

# 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

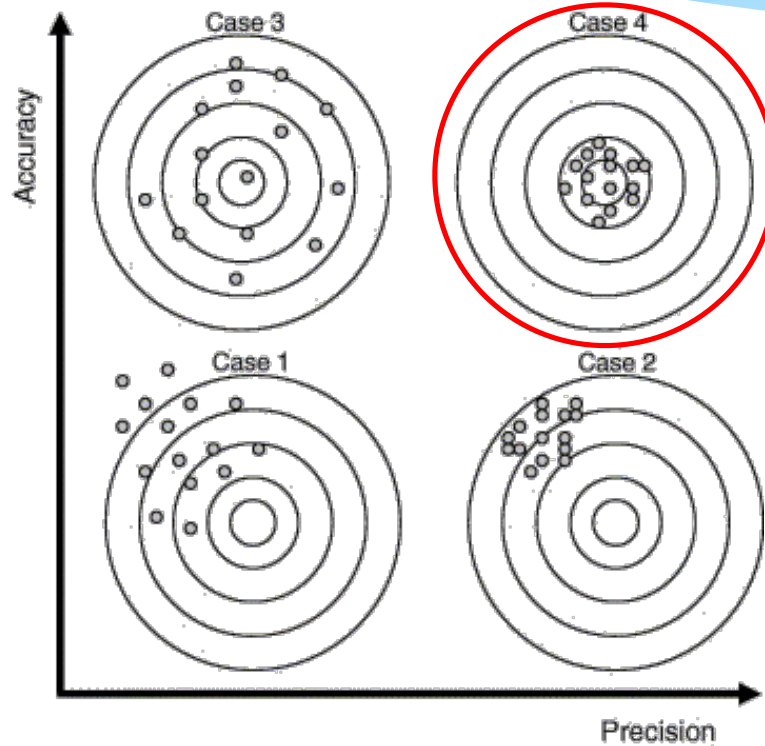
e.g. missing values

## † 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

**Personal History:**

Name:

Age:

Nationality:  Saudi  Other:

Sex:  male  female

Occupation:

Marital Status:  Single  married  Other:

Residence:

Admission Date:

Admission type:  Emergency  Elective

Ward / Bed:

**The Complaint:**

Complaint(s)

1.	2.	3.

Duration

**History of Presenting Illness:**

Onset


Duration


Onset


Duration


Onset

# We need better access to clinical data [1]

- † Missing clinical information during primary care visits (Smith, 2005)
  - † Information reported missing in 13.6% of clinical visits
    - † Available but outside system in 52% of instances
    - † Estimated to adversely effect patients 44% of time
    - † Unsuccessful searching for it took >5 minutes 35% of time
- † Physicians have two unmet information needs for every three patients (Gorman, 1995; Ely, 1999)
- † Secondary use of clinical data (Safran, 2007)

# 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

zztest.cpoe X

EPIC: Unknown: See Alerts / Adv Di... Selected Encntr: UWMC 2-UWMC Emergency Dept: Emergency: 04  
22 years DOB: [redacted] Allergies: No Known Allergies PCP(s): TTBre

**Menu** | **Menu - Inpatient** | **Add button to add an order** | **Med Recon** | **Status Bar** | **Refresh Button**

Orders | Document In Plan | Reconciliation | Check Interactions | Print | 5 minutes ago

Orders Pane

View

- Orders for Signature
- Plans
  - Document In Plan
  - Medical
    - Central Venous Access Device (CV)
    - NonFormulary (Initiated)**
  - Orders
    - Communication
    - Admit / Tx / Disch
    - Code Status / Precautions**
    - Vitals / Monitoring
    - Pt Care / Nursing
    - Respiratory
    - Activity
    - Diet / Nutrition**
    - Infusions / TPN
    - Medications**
    - Lab / Path**
    - Radiology
    - Diagnostics Other**
    - Consults / Therapies
    - DME / Supplies
    - Scheduling**
    - Non Categorized
  - Medication History
  - Reconciliation History

