Clinical Data

Amr Jamal, MD, SBFM, ABFM, MRCGP, GCMI, MBI
Assistant Professor and Consultant
Family Physician and Clinical Informatician
Head of Medical Informatics & e-Learning Unit (MIELU)
King Saud University, College of Medicine

Nasriah Zakaria

Assistant Professor Medical Informatics & e-Learning Unit (MIELU) King Saud University, College of Medicine



What are clinical data? [1]

- A datum is a single observation of a patient
- Clinical data are a collection of observations about a patient
- Each datum has five elements:
 - the patient (Amr Jamal)
 - the attribute (heart rate)
 - the value of the attribute (52 beats per minute)
 - the time of the observation (1:00 pm on 1/1/2015)
 - the method by which the attribute was obtained (heart monitor)



Types of clinical data [1]

- * Narrative: recording by clinician- maternity history
- * Numerical measurements: blood pressure, temperature
- Coded data: selection from a controlled terminology system example being the term MI that may mean myocardial infarction or mitral insufficiency
- * Textual data: other results reported as text
- * Recorded signals: EKG, EEG
- Pictures: radiographs, photographs, and other images



Use of clinical data [1]

- Form basis of historical record
- Support communication among providers
- Anticipate future health problems
- Record standard preventive measures
- Identify deviations from expected trends example being a growth chart
- Coding and billing
- Provide a legal record
- Support clinical research



Types of clinical data documents [1]

- History and physical examination:
 - by a clinician
- Progress notes
 - update of progress by primary, consulting, and ancillary providers
- Reports
 - by specialists, ancillary providers
- Typical paper chart maintains all patient notes in chronological order, sometimes separated into different components



Assessment of a stable patient [1]

- Chief complaint
- History of the present illness
- Past medical history
- Social history
- Family history
- Review of systems
- Physical examination
- Investigations –lab, x-ray, other
- Assessment plan



Some complications of data [1]

- Circumstances of observation
 - e.g., how was heart rate taken? pulse? EKG?
- Uncertainty
 - how accurate is patient reporting, measurement, device?
- **Time**
 - what level of specificity do we need?

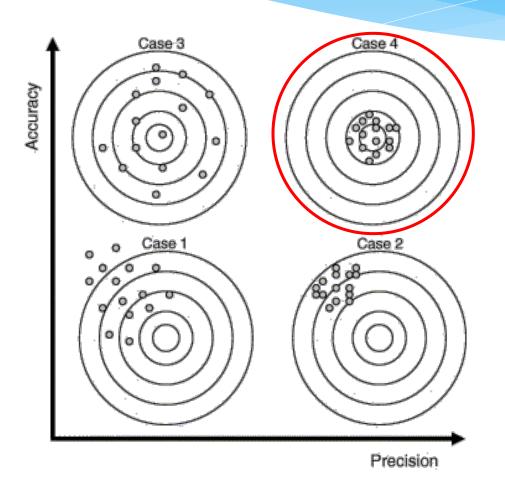


Some complications of data[2]

- Duplication
 - e.g., multiple records in different departments
- Outdated e.g. missing values
- Incorrectly formatted does not follow standards



Imprecision vs. Inaccuracy [5]





Structure of clinical data [1]

- Medicine lacks uniform structured vocabulary and nomenclature as does Physics and Chemistry
- Standardization and computerization of data is benefited by standard representations (Cimino, 2007)
- Counter-arguments are "freedom of expression" and "art of medicine"
- Narrative information when expressed in many ways can be ambiguous



Date:

Perso	nal History:				
Name:					
Age: Nationa	lity: 🗆 Saudi	☐ Other:			
	male 🖵 female	a omer.			
Occupat					
Marital	Status: 🖵 Single	married	Other		
Residen	ce:				
Admicci	on Date:				
	on Date: on type: 🖵 Emergen	cv 🗆 Elec	tive		
Ward / 1		- 2.000			
The Co	omplaint:				
aint(s)	1.	2.		3.	
(-)	-				
on					
Histo	ry of Present	ing Illne	ess:		
ion					
e					
ency					
ime					
tv					

