

Component 4/Unit 6d
Topic IV: Design a simple relational database using data modeling and normalization

- Description and Information Gathering
- Data Model
- Normalization, Functional Dependencies and Constraints
- Final design, relationships, Primary keys and Foreign keys

Description of Database

- The database we are going to design is to keep track of new medications that are in trial testing. We need to keep track of the medications, the trials for those medications and the clinical institutions that are doing the testing.

Information Gathering

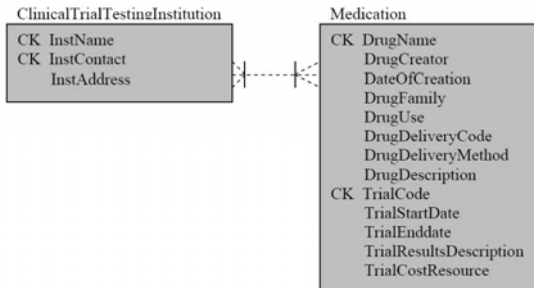
- Through meetings with users and looking at forms and reports it was determined that certain data about a clinical institution needed to be kept in the data base.
 - Name of the institution
 - Contact information
 - Address

Information Gathering Continued

Information about the medication needed from the database.

Drug name	Trial start date
Drug creator	Trial end date
Date of Creation	Trial results description
Drug family	Trial cost resource
Drug use	
Drug description	
Trial code	

Data Model - First Attempt



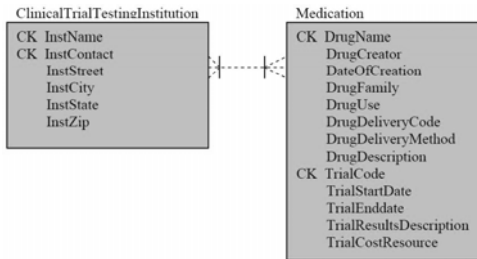
Next the database is normalized.

Normalization

- A database is normalized to eliminate data anomalies.
 - Deletion anomaly
 - Insertion anomaly
- Functional dependencies
- Constraints
 - Data rules that must be followed
- Referential Integrity Constraint

Functional Dependencies

- Data within a row can be shown to be dependent on a candidate key



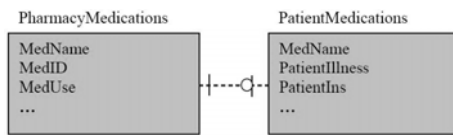
Component 4/Unit 6d

Health IT Workforce Curriculum
Version 1.0/Fall 2010

7

Referential Integrity Constraint

- An attribute of one entity is a subset of an attribute of another entity.



Component 4/Unit 6d

Health IT Workforce Curriculum
Version 1.0/Fall 2010

8

First Normal Form (1NF)

- Definition of a relation
 - Data rows within an entity must be unique and connected to the entity (can't have data in a relation that is associated with something else).
 - Columns are uniquely named and contain only one piece of data (attribute)
 - The sequence of rows and columns is not important.

Component 4/Unit 6d

Health IT Workforce Curriculum
Version 1.0/Fall 2010

9

Putting the Example database in First Normal Form

- There is nothing to indicate that rows in the entities are not unique and the attributes are connected to each entity
- Columns are uniquely named, but Address in the entity ClinicalTrialInstitution contains more than one piece of data (Street, City, State and Zip)

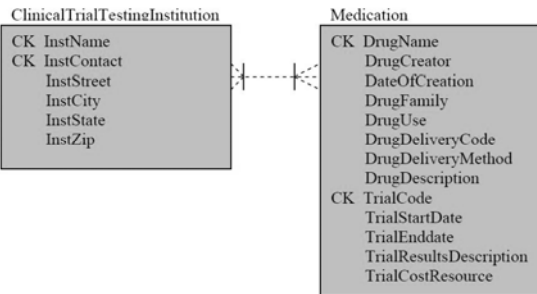


Component 4/Unit 6d

Health IT Workforce Curriculum
Version 1.0/Fall 2010

10

Putting a Database In 1NF



Component 4/Unit 6d

Health IT Workforce Curriculum
Version 1.0/Fall 2010

11

Second Normal Form (2NF)

- The second normal form eliminates deletion and insertion anomalies that are due to having an attribute or attributes dependent on something other than the key.
- This is especially true for a composite keys. To be in second normal form attributes must be dependent on the whole key.