**Clinical Categories**

View Pane

Order Name	Status	Details
<b>Code Status / Precautions</b>		
<input type="checkbox"/> 60' Code Status	Discontinued	07/14/11 9:36:00, Code Status: DNR / DNI
<b>Diet / Nutrition</b>		
<input type="checkbox"/> 60' Full Liquid Diet (Diet Full ...	Discontinued	07/12/11 13:53:00, SEC DIET TYPE: Carbohydrate Managed Diet
<input type="checkbox"/> 60' Clear Liquid Diet (Diet Cl...	Discontinued	07/12/11 13:51:00
<input type="checkbox"/> 60' Clear Liquid Diet (Diet Cl...	Completed	07/12/11 13:29:00
<b>Medications</b>		
<input type="checkbox"/> 60' prasugrel	Discontinued	10 mg, PO, Daily, Start: 07/15/11 9:00:00, Tablet
<input checked="" type="checkbox"/> 60' NonFormulary - Med (Lipitor)	Ordered	Lipitor, PO, Daily, 07/11/11 13:16:00 pt to take own meds
<b>Lab / Path</b>		
<input checked="" type="checkbox"/> 60' Complete Blood Count (...	Ordered	07/18/11 16:56:00, Routine, Stop: 07/18/11 16:56:00
<b>Diagnostics Other</b>		
<input type="checkbox"/> 60' Lung Volumes	Deleted	DX: Abnormal Chest X-ray, QUESTIONS ANSWERED: asdf, METH...
<input type="checkbox"/> 60' Spirometry	Discontinued	DX: Pleural Effusion   Pneumonia Unspecified   Preoperative Respir...
<input type="checkbox"/> 60' Somatosensory Evoked ...	Deleted	
<b>Scheduling</b>		
<input checked="" type="checkbox"/> Schedule Laboratory Or...	Ordered	Priority: RT, Any
<input checked="" type="checkbox"/> Schedule Chemo Teach	Ordered	Priority: ROUTINE, Any

**Initiate and Sign buttons can be found in the lower right corner of the screen.**

Sign

Select Patient **Rx** New Rx Perio Chart Tooth Chart

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16  
32 31 30 29 28 27 26 25 24 23 22 21 20 19 18 17

PSR 233222  
8-watch  
15 unerupted  
Patient wants his work all done before the middle of July!!!!  
Talked about whitening

**Patient Info**

ABC0	A
Billing Type	Standard Account
Referred From	yellow pages
Date First Visit	03/17/2005
Pri Ins	Delta Dental of CA. (pending)
Sec Ins	
Med Urgent	
Medical Summary	Acid Reflux High BP
Service Notes	No Flo
Medications	none

Enter Treatment Missing Teeth Movements Primary Planned Appointment Show

Diagnosis:  Procedure List  Or Type ADA Code  OK

None  
Caries  
Recurrent (Car)  
Incipient (Car)  
Defect (or miss fill)  
Missing (tooth struc)  
Irrevers. Pulp.  
Revers. Pulp.  
Necrotic  
Apical Perio  
Abscess  
Carious Pulp Exp  
Cracked Tooth