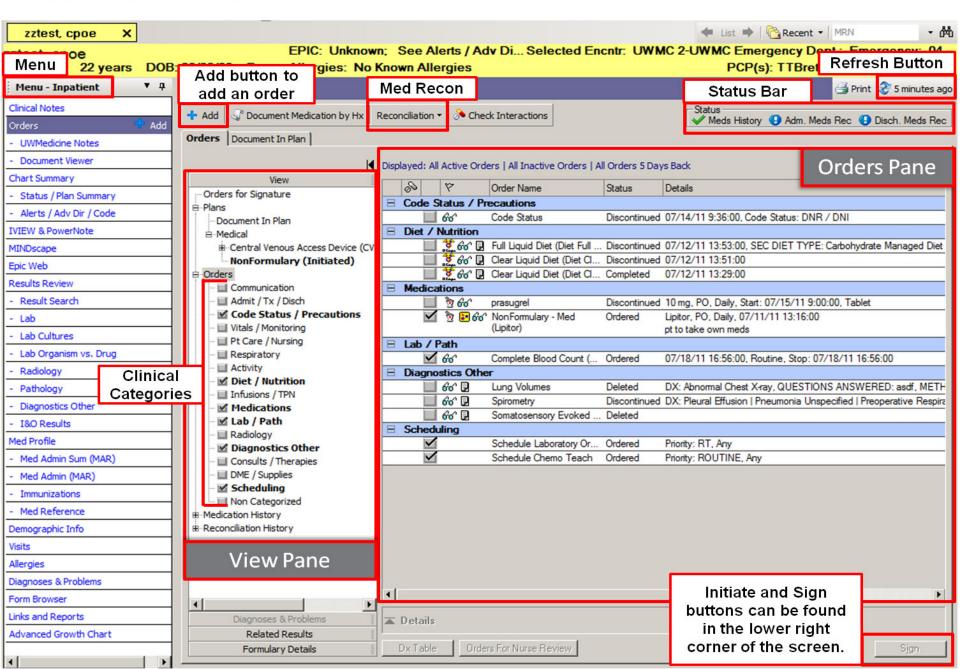
Data entry [1]

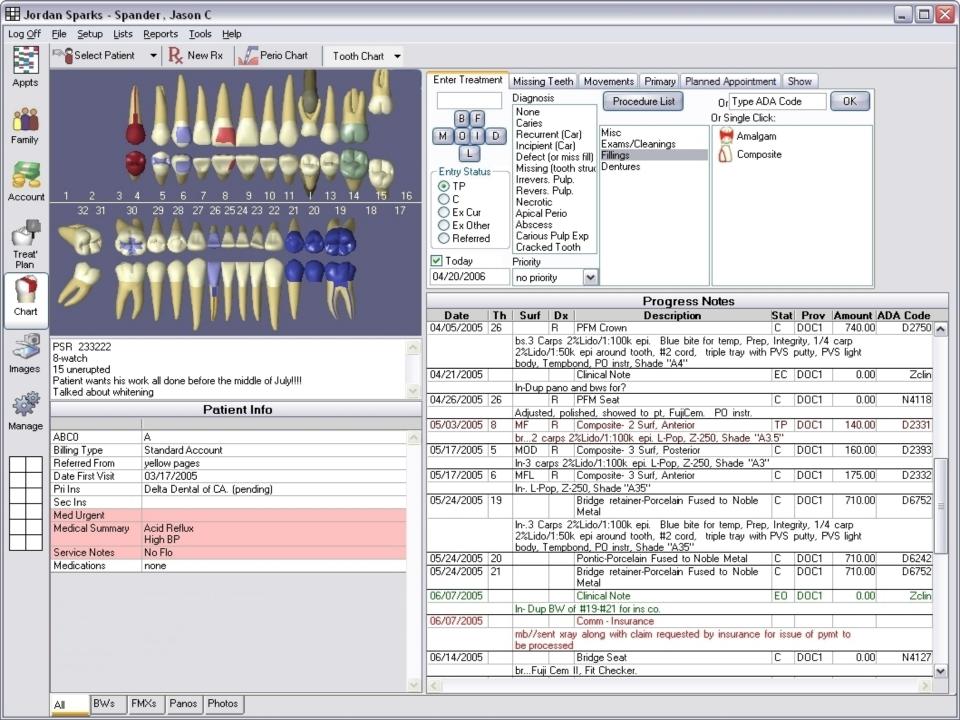
- General categories of data entry:
 - Free-form entry by historical methods:
 - writing
 - dictation
 - typing
 - Structured (menu-driven) data entry by mouse or pen
 - Speech recognition for either of above





