



Data Science in Healthcare

Juan F. Arias

2022 ECC Conference

June 12 - 14

Outline

Introduction

Data Science in Healthcare:

- Data Capture
- Data Cleanup
- Data Storage
- Data Analysis
- Predictive Analytics

Questions









MEDICAL REPORT
 02-08-38 MALE

: 02 : 43 : 080
 : 586 : 89 : 403
 : 253 : 684 : 01
 : 99 : RP : 809






Data Capture







Data Capture

- Genetic Attributes
- Demographic Attributes
- Clinical Attributes
- Wellness Attributes
- Administrative Attributes
- Aggregate Attributes





EXPLANATION OF BENEFITS

THIS IS NOT A BILL

Date: 06/20/13
Policy: 101-112-208

Provider of Service	Billed Charges	Not Covered	Deductible	Copy	Total Paid
HOSPITAL 5/20/13- 5/20/13	91.00 20.76	4-11-208 101-112-208			

If you have any questions, please write or call our Customer Service Department at 1-800-XXX-XXXX.

ISA*00* *00* *12*SENDERID
*14*0073268795005 *960807*1548*U*00401*000000020*0*T*>~
GS*PO*SENDERID*007326879*19960807*1548*000001*X*004010~
ST*850*000000001~
BEG*00*SA*A99999-01**19971207~
REF*VR*54321~
ITD*01*3*1.000**15**16~
DTM*002*19971219~
N1*BT*BUYSNACKS INC.*9*1223334444~
N3*P.O. BOX 0000~
N4*TEMPLE*TX*76503~
N1*ST*BUYSNACKS PORT*9*1223334445~
N3*1000 N. SAMPLE HIGHWAY~
N4*ATHENS*GA*30603~
PO1**16*CA*12.34**CB*000111111*UA*002840022222~
PID*F****CRUNCHY CHIPS LSS~
PO4*48*7.89*LB~
PO1**13*CA*12.34**CB*000555555*UA*002840033333~
PID*F****NACHO CHIPS LSS~
PO4*48*8.9*LB~
PO1**32*CA*12.34**CB*000666666*UA*002840044444~
PID*F****POTATO CHIPS~
PO4*72*6.78*LB~
PO1**51*CA*12.34**CB*000874917*UA*002840055555~
PID*F****CORN CHIPS~
PO4*48*8.9*LB~
PO1**9*CA*12.34**CB*000874958*UA*002840066666~
PID*F****BBQ CHIPS~
PO4*48*4.5*LB~
PO1**85*CA*12.34**CB*000874990*UA*002840077777~
PID*F****GREAT BIG CHIPS LSS~
PO4*48*4.56*LB~
PO1**1*CA*12.34**CB*000875088*UA*002840088888~
PID*F****MINI CHIPS LSS~
PO4*48*4.56*LB~
CTT*7~
SE*34*000000001~
GE*1*000001~
IEA*1*000000020~