Entry Status:  
 TP  
 C  
 Ex Cur  
 Ex Other  
 Referred

Misc  
Exams/Cleanings  
Fillings  
Dentures

Or Single Click:  
 Amalgam  
 Composite

Today Priority  
04/20/2006 no priority

**Progress Notes**

Date	Th	Surf	Dx	Description	Stat	Prov	Amount	ADA Code
04/05/2005	26		R	PFM Crown	C	DOC1	740.00	D2750
				bs.3 Carps 2%Lido/1:100k epi. Blue bite for temp, Prep, Integrity, 1/4 carp 2%Lido/1:50k epi around tooth, #2 cord, triple tray with PVS putty, PVS light body, Tempbond, PD instr, Shade "A4"				
04/21/2005				Clinical Note	EC	DOC1	0.00	Zclin
				In-Dup pano and bws for?				
04/26/2005	26		R	PFM Seat	C	DOC1	0.00	N4118
				Adjusted, polished, showed to pt, FujiCem. PD instr.				
05/03/2005	8	MF	R	Composite- 2 Surf, Anterior	TP	DOC1	140.00	D2331
				br...2 carps 2%Lido/1:100k epi. L-Pop, Z-250, Shade "A3.5"				
05/17/2005	5	MDD	R	Composite- 3 Surf, Posterior	C	DOC1	160.00	D2393
				In-3 carps 2%Lido/1:100k epi. L-Pop, Z-250, Shade "A3"				
05/17/2005	6	MFL	R	Composite- 3 Surf, Anterior	C	DOC1	175.00	D2332
				In. L-Pop, Z-250, Shade "A35"				
05/24/2005	19			Bridge retainer-Porcelain Fused to Noble Metal	C	DOC1	710.00	D6752
				In-3 Carps 2%Lido/1:100k epi. Blue bite for temp, Prep, Integrity, 1/4 carp 2%Lido/1:50k epi around tooth, #2 cord, triple tray with PVS putty, PVS light body, Tempbond, PD instr, Shade "A35"				
05/24/2005	20			Pontic-Porcelain Fused to Noble Metal	C	DOC1	710.00	D6242
05/24/2005	21			Bridge retainer-Porcelain Fused to Noble Metal	C	DOC1	710.00	D6752
06/07/2005				Clinical Note	EO	DOC1	0.00	Zclin
				In-Dup BW of #19-#21 for ins co.				
06/07/2005				Comm - Insurance				
				mb//sent xray along with claim requested by insurance for issue of pymt to be processed				
06/14/2005				Bridge Seat	C	DOC1	0.00	N4127
				br...Fuji Cem II, Fit Checker.				

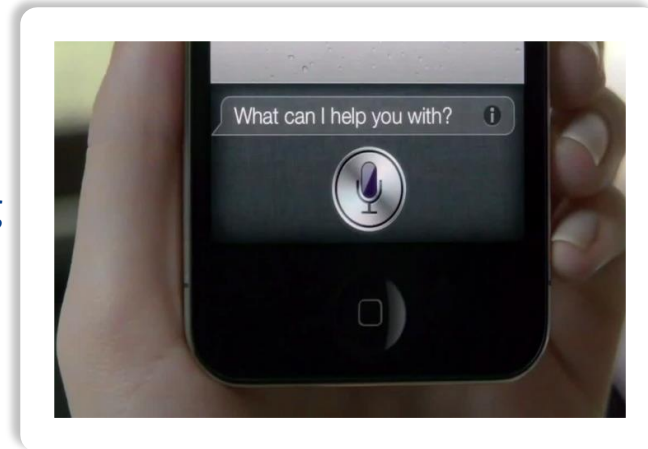
# Structured or menu-driven data entry

- † 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



# 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)



iPad 1:47 PM 99%

**Katelyn Gleason**

Chart ID: GLKA000007 Gender: F Age: 25 DoB: 02/15/1986 555-555-5555

Chief Complaint: unspecified pain or illness

Temperature: 98.0 °F Pulse: 60 bpm Blood Pressure: 110 / 65 mmHg Respiratory Rate: 20 rpm Oxygen Saturation: 94 %

Height: 68 in Weight: 130 lbs BMI: 19.76 Pain (1-10): 2 Smoking Status: Never Smoker

**Clinical Checklist**

General WNL

General Comments

HEENT WNL

HEENT Comments

Skin WNL

Skin Comments

Speech to text in progress

0:00:08

Tap to end

coarse hair,

Speech to Text

dr chrono Lock New eRx eRx Refills Chat Help



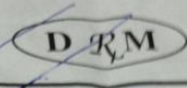
# Coded vs. free-text data [1]

## **Coded data:**

Documentation of discrete data from controlled vocabulary

## **Free text:**

Alphanumeric data that are unstructured, typically in narrative form



وصفة طبية  
Prescription

13 APR 2013

Date : .....

Name : Mamsha al namin M.R. No. : 134415

Age : .....  M  F

Diagnosis : .....

