrhage due to birth injury(P12.2)
        - [+] ICD10 Monitoring injury of scalp of newborn(P12.4)
        - [+] ICD10 Other birth injuries to scalp(P12.8)
      - [–] ICD10 Birth injury to skeleton(P13)
        - [+] ICD10 Birth injuries to other parts of skeleton(P13.8)
        - [+] ICD10 Birth injury to femur(P13.2)
        - [+] ICD10 Birth injury to other long bones(P13.3)
        - [+] ICD10 Birth injury to skeleton unspecified(P13.9)
        - [+] ICD10 Fracture of clavicle due to birth injury(P13.4)
        - [+] ICD10 Fracture of skull due to birth injury(P13.0)
        - [+] ICD10 Other birth injuries to skull(P13.1)
      - [–] ICD10 Intracranial laceration and haemorrhage due to birth injury(P10)
        - [+] ICD10 Cerebral haemorrhage due to birth injury(P10.1)
        - [+] ICD10 Intraventricular haemorrhage due to birth injury(P10.2)
        - [+] ICD10 Other intracranial lacerations and haemorrhages due to birth injury(P10.8)
        - [+] ICD10 Subarachnoid haemorrhage due to birth injury(P10.3)

JOHNS HOPKINS MEDICINE

Linearizations for multiple use-cases Morbidity, Mortality, Quality, …



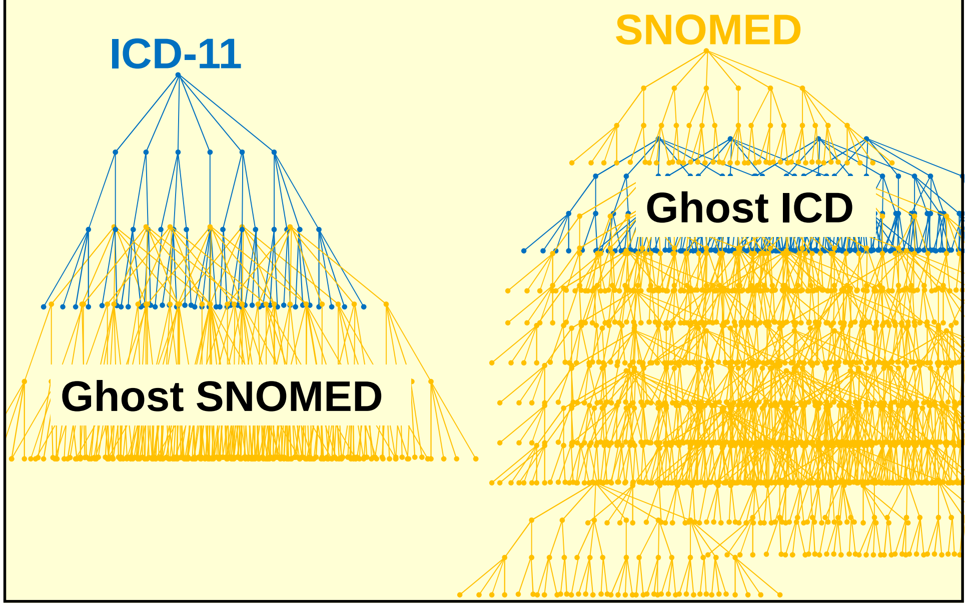
JOHNS HOPKINS MEDICINE

Relationship with IHTSDO SNOMED content