Component 4/Unit 6d

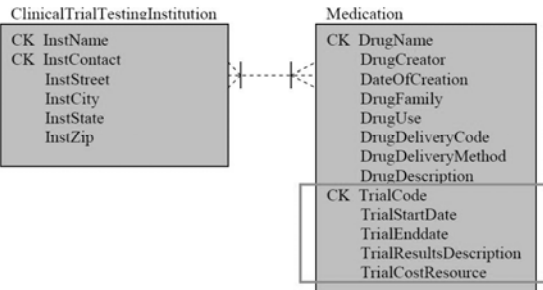
Health IT Workforce Curriculum
Version 1.0/Fall 2010

12

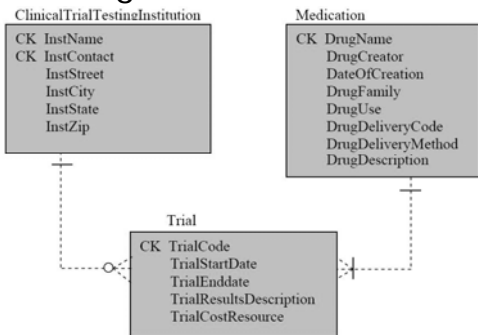
Second Normal Form Continued

- A relation is in second normal form if all its non-key attributes are dependent on the entire key.
- A relation in second normal form must also be in first normal form.

Putting The Trial DB In 2NF



Putting The Trial DB In 2NF



Third Normal Form (3NF)

- The third normal form eliminates deletion and insertion anomalies that are due to having an indirect dependency where an attribute is indirectly dependent on the key
- The attribute is directly dependent on an attribute that is dependent on the key
- This indirect dependency on the key is called a **transitive dependency**

Component 4/Unit 6d

Health IT Workforce Curriculum
Version 1.0/Fall 2010

16

Third Normal Form Continued

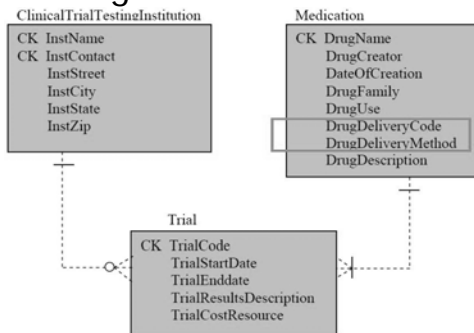
- A database is said to be in third normal form if there are no transitive dependencies
- A database in third normal form must also be in second and first normal forms
- Many Database Administrators (DBA) consider third normal form to be sufficient for most business and health care databases. Putting the database in a higher level of normalization may actually make the database less efficient.

Component 4/Unit 6d

Health IT Workforce Curriculum
Version 1.0/Fall 2010

17

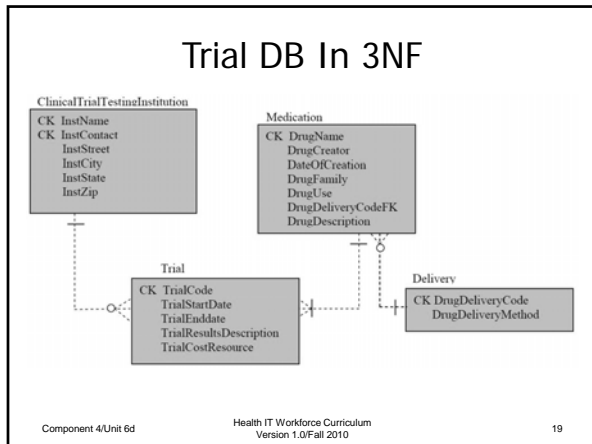
Putting The Trial DB In 3NF



Component 4/Unit 6d

Health IT Workforce Curriculum
Version 1.0/Fall 2010

18



Other Normal Forms

- DBAs sometimes have to troubleshoot problems and on occasion will use normal forms beyond third normal form.
- A database can be de-normalized to solve some slow response problems.
- Boyce-Codd Normal Form
 - A determinant is an attribute that determines another attribute
 - DB is in Boyce-Codd form if every determinant is a candidate key

Component 4/Unit 6d Health IT Workforce Curriculum Version 1.0/Fall 2010 20

Other Normal Forms

- Fourth Normal Form
 - Fixes an update anomaly that involves a multi-value dependency
 - A multi-value dependency exists when there are a minimum of three attributes, two of the attributes are multi-valued and the values of the two multi-value attributes depend only on a 3rd attribute.
 - This situation is rare
 - DB is in 4th normal form when there are no multi-value dependencies

Component 4/Unit 6d Health IT Workforce Curriculum Version 1.0/Fall 2010 21

Other Normal Forms

- Fifth Normal Form or Project-Join Normal Form (PJNF)
 - Generalization of multi-valued dependencies
 - Difficult to deal with
 - Extremely rare
- Domain Key Normal Form (DKNF)
 - Generalization of other non-time constraints
 - Difficult to deal with and rarely used

Evolution of the Data Model

- Data model progresses from being volatile with many changes to a DB design with little change or surprises
- In the final design entities will become tables, relationships will show minimum and maximum cardinality and primary/foreign keys are chosen

Final Design of the Trial DB