R  
x

*Elan / ...*

*1001*

Doct

Signa



# 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”

**Patient:** John Dokes **Age:** 47 **DOB:** 03/14/1960  
**Current Provider:** Joseph Barclay MD **Gender:** Male **Current Encounter:** 06/26/2007

**New patient**  
 **Established patient**

**Specialty** IM  
**Visit Type** Office Visit  
**Historian** self

[Referring MD | PCP Info](#)  
[Alerts](#) [Patient Service info](#)

**Reason(s) for visit** **Brief Visit** **Chronic Problem List** **Add new problem**

cough	F/U	<input type="checkbox"/>	Chronic Problem	Code
headache	F/U	<input type="checkbox"/>		
	F/U	<input type="checkbox"/>		
	F/U	<input type="checkbox"/>		
	F/U	<input type="checkbox"/>		

[Add to today's assessments ?](#)

**Vitals** **Vital Signs Outside Normal Range** [Add New Vital Signs](#) [Expand Vital Signs](#)

Date / Time	Temp F	Temp C	BP	Pulse	Rhythm	Respiration	Ht In	Ht Cm	Wt Lb	Wt Kg	Cont
06/26/2007 12:00 PM	96.4		130/90	80	regular	16	71.0		216.00		dress

**Medications**  **No Medications** **Comment** **Allergies**  **No Known Allergies** **Comment**

Medication	Dose	Sig Codes	Start Date	Stop Date	Ingredient/Allergen	Brand Name
SIMVASTATIN	10MG	1T PO OD	!!	!!		

**Health Monitor:** [Set Health Maintenance Protocols](#) [Set Disease Management Protocols](#) **Tobacco User:**  yes  quit

	Due:		Due:		Due:		Due:
Physical Exam	!!	Tetanus	!!	Eye Exam	!!	ALT/AST	!!
Lipid Panel	06/26/2007	PSA Test	!!	Foot Exam	!!	CPK	!!
Colonoscopy	!!			HgbA1C	!!	Urinalysis	06/26/2007
Sigmoidoscopy	!!			BMP Fasting	!!	Urine Micro	!!
FOBT x3	!!			EKG	06/26/2007	TSH	!!
Influenza Vac	!!			Stress Test	!!	PFT	!!
Pneumo Vac	!!			Echocardiogram	06/26/2007	Chest X-ray	!!

- HOME
- Demographics
- Record Vital Signs
- Nurse Documentation
- Chart Summary
- View Results

- Allergies
- Immunizations
- Past Medical History
- Family History
- Social History
- Health Maintenance
- HPI / Problem List
- Review of Systems
- Physical Exam
- Procedures
- Assessment
- Disease Management
- Plan / Lab / OS / Diag
- Document Library
- E&M Coding

- Coumadin
- Adult Office Visit
- Echocardiogram
- Nutrition Assessment
- Stress Master
- Stress Nuclear

Preview Offline

Navigation pane with folders and icons:

- New
- Lock
- 06/26/2007 12:00 PM
  - Master Im
  - Master Im Vitals
  - Medication
  - Adult Office Visit
  - Disease Mngt
- Custom
- Grid of icons for various medical functions (e.g., chart, notes, vitals, stress, etc.)

**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

Log Files (0) Images (2) Calculator

ATTENTION: Brain scan overdue:

Admissions Respiratory Nutrition Other Treatments Test Results

Admission Planning Discharge

Liverpool Hospital

Admitted: 28/02/08 at 4 hours

Admission Age 0 Corr.GA 26 Weight 1070 76% HC 25.5 71% Length 35 56%

Date & Time 28/02/2008 16:30 Hospital Liverpool Hospital MRN 123432  
 Bed 01 To NICU Reason(s) for Admission Prematurity  
 Consultant Ian Callander Insurance Hospital Respiratory Distress

**Fluids / Feeds**

160 ml/kg/day  
 TPN 10% Fat 3g  
 14x2 EBM 24cal (134)

**Joundice**

09/03 SBr 135 Bilblanket  
 ceased 08/03

**Other**

01/03 Mod PDA  
 POSSIBLE NEC

**Treatments**

Pentavite, Folic Acid  
 Longline,

**Test Results**

09/03 Na 136  
 09/03 Hb 135  
 09/03 Plat 265  
 02/03 HUS IVH II  
 01/04 Eyes ROP I  
 Opened 01 Apr 12:27

**MATERNAL HISTORY**

Ann is a 28 year old G2 P1 (now) woman whose blood group is O positive. She was booked to deliver at Campbelltown Hospital under the care of Kaisher however delivered at Liverpool Hospital under the care of Dr Peter Hammill. She had a history of essential hypertension. This pregnancy was complicated by hypertension of pregnancy, fetal growth restriction, Bilateral Renal Pelvis dilatation 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.