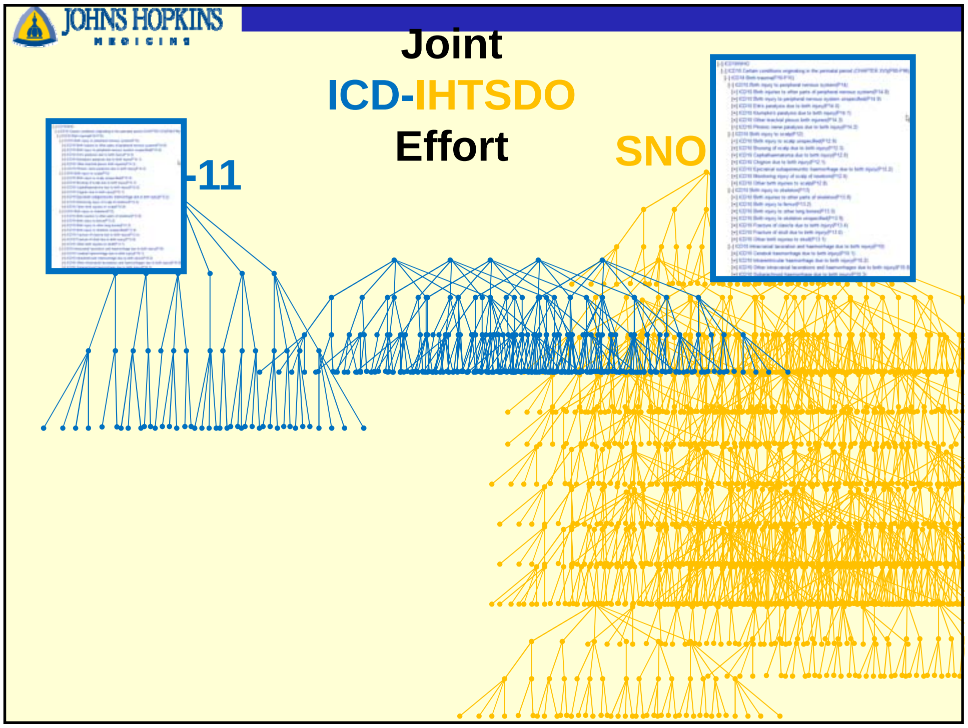
- IHT (SNOMED) will require high-level nodes that aggregate more granular data
  - Use-cases include mutually exclusive, exhaustive,…
  - Sounds a lot like ICD
- ICD-11 will require lower level terminology for value sets which populate content model
  - Detailed terminological underpinning
  - Sounds a lot like SNOMED
- Memorandum of Agreement – July 2010!
  - WHO right to use for authoring and interpretation

JOHNS HOPKINS MEDICINE

Potential Future States (2007)

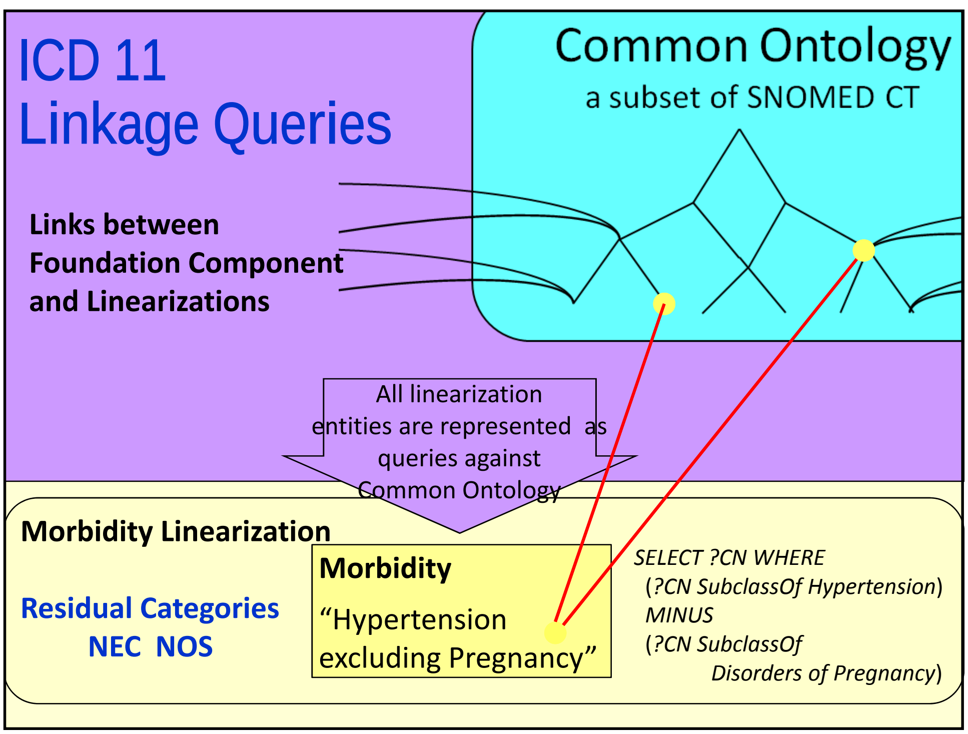


JOHNS HOPKINS MEDICINE



Common Ontology

- Joint effort between WHO and IHTSDO
- A subset of SNOMED CT
- Provides semantic anchoring of ICD11 Foundation Component
  - Semantic backbone
- All ICD11 Foundation Component elements will be defined by “query expressions” against the Common Ontology



Where is This Going?

- Biomedical practice and research are data, information, and knowledge intensive
- Comparable and consistent data representation are pre-requisite for efficient clinical analytics
- Data standards are needed to support comparable and consistent information
- ICD-11 promises to become a multi-use standard for the 21st century