ORCA CPOE order screen





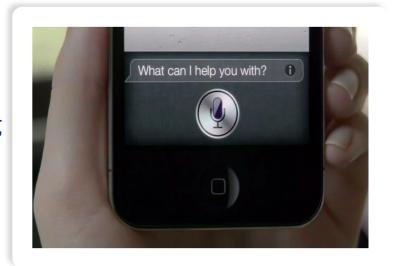
Structured or menu-driven data entry

- Many attempts from old (Greenes, 1970; Cimino, 1987; Bell, 1994) to new (Oceania; OpenSDE – Los, 2005)
- Can be done via mouse or pen, with typing
- Benefits
 - Data codified for easier retrieval and analysis
 - * Reduces ambiguity if language used consistently
- Drawbacks
 - In general, more time-consuming
 - Requires exhaustive vocabulary
 - Requires dedication to use by clinicians
- *Alternative: Processing free text with natural language processing and tagging text (in XML) (Johnson, 2008)



Speech recognition for data entry [1]

- Most common use is for narration
 - e.g., computer dictation of clinical notes
- An advantage is instant availability of dictated content
- Continuous speech recognition now is commercial reality
 - Speaker-dependent systems require user training
 - Speaker-independent are systems less accurate
- Many established systems on the market that operate on:
 - front-end (used by clinician) or
 - back-end (process dictations) (Brown, 2008)









Coded vs. free-text data [1]

* Coded data:

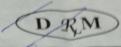
Documentation of discrete data from controlled vocabulary

* Free text:

Alphanumeric data that are unstructured, typically in narrative form



DE RMA CLINICS
DERMATOLOGY - PLASTIC SURGERY
HAIR TRANSPLANT - LASER CENTER



مراش جلدية جراحة تهميل وراش جلدية جراحة تهميل وراعة شعر مركز العلاج بالليزو

وصفةطبية

QB APR 2013

Prescription

Rx			1
6	no	c	in
		e ,	1001
A			

Narratives tell a story.

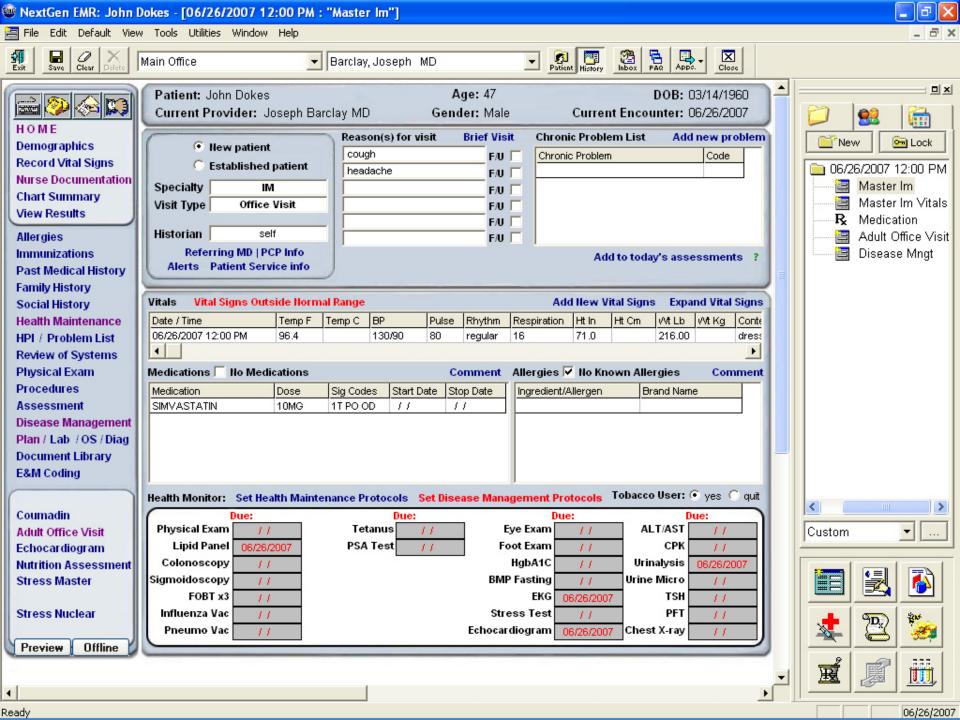
A narrative tells a story

- See the patient through a description
- Complicated events are easier to describe in text

Undifferentiated problems

- Interpretation.
 - "only a human can prioritize and determine what the chief complaint really is"