**PERINATAL HISTORY**

Ian was born at 13:00 hours with a birth weight of 1070 grams (76th centile). Apgars were 3 at 1 minute and 7 at 5 minutes 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 text to delete)

Hospital Episodes	MRN	Admitted	Discharged
Liverpool Hospital	123432	28 Feb 2008 16:30	
NETS	PD12345	28 Feb 2008 15:00	28 Feb 2008 16:30
Campbelltown Hospital	222222	28 Feb 2008 13:00	28 Feb 2008 15:00

Add Another Admission

Delete MRN .. then click again to Delete Episode

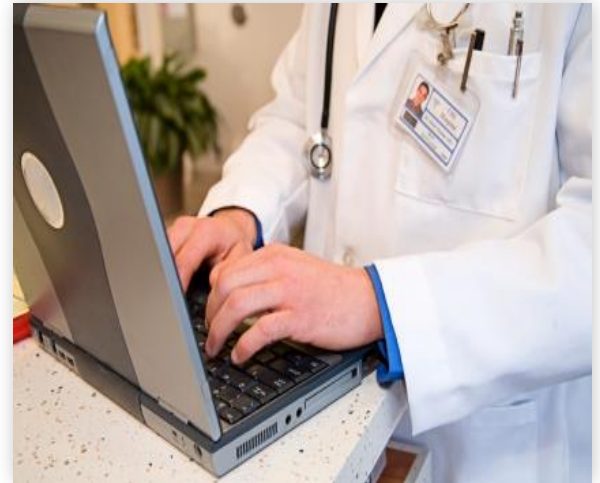
Add Twin

local form



# Issues with coded data

- † “pick from a list” allows wrong selection
- † compliance concerns
- † over documentation for care
- † cloning





# 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

Database

Radiology

Registration

Financial

File

Name	Age	Medical Summary
Abdullah AlSaif	21	BP
Khalid AlQahtani	34	Acid Reflux
Maryam Badr	42	Pneumonia
Reem Alowais	32	Allergies

Record

Name	Age	Medical Summary
Abdullah AlSaif	21	BP

Field

BP (Medical summary field)



# Relational DBMS [1]

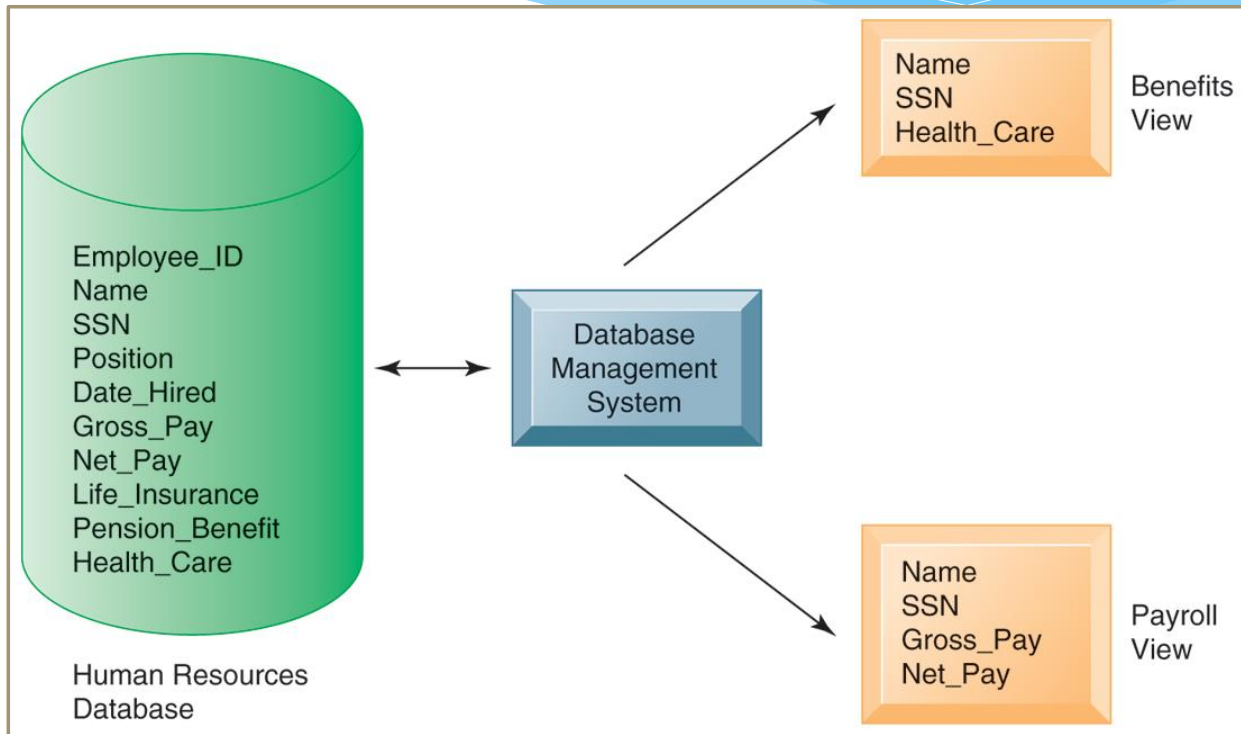
† Relational model links records to tables

