COLUMBIA UNIVERSITY | MAILMAN SCHOOL of PUBLIC HEALTH
BIOSTATISTICS

In-Class Assignment 13

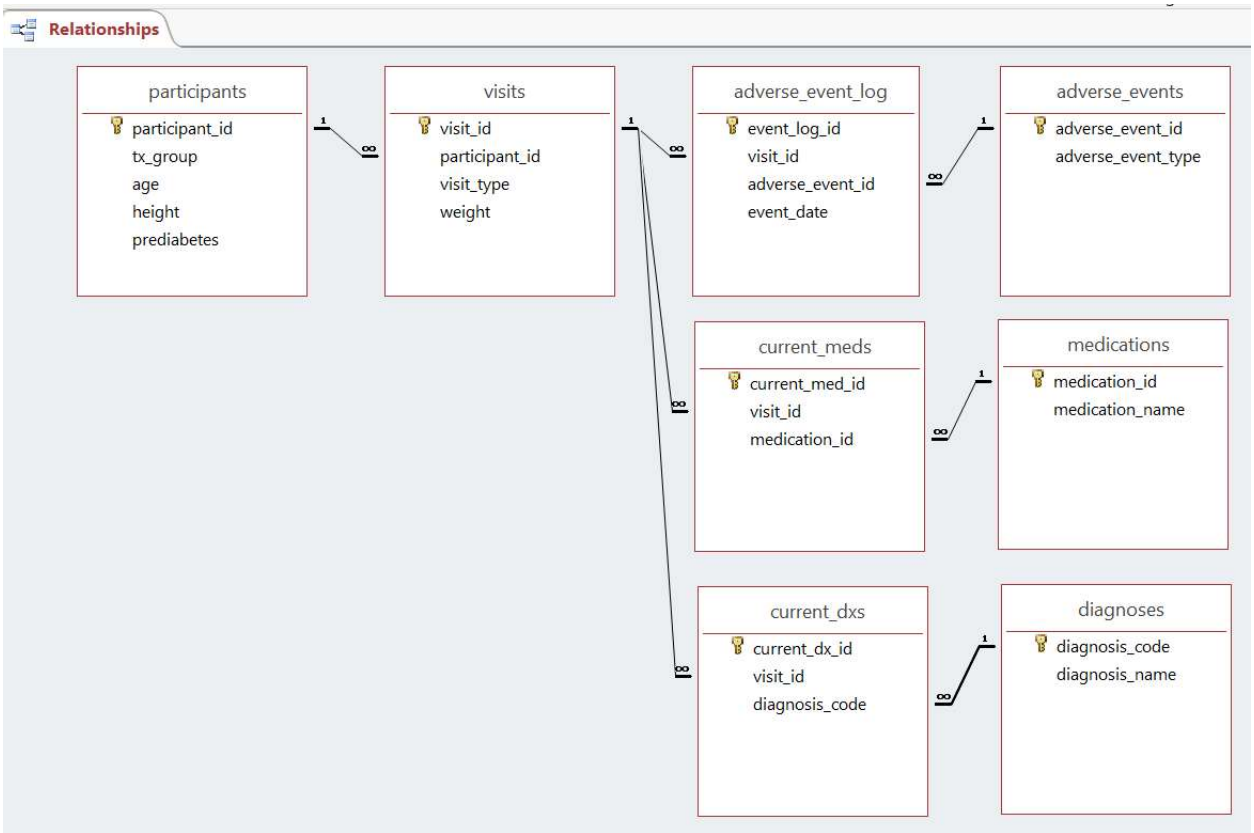
Creating Schemas and Tables in MySQL

Read the scenario below, then follow the instructions and answer the questions that follow. Add your SQL solutions in the relevant sections of the **In-Class Assignment 13.sql** SQL script, along with a **pdf printout (or screenshot if pdf not possible) of your EER diagram**, and submit to Canvas by the deadline listed above. Save your file frequently to avoid losing work!

Scenario:

You were previously asked by a research group to help build a database for the data they are collecting for a randomized controlled trial (RCT) examining the efficacy of a new weight-loss drug called WL001A. You will now build the same database in MySQL and populate the tables with data. Relevant information about the study is included below to help guide your design of the database.

Create a schema named **weight** and add the appropriate tables based on the relationship diagram from the previously built Access database, shown below. Again, use **Autonumber primary keys**.



NOTE: you will not be able to add combo boxes in MySQL, but prepare any fields that you would normally make a combo box with the appropriate data type. Also, do not worry about validation rules at this stage of the database development.

In-Class Assignment 13

Guidelines:

- Follow appropriate SQL style conventions when creating tables.
- Create parent tables before child tables so that you can apply foreign keys when creating child tables.
- Choose appropriate data types for each column and consider best practices for minimizing storage requirements
- For numeric fields, be sure to limit to an appropriate number of digits, and choose UNSIGNED when applicable
- Assume that text fields do not require more than 255 characters
- Add indexes where specified below, in addition to where needed to enforce uniqueness
- Allow Cascade Update but not Cascade Delete.
- Once all tables and relationships have been created, **generate an EER diagram** and organize it so that the tables appear in a similar arrangement to the relationships window of the same tables in Access

Study Background:

- This study began on 6/15/2017. Participants are randomized either to WL001A or placebo and followed for 20 weeks over 6 visits (Baseline, Week 4, Week 8, Week 12, Week 16, Week 20). Their treatment arm will not change over the course of the study. Treatment arm and visit week should be indexed.
- The following measurements are only taken once during the course of the study, at baseline:
    - Age (years) → Inclusion criteria: 18-55 years (integers only)
    - Height (meters), collected up to 2 decimal places (e.g., 66 inches = 1.68 meters)
    - Prediabetes status (Yes/No)
- At each follow-up visit, the following information will be collected:
    - Participant's current weight (kg), collected up to 2 decimal places (e.g., 250 lbs. = 113.40 kg). This should be indexed.

- Log of recent adverse events the participant experienced (indexed) since the last visit, and the date(s) (indexed) on which each event occurred (collect all that apply).
  - Excessive fatigue
  - Reaction to medication
  - Abnormal lab results
  - Hospitalization
  - Other
- Log of medications (indexed) that the participant is taking at the current visit
- Log of diagnoses (indexed) that have been confirmed for the participant at the current visit