CLINICAL TRIAL DATA MANAGEMENT SYSTEM -

A COMPREHENSIVE DATABASE FOR CLINICAL RESEARCH AND TRIAL MANAGEMENT

Prepared by:

Sheethal C Jayaram

Data Analyst

Date of Completion:

May 05, 2024

Project Direction Overview

Overview:

This database is designed to serve the needs of researchers, clinical trial coordinators, and healthcare professionals involved in conducting and managing clinical trials. It contains comprehensive data related to clinical trials, participants, adverse events, principal investigators, and statistical analyses.

Target Audience:

The primary users of this database include researchers, clinical trial coordinators, principal investigators, healthcare professionals, and regulatory authorities involved in clinical research and trial management.

Data Content:

The database contains structured information on various aspects of clinical trials, including trial design details such as trial name, objectives, phase, start and end dates, principal investigator details, and sponsor information. It also includes participant information such as demographics, medical history, enrollment status, adverse events reported during the trial, and statistical analyses conducted.

Usage:

The database will be used for tracking and managing clinical trial data, facilitating participant enrollment, monitoring trial progress, analyzing trial outcomes, and generating reports for regulatory submissions and publications. It will support decision-making processes related to trial design, participant recruitment, safety monitoring, and data analysis.

Importance:

This database is crucial for advancing medical research, improving patient care, and ensuring the safety and efficacy of new medical interventions. By providing a centralized repository of comprehensive and accurate clinical trial data, it enables efficient data management, enhances transparency and accountability, and supports evidence-based decision-making in healthcare.

Interest:

Our interest in this database stems from the recognition of the critical role that clinical trials play in advancing medical knowledge and improving patient outcomes. By developing and maintaining a robust database infrastructure for clinical trial management, we aim to contribute to the advancement of medical science and the improvement of public health. Additionally, the database presents an opportunity to leverage data analytics and machine learning techniques to derive valuable insights and inform future research directions. This is something which interests us as analysts.

Use Cases and Fields

- Setting up a new clinical trial
- Enrolling a new participant in the database
- Enrolling a participant in a specific trial
- Adding a new principal investigator to the database
- Reporting an adverse event during a trial
- Conducting statistical analysis on trial data

1. **Trial Design Table**

Field	**What it stores**	**Why it's needed**
TrialID (PK)	Unique identifier for each clinical trial	This is the Primary key of this table. We need this to keep track of the trials
TrialName	Name of the clinical trial	We need it to name the trials as it keeps track of them
Objective	Main objective or purpose of the trial	It is useful to get to know why the trials are being conducted and thus will help us understand the need.
Phase	Clinical trial phase (I, II, III, or IV)	It is necessary to know what phase the trials are at.
StartDate	Start date of the trial.	It is necessary to know the start dates of the trials.
EndDate	Estimated end date of the trial	It is necessary to know the end dates of the trials.
PrincipalInvestigatorID (FK)	Lead researcher or physician	It is necessary to know who is conducting the trials and to link it to the Principal

		Investigator Table.
Sponsor	Organization or entity funding the trial	It is useful to know who is sponsoring the trials and also link to a different database that keeps track of the sponsors.

2. **Participant Table:**

Field	What it stores	Why it's needed
ParticipantID:	Unique identifier for each participant. (PK)	This field is required as a unique identifier for each participant, facilitating data tracking and ensuring accurate record keeping.
ParticipantFirstName	First Name of participant	This field is necessary to identify participants individually and distinguish between them in the database.
ParticipantLastName	Last Name of participant	This field is necessary to identify participants individually and distinguish between them in the database.
Dob	Date of birth of the participant	This is essential for determining age-related eligibility criteria for participation in clinical trials and understanding age-related factors influencing health outcomes.
Gender	Gender identity of the participant.	Gender identity is important for analyzing gender-specific differences in response to treatments and assessing gender-related health disparities.
MedicalConditions	Any significant medical conditions or comorbidities	This helps researchers understand the impact of