Liverpool Hospital Neonatal Database - Clinical Data Entry Ian TEST DOB 28/2/2008 GA 26+2 BW 1070 Liverpool 2170 MRN 123432 Day 33 - Corrected GA 31+0 1250g on 01/04 Files (0) Calculator Log Images (2) ATTENTION: Brain scan overdue: Admissions Respiratory Nutrition Other Treatments Test Results Current Status Respiratory Support Admission Planning Discharge CPAP /5 , FiO2 29 Admitted: 28/02/08 at 4 hours Liverpool Hospital 76% HC 25.5 71% Length 35 Weight 1070 Admission Corr.GA 26 Age 0 Fluids / Feeds MRN 123432 Date & Time 28/02/2008 16:30 | Hospital Liverpool Hospital 160 ml/kg/day TPN 10% Fat 3q NICU Prematurity Bed Reason(s) for 14x2 EBM 24cal (134) Admission Consultant Ian Callander Insurance Hospital Respiratory Distress Jaundice 09/03 SBr 135 Biliblanket MATERNAL HISTORY ceased 08/03 Ann is a 28 year old G2 P1 (now) woman whose blood group is O positive. She was booked to deliver at Campbelltown Hospital Other under the care of Kaisher however delivered at Liverpool Hospital under the care of Dr Peter Hammill. She had a history of essential 01/03 Mod PDA hypertension. This pregnancy was complicated by hypertension of pregnancy, fetal growth restriction, Bilateral Renal Pelvis dilatation POSSIBLE NEC 5 - 10mm, GBS +ve swab, fever, abnormal Dopplers, prolonged rupture of membranes for 2 days, clinically suspected chorioamnionitis. Ann was treated with antenatal steroids, tocolytics, and antihypertensive drugs. Following the spontaneous onset of labour, she proceeded to a vaginal delivery. Antibiotics were given before delivery. Treatments PERINATAL HISTORY Pentavite, Folic Acid Ian was born at 13:00 hours with a birth weight of 1070 grams (76th centile). Appars were 3 at 1 minute and 7 at 5 minutes Longline, respectively treated with intubation and ventilation. The arterial cord pH was 7.24 and the base excess -6, Ian was then retrieved to Added to Worksheet 01/03 Orders on Worksheet 01/03 This is freetext Freetext orders (double click on Test Results text to delete) 09/03 Na 136 09/03 Hb 135 09/03 Plat 265 Hospital Episodes MRN Admitted Discharged Add Another Admission 02/03 HUS IVH II Liverpool Hospital 123432 28 Feb 2008 16:30 01/04 Eyes ROP I PD12345 | 28 Feb 2008 15:00 | 28 Feb 2008 16:30 **NETS** Opened 01 Apr 12:27 Delete MRN .. then click again to Campbelltown Hospital 28 Feb 2008 13:00 | 28 Feb 2008 15:00 222222 Delete Episode Add Twin 1 local form

Issues with coded data

- "pick from a list" allows wrong selection
- compliance concerns
- over documentation for care
- cloning



Coded clinical data enables EHR advanced functionality

- Alerts
- Clinical Decision Support
- Best documentation practices
- Multi-media reporting
- Multiple output formats
- Data mining



Data Management [2]

File Organization concepts

- Database: A set of related files
- File: Collection of records of same type
- Record: A set of related field
- Field: Words and numbers



Registration Radiology Database Financial Name **Medical Summary** Age File Abdullah AlSaif BP 21 34 Acid Reflux Khalid AlQahtani Maryam Badr 42 Pneumonia 32 Reem Alowais Allergies Record Name **Medical Summary** Age Abdullah AlSaif BP 21 Field BP (Medical summary field)

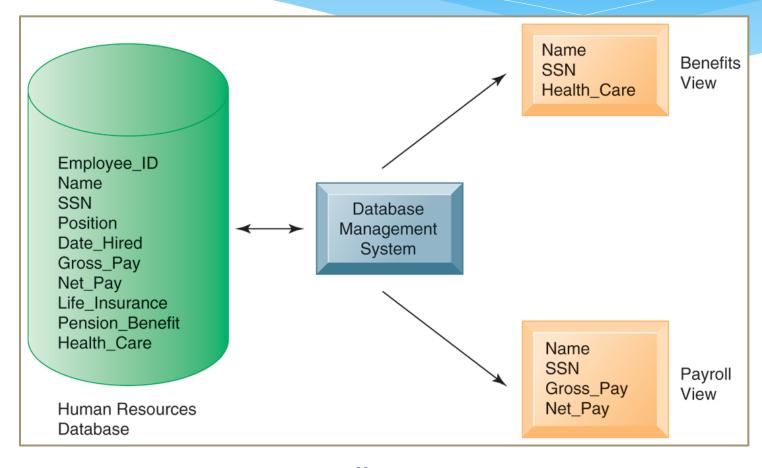


Relational DBMS [1]