Data Cleanup

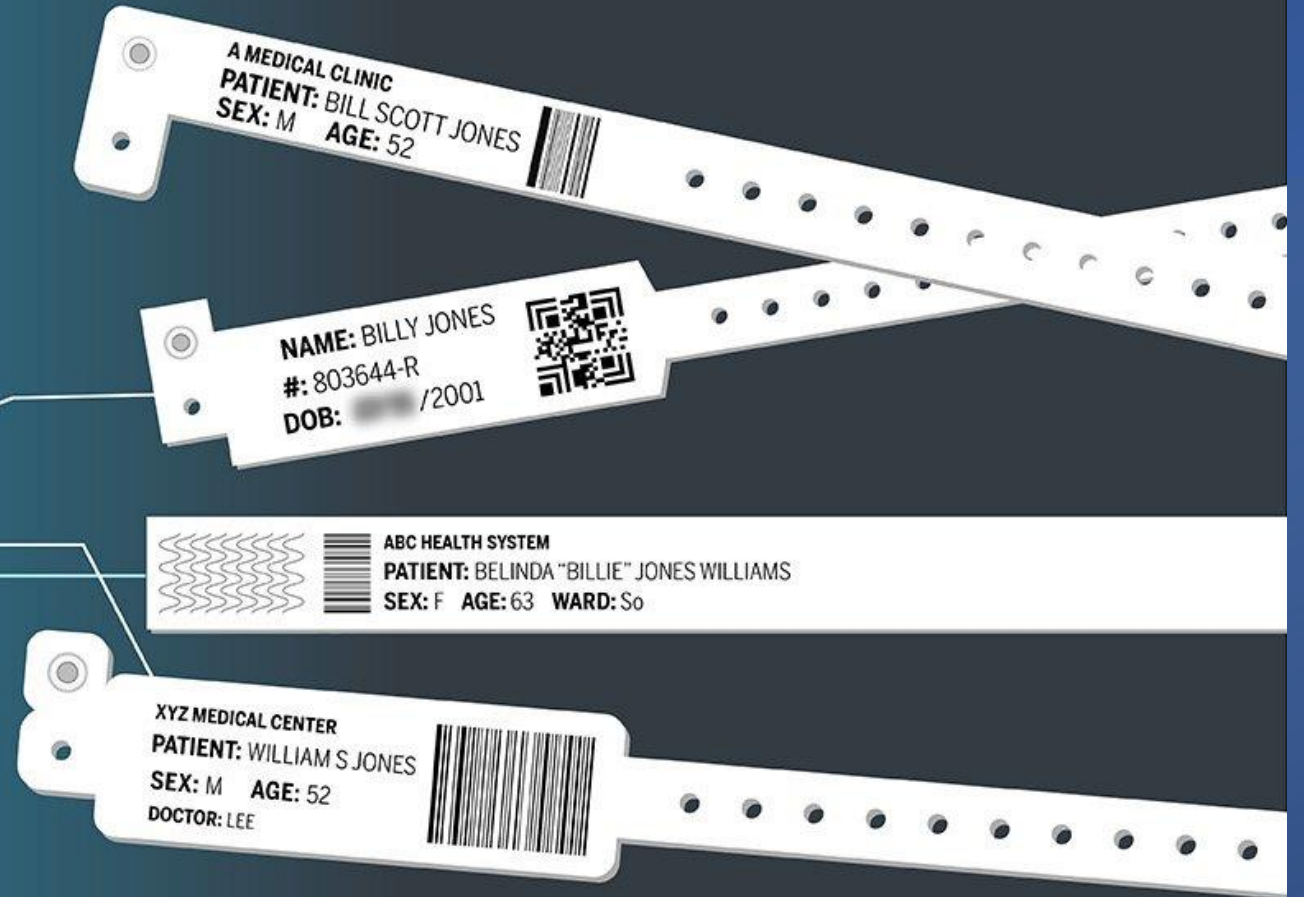
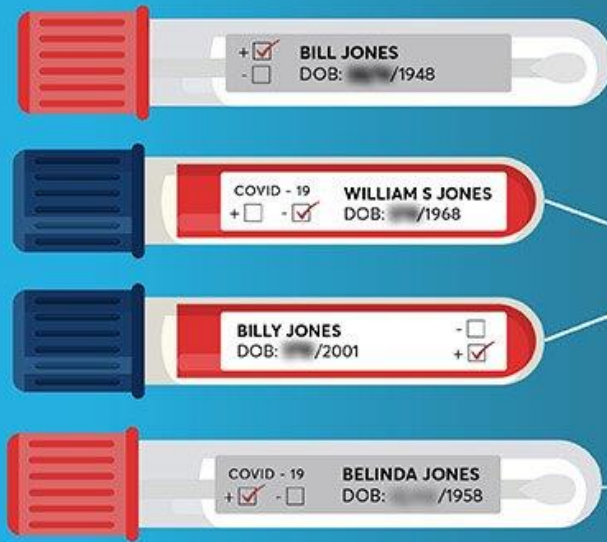


Data Cleanup

- Data Entry Errors
- Diversity and Evolving Standards
- Data Staging Errors
- Entity Resolutions



The Importance of Matching Patient Records

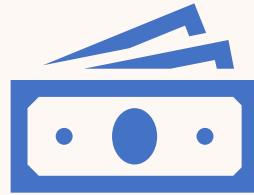


Matching patient records from disparate sources is not only achievable, but fundamental to stem the tide of the current pandemic and allow for fast action for future highly contagious viruses.

Patient Records Matching



Patient Harm



Financial Cost



Solutions with Patient
Security Risk



Data Storage



Data Storage

- Relational Databases
- NoSQL Databases
- Graph Databases
- Cloud Storage
- Security Considerations





Data Analytics

Data Analytics

- Measurements, Metrics and Indicators
- Population Health Analytics





Predictive Analytics



Predictive Analytics



Decision Trees



Regression
Analysis



Neural
Networks



Questions?