	that may impact the participant's health status or treatment response.	pre-existing health conditions on trial outcomes and tailor interventions accordingly.
MedicationHistory	Current and past medications taken by the participant, including prescription drugs, over-the-counter medications, and supplements.	This field provides critical information about the participant's current and past medications, aiding in the assessment of potential drug interactions and treatment safety.
BMI (Body Mass Index):	A measure of body fat based on height and weight.	This is a useful measure for assessing participants' overall health status and evaluating the impact of weight-related factors on treatment efficacy.
Allergies:	Any known allergies or adverse reactions to medications, foods, or environmental factors.	This is crucial for ensuring participant safety during the trial and avoiding adverse reactions to medications or interventions.
VitalSigns:	Key physiological measurements such as blood pressure, heart rate, and respiratory rate.	This allows for ongoing monitoring of participants' health status throughout the trial and helps detect any adverse reactions or changes in health condition promptly.

3. **ParticipantEnrollment Table:**

Field	**What it stores**	**Why it's needed**
EnrollmentID (PK)	Unique identifier for each enrollment	It is necessary to track the unique enrollments to help ensure smooth tracking.

ParticipantID (FK)	Associated participant identifier	It is important to know the participants who participated
TrialID (FK)	Associated trial identifier	It is important to know the trials and link them with the trial table.
ConsentDate	Date when consent was obtained	It is necessary to know the consent date obtained from the participants. Consent is the first step of the clinical trials.
EligibilityStatus	Indicates if the participant meets the trial's eligibility criteria	It is useful to know if the participant is eligible for the trials. If they are not, they should not be participating in the trials.
EnrollmentDate	Date when the participant was enrolled in the trial	It is useful to know when the participant was enrolled in the trial.
WithdrawalDate	Date, if any, when the participant withdrew from the trial	It is useful to know if a participant withdrew from a trail and how many people withdrew.
ReasonForWithdrawal	Reason for withdrawal, if applicable	It is useful to know why they withdrew from the trial for future references.

4. **PrincipalInvestigator Table**

Field	**What it stores**	**Why it's needed**
PrincipalInvestigatorID:	Unique identifier for each principal investigator. PK	This field is necessary as a unique identifier for each principal investigator, enabling efficient data management and referencing in the database.
FirstName	First name of the principal investigator.	This field is required to accurately identify and

		distinguish between principal investigators, ensuring clear communication and record-keeping.
LastName:	Last name of the principal investigator.	This field is required to accurately identify and distinguish between principal investigators, ensuring clear communication and record-keeping.
Email:	Email address of the principal investigator.	This is essential for communication purposes, allowing for efficient correspondence regarding trial management and updates.
Phone:	Phone number of the principal investigator.	This allows for direct contact with the principal investigator, facilitating timely communication and coordination of trial-related activities.
Institution:	Institution or organization affiliation of the principal investigator.	This field indicates the affiliation of the principal investigator with a specific institution or organization, providing context for their professional standing and institutional support.
Department:	Department within the institution where the principal investigator works.	This helps understand their institutional role and potential areas of collaboration.
Specialty:	Specialty or area of expertise of the principal investigator.	This aids in selecting investigators with relevant skills and knowledge for specific types of clinical trials.
ExperienceYears:	Number of years of experience as a principal investigator.	This is important for assessing the investigator's expertise, reliability, and track

		record in conducting clinical research.

5. **Conditions ID Table**

Field	**What it stores**	**Why it's needed**
ConditionID(PK)	Unique identifier for each condition	Primary key to ensure a unique identifier for each entry.
ParticipantID(FK)	Foreign key referencing ParticipantID from the Participant table	Establishes a link to the Participant table for relational integrity.
ConditionName	Name or description of the medical condition	Stores the name or description of the medical condition
Allergies	Any known allergies or adverse reactions related to medical condition	Stores information about allergies or adverse reactions associated with the medical condition.
MedicationHistory	Details of current and past medications associated with the medical condition	Records information about the medications related to the medical condition.

6. **Adverse Events Table**

Field	**What it stores**	**Why it's needed**
EventID (PK)	Unique identifier for each adverse event reported	It is necessary to track the unique events.