† 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)

\*Murdoch, T. , Detsky, A. (2013) The Inevitable Application of Big Data to Health Care  
*JAMA*. 2013;309(13):1351-1352. doi:10.1001/jama.2013.393.



# The FOUR V's of Big Data [3]

## 40 ZETTABYTES

(43 TRILLION GIGABYTES) of data will be created by 2020, an increase of 300 times from 2005



## Volume SCALE OF DATA

It's estimated that **2.5 QUINTILLION BYTES** (2.3 TRILLION GIGABYTES) of data are created each day



Most companies in the U.S. have at least **100 TERABYTES** (100,000 GIGABYTES) of data stored



**6 BILLION PEOPLE** have cell phones



WORLD POPULATION: 7 BILLION

## The FOUR V's of Big Data

From traffic patterns and music downloads to web history and medical records, data is recorded, stored, and analyzed to enable the technology and services that the world relies on every day. But what exactly is big data, and how can these massive amounts of data be used?

As a leader in the sector, IBM data scientists break big data into four dimensions: **Volume, Velocity, Variety and Veracity**

Depending on the industry and organization, big data encompasses information from multiple internal and external sources such as transactions, social media, enterprise content, sensors and mobile devices. Companies can leverage data to adapt their products and services to better meet customer needs, optimize operations and infrastructure, and find new sources of revenue.

By 2015, **4.4 MILLION IT JOBS** will be created globally to support big data, with 1.9 million in the United States



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

**150 EXABYTES** (161 BILLION GIGABYTES)

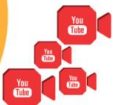


**30 BILLION PIECES OF CONTENT** are shared on Facebook every month



By 2014, it's anticipated there will be **420 MILLION WEARABLE, WIRELESS HEALTH MONITORS**

**4 BILLION+ HOURS OF VIDEO** are watched on YouTube each month



**400 MILLION TWEETS** are sent per day by about 200 million monthly active users

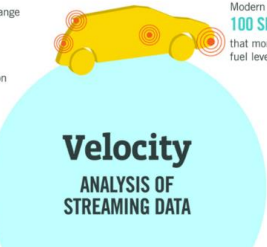


## Variety DIFFERENT FORMS OF DATA

The New York Stock Exchange captures **1 TB OF TRADE INFORMATION** during each trading session



## Velocity ANALYSIS OF STREAMING DATA



By 2016, it is projected there will be **18.9 BILLION NETWORK CONNECTIONS** - almost 2.5 connections per person on earth

Modern cars close to **100 SENSORS** that monitor items such as fuel level and tire pressure



**1 IN 3 BUSINESS LEADERS** don't trust the information they use to make decisions



Poor data quality costs the US economy around **\$3.1 TRILLION A YEAR**



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

## 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 **adapted with permission** 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-01.ibm.com/software/data/bigdata/what-is-big-data.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>