- Relational model links records to tables
- Allows efficiencies
- One-time information (e.g., demographics) stored only once
- Complex queries easier to construct and carry out
- ❖ Most query capabilities are based on **Structured Query Language (SQL)**-special language in relational database



Relational DBMS [2]





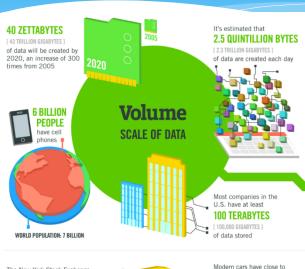


Big Data [3]

- Science of Data Management & analysis
- * "to convert Vast information and knowledge in organisation to achieve their objectives" (Murdoch et al, 2013*)
- What is BIG/VAST? Zettabytes (10^21 gigabytes) to Yottabytes (10^24 gigabytes)
- Used in Astronomy, Search Engines, Financial, Politics and now in Biomedicine
- Example of Big Data is Bioinformatics (genome, proteomic)



The FOUR V's of Big Data [3]



The New York Stock Exchange captures 1 TB OF TRADE

100 SENSORS

that monitor items such as uel level and tire pressure

Velocity

By 2016, it is projected there will be 18.9 BILLION **NETWORK** CONNECTIONS

during each trading session

- almost 2.5 connections per person on earth

ANALYSIS OF STREAMING DATA



The FOUR V's of Big Data

break big data into four dimensions: Volume, **Velocity, Variety and Veracity**

4.4 MILLION IT JOBS



As of 2011, the global size of data in healthcare was estimated to be



Variety DIFFERENT

30 BILLION PIECES OF CONTENT

are shared on Facebook every month









By 2014, it's anticipated there will be

420 MILLION WEARABLE, WIRELESS **HEALTH MONITORS**

4 BILLION+

are watched on YouTube each month



are sent per day by about 200 million monthly active users

1 IN 3 BUSINESS

don't trust the information they use to make decisions



in one survey were unsure of how much of their data was



Poor data quality costs the US economy around

\$3.1 TRILLION A YEAR



Veracity UNCERTAINTY OF DATA





Big Data in healthcare [3]

- * "80% of medical data is unstructured and is clinically relevant.
- The data reside in multiple places like individual EMRs, lab and imaging systems, physician notes, medical correspondence, claims, customer relations management systems and finance."



Sources of BIG DATA [4]

- Clinical Data from CPOE
- Clinical decision support systems (Written notes & prescriptions)
- Imaging systems: PACS, Radiology Information systems
- Sensor data (monitoring vital signs)
- Social media data- Tweets from Twitter, wall and status updates on Facebook
- Emergency care data
- Literature from medical journal



Healthcare BIG data problems to be solved [4]

- * Patient profiles and the health outcomes- identify the effective treatments
- * For public health- identify individuals who would get preventive care or lifestyle changes
- * Analysing literature on medical procedure to determining which care protocols work best
- * Creating mobile apps to manage diabetes. Via Data analytics, we are able to monitor the healthcare outcomes improvements
- * Analysing social network communication among support group members- to understand how non-profit organization can interact and provide help



In summary,

- Types of clinical data
- Types of clinical data documents
- Use of clinical data
- Access to clinical data
- Data entry
- Coded vs. free-form data
- Speech recognition
- Big Data
- Database Management



Acknowledgement



* Notes are <u>adapted with permission</u> from Professor Hersh, Oregon Health and Science University (OHSU), Oregon, USA



References



- [1] Hersh, W. (2014). Notes from 10x10 Medical Informatics certificate, Oregon Health & Science University
- [2] Laudon & Laudon (2011), Management Information Systems, Prentice Hall
- [3] IBM website:http://www-o1.ibm.com/software/data/bigdata/what-is-bigdata.html
- [4] Ragupathi W. & Ragupathi V. (2014). Big Data Analytics in Healthcare: Promise and Potential. Health Information Science and Systems

http://www.hissjournal.com/content/2/1/3

[5] http://www.sophia.org/tutorials/accuracy-and-precision--3