ParticipantID (FK)	Associated participant identifier.	It is important to know the participants who participated in a trial.
TrialID (FK)	Associated trial identifier	It is important to know the trials and to link back to the trial table.
DateReported	Date when the event was reported	It is very important to know when the incident happened.
Description	Detailed description of the adverse event.	The description will help explain the situation clearly.
Severity	Severity rating of the adverse event	It helps to understand the severity of the situation and its impact. It also helps track which trials were more severe on the participants.
ActionTaken	Actions taken in response to the adverse event	It is necessary to track the actions taken by the management for clarity and for future reference.
Outcome	Outcome of the adverse event.	It is useful to keep track of the outcome as it will help in the future events

7. **Statistical Analysis Table**

Field	**What it stores**	**Why it's needed**
AnalysisID	Unique identifier for each analysis (PK)	This field serves as a unique identifier for each analysis, facilitating data organization and referencing in the database.
TrialID	Associated trial identifier. (FK)	This allows for easy retrieval and aggregation of analysis results based on specific trials.

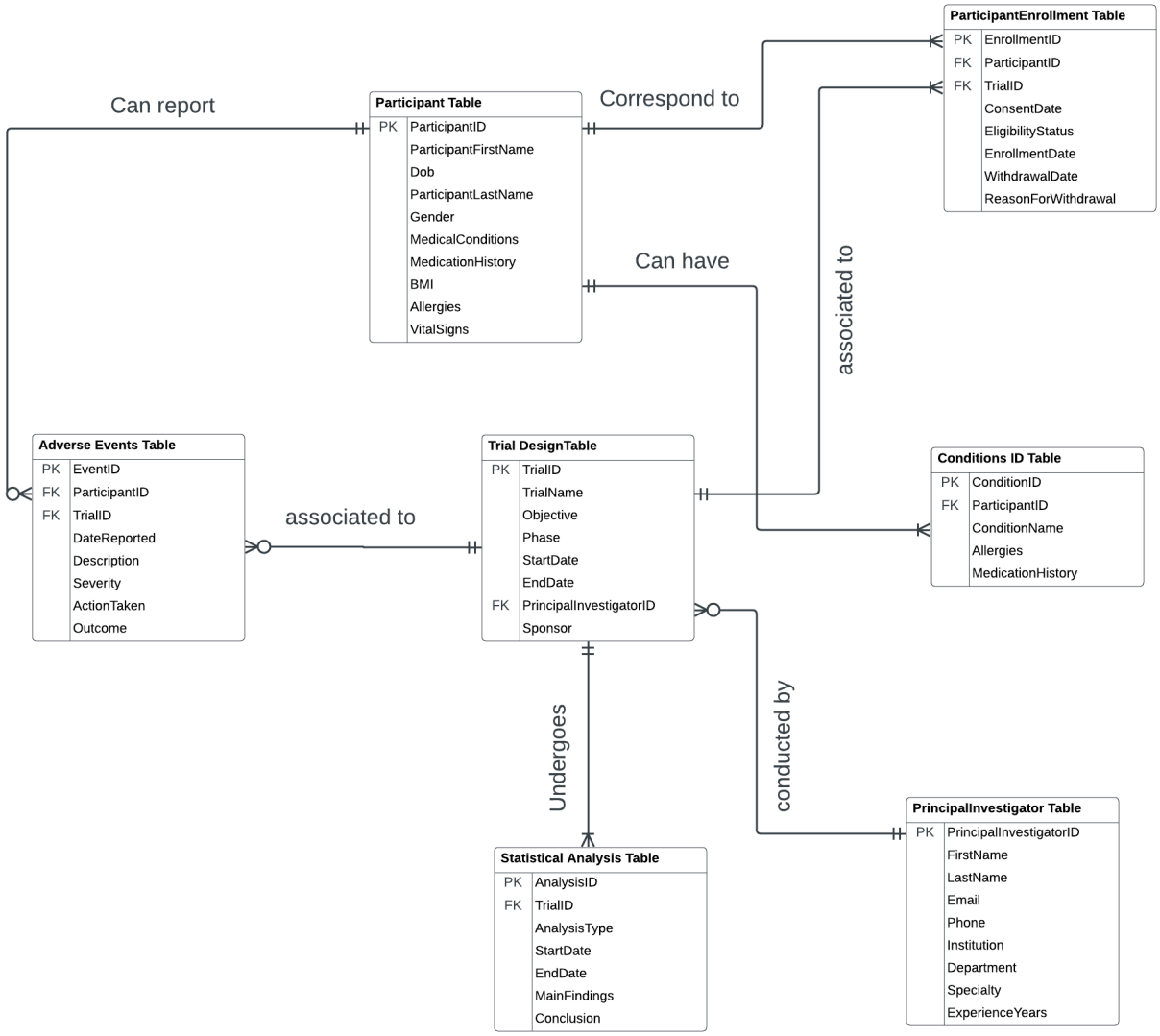
AnalysisType	Type of statistical analysis performed (e.g., intention-to-treat, per-protocol).	This field provides insight into the methodology used, aiding in result interpretation and replication.
StartDate	Start date of the analysis period.	This helps establish the timeframe for data collection and analysis, ensuring accuracy and consistency.
EndDate	End date of the analysis period.	This delineates the conclusion of data collection and analysis, providing clarity on the study timeline.
MainFindings	Summary of the main findings from the analysis.	This field allows for quick review and understanding of key results without needing to delve into detailed reports.
Conclusion	Conclusions drawn from the analysis.	This field provides actionable insights and recommendations based on the study findings, guiding future research directions and clinical decisions.

Structural Database Rules

1. One clinical trial can have multiple enrollments
2. One enrollment ID can correspond to only one participant.
3. One participant can correspond to multiple enrollment IDs (trials) over time.
4. Each participant can be enrolled in one clinical trial at a time.
5. Each participant can report multiple adverse events during a trial
6. Each clinical trial is conducted by one principal investigator
7. A principal investigator can oversee no trials or multiple trials.
8. Each participant can have multiple medical conditions, medications, and allergies recorded in their profile.
9. Each trial undergoes at least 1 statistical analysis.
10. Atleast one to multiple analyses can be conducted for a single trial.
11. Each adverse event is associated with one trial.
12. Each adverse event is associated with one participant.
13. A trial can have multiple adverse events reported.

14. A participant can experience none or multiple adverse events.
15. A trial can be an Interventional Trial or an Observational Trial, both or none.
16. A participant can be an adult or a minor.
17. An adverse event can be mild, moderate, or severe.

Entity Relationship Diagram



Specialization-Generalization Structural Database Rules:

1. Participant Specialization:

General Entity: Participant

Description: Represents all individuals enrolled in the database with general attributes such as ID, name, DOB, gender, and medical conditions.

Specialized Entities:

a) **Adult Participant**

Steps:

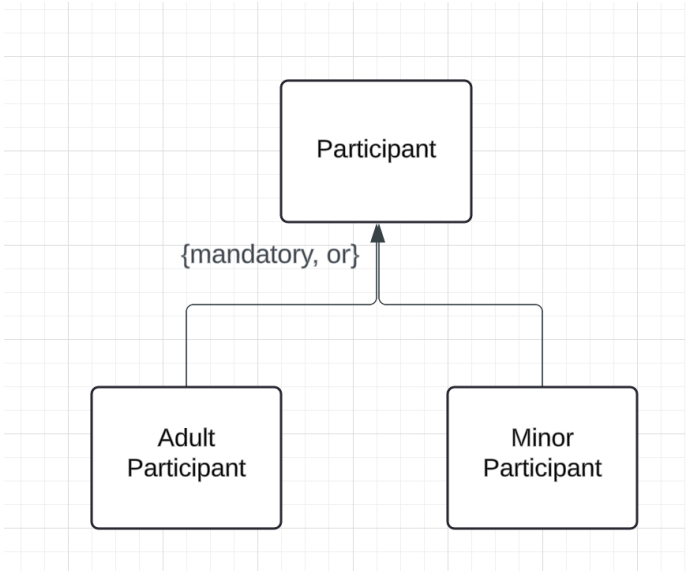
1. If the participant's age is 18 years or older, they are classified as an Adult Participant.
2. Adult Participants must provide their own consent to participate in trials.
3. Additional attributes for Adult Participants: ConsentStatus.

b) **Minor Participant**

Steps:

1. If the participant's age is under 18 years, they are classified as a Minor Participant.
2. Minor Participants require consent from a guardian or parent to participate in trials.
3. Additional attributes for Minor Participants: GuardianName, GuardianContactInfo.

Note: Participant classification into Adult or Minor is based on the Date of Birth (Dob) attribute.



2. Trial Type Specialization:

General Entity: Clinical Trial (Trial Design table)

Description: Represents all clinical trials with attributes such as ID, name, objective, phase, and dates.

Specialized Entities:

a) **Interventional Trial**

Steps:

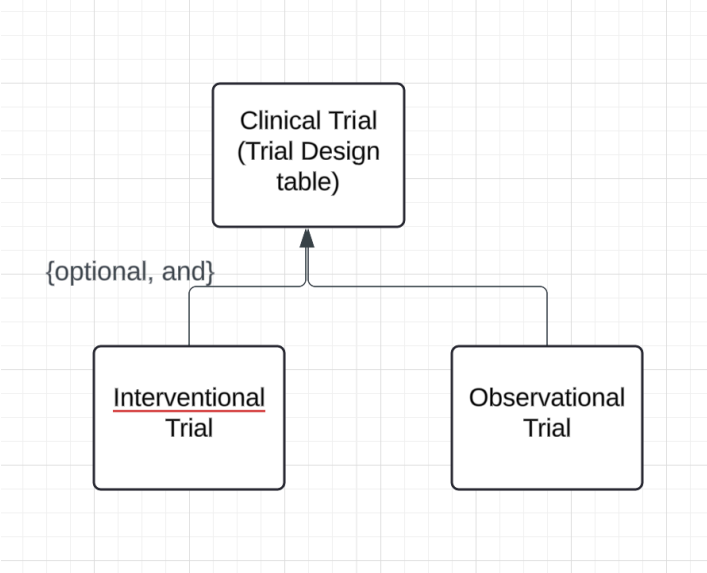
1. Trials focusing on the effects of interventions on participants are classified as Interventional Trials.
2. Additional attributes for Interventional Trials: InterventionType.

b) **Observational Trial**

Steps:

1. Trials observing participants without intervention are classified as Observational Trials.
2. Additional attributes for Observational Trials: ObservationMethod.

Note: Trial classification is based on the Objective attribute, specifying whether the trial is interventional or observational.



3. Adverse Event Severity Generalization:

Generalized Entity: Adverse Event

Description: Represents adverse events reported during trials, categorized by severity.

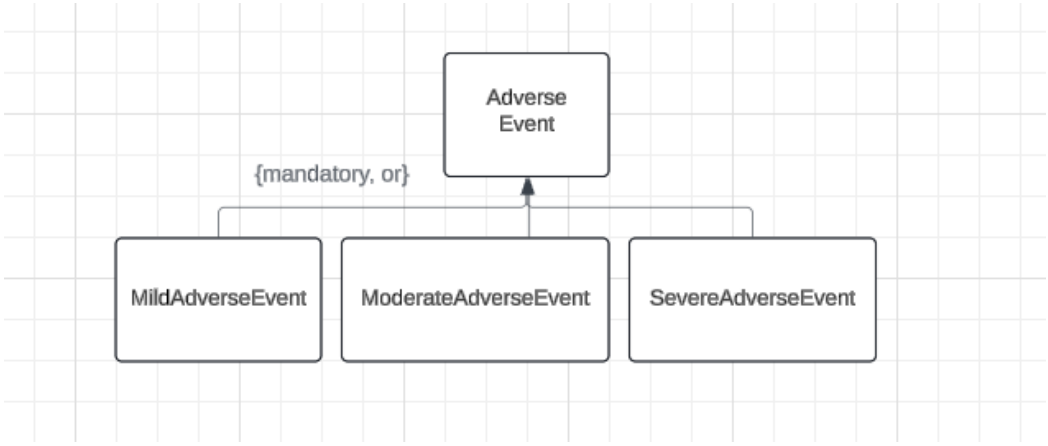
Steps:

1. All adverse events are recorded with a severity level.
2. Events are classified into Mild, Moderate, or Severe based on the SeverityLevel attribute.

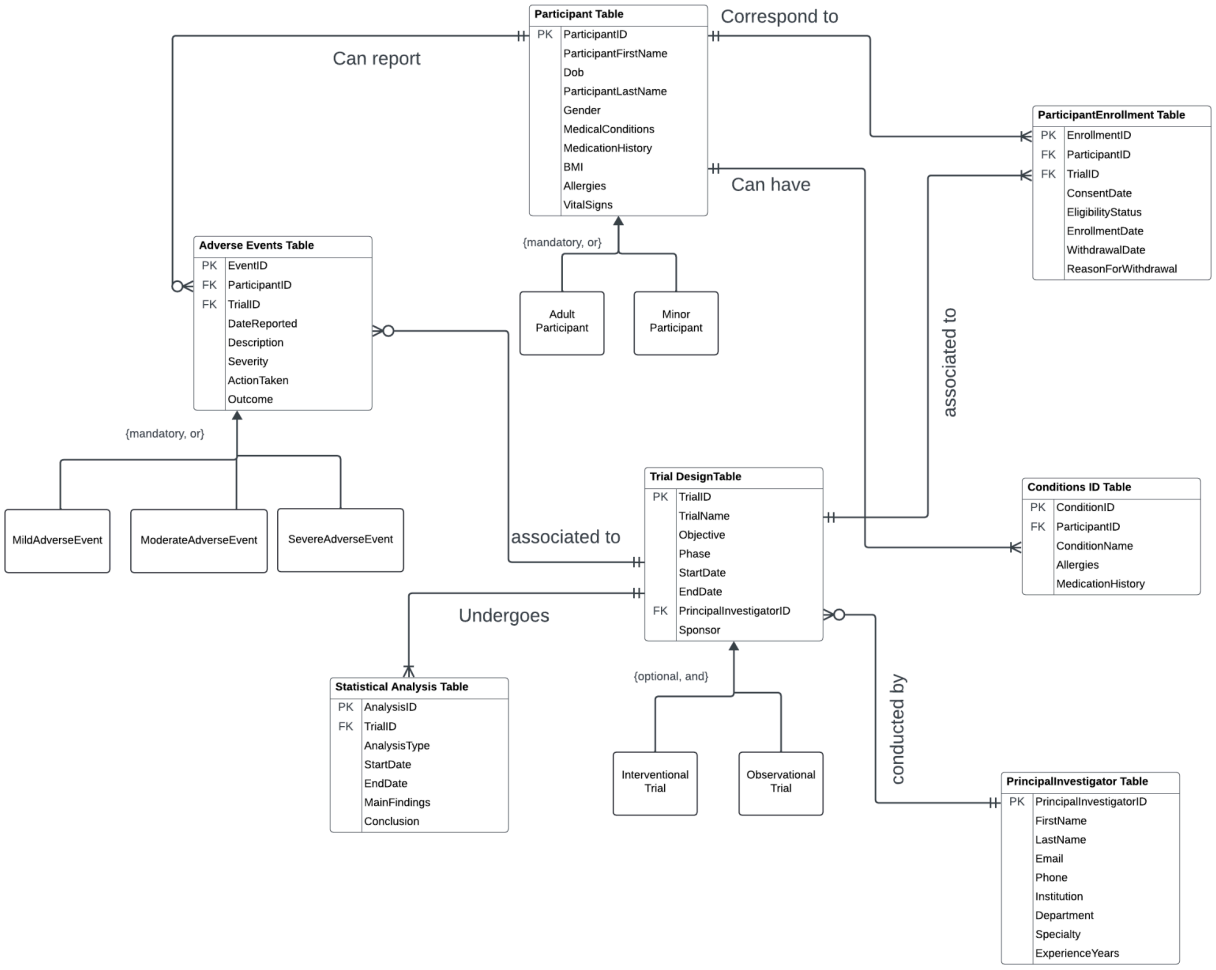
3. A single Adverse Event entity can encompass all severity levels, with specific handling based on the severity.

Note: This generalization allows for a unified approach to managing adverse events while still recognizing their different impacts.

Specific Entities: MildAdverseEvent, ModerateAdverseEvent, SevereAdverseEvent



Initial DBMS Physical ERD:



Adding Data Types to Attributes:

Table Name	Attribute	Data Type	Reasoning
Trial Design Table	TrialID (PK)	VARCHAR(10)	Unique identifier for trials with alphanumeric characters, maximum length 10.
	TrialName	VARCHAR(255)	Descriptive names for trials with a maximum length of 255 characters.

	Objective	TEXT	Longer descriptions of trial objectives.
	Phase	VARCHAR(20)	Represents phase of the trial.
	StartDate	DATE	Stores start date of the trial.
	EndDate	DATE	Stores estimated end date of the trial.
	PrincipalInvestigator ID (FK)	INT	Numerical identifier linking to Principal Investigator Table.
	Sponsor	VARCHAR(255)	Descriptive names for sponsors with a maximum length of 255 characters.
Participant Table	ParticipantID (PK)	VARCHAR(10)	Unique identifier for participants with alphanumeric characters, maximum length 10.
	ParticipantFirstNam e	VARCHAR(50)	First name of participants with a maximum length of 50 characters.
	ParticipantLastNam e	VARCHAR(50)	Last name of participants with a maximum length of 50 characters.
	Dob	DATE	Stores date of birth of participants.
	Gender	VARCHAR(5)	Represents gender identity.
	MedicalConditions	TEXT	Longer descriptions of medical conditions.

	MedicationHistory	TEXT	Longer descriptions of medication history.
	Allergies	TEXT	Longer descriptions of allergies.
	BMI	DECIMAL(5,2)	Stores Body Mass Index with precision of 5 digits and 2 decimal places.
ParticipantEnrollment	EnrollmentID (PK)	VARCHAR(10)	Unique identifier for enrollments with alphanumeric characters, maximum length 10.
	ParticipantID (FK)	VARCHAR(10)	Associated participant identifier with alphanumeric characters, maximum length 10.
	TrialID (FK)	VARCHAR(10)	Associated trial identifier with alphanumeric characters, maximum length 10.
	ConsentDate	DATE	Stores date when consent was obtained.
	EligibilityStatus	VARCHAR(20)	Represents participant eligibility status.
	EnrollmentDate	DATE	Stores date when participant was enrolled in the trial.
	WithdrawalDate	DATE	Stores date when participant withdrew from the trial.
	ReasonForWithdrawal	TEXT	Longer descriptions of reasons for withdrawal.

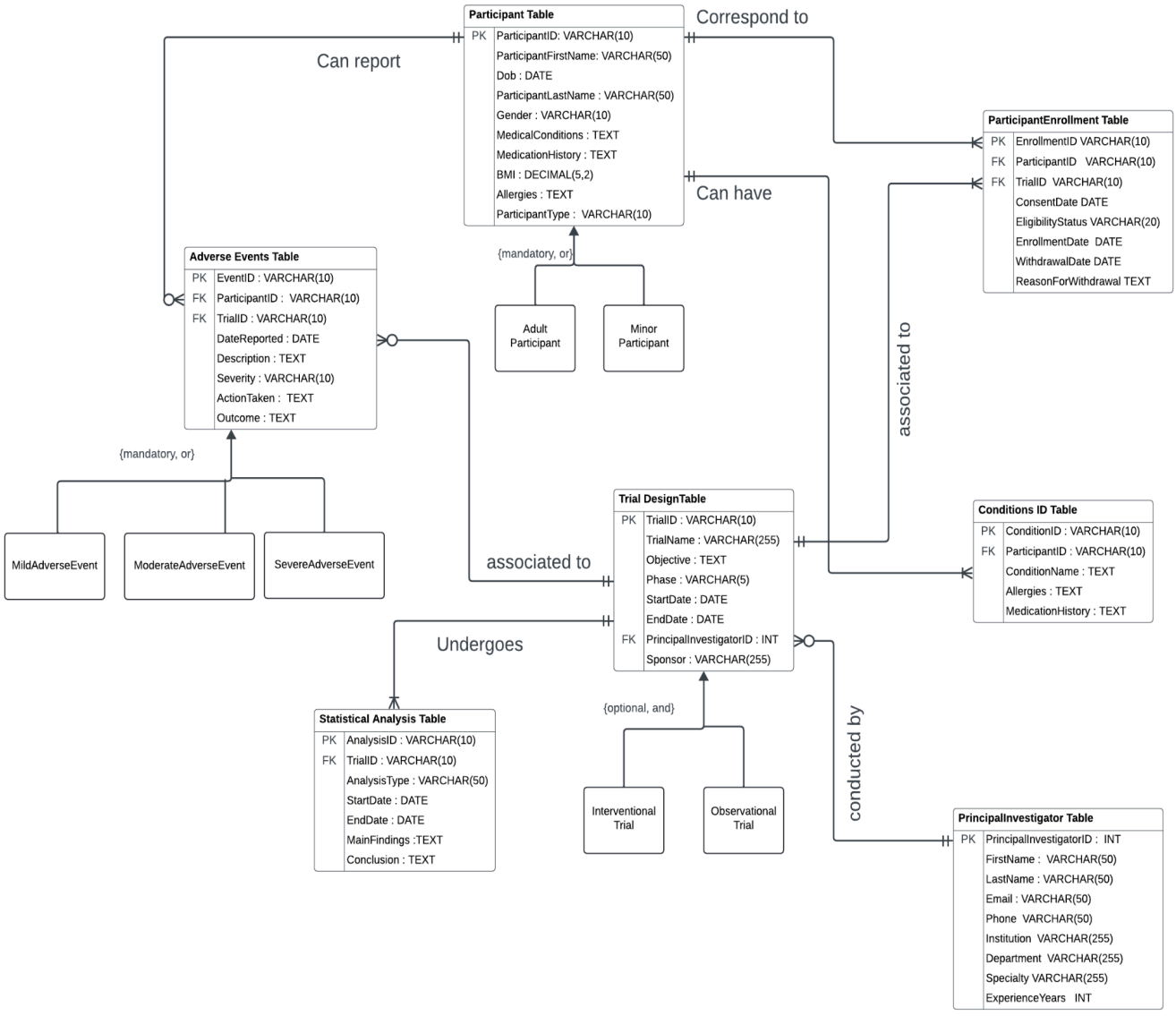
PrincipalInvestigator	PrincipalInvestigatorID (PK)	INT	Unique identifier for principal investigators as numerical value.
	FirstName	VARCHAR(50)	First name of principal investigators with a maximum length of 50 characters.
	LastName	VARCHAR(50)	Last name of principal investigators with a maximum length of 50 characters.
	Email	VARCHAR(50)	Email address of principal investigators with a maximum length of 50 characters.
	Phone	VARCHAR(20)	Phone number of principal investigators with a maximum length of 20 characters.
	Institution	VARCHAR(255)	Affiliation institution of principal investigators with a maximum length of 255 characters.
	Department	VARCHAR(255)	Department within institution of principal investigators with a maximum length of 255 characters.
	Specialty	VARCHAR(255)	Specialty or expertise of principal investigators with a maximum length of 255 characters.
	ExperienceYears	INT	Number of years of experience as principal investigator.

Conditions ID Table	ConditionID (PK)	VARCHAR(10)	Unique identifier for conditions with alphanumeric characters, maximum length 10.
	ParticipantID (FK)	VARCHAR(10)	Associated participant identifier with alphanumeric characters, maximum length 10.
	ConditionName	TEXT	Longer descriptions of medical conditions.
	Allergies	TEXT	Longer descriptions of allergies.
	MedicationHistory	TEXT	Longer descriptions of medication history.
Adverse Events Table	EventID (PK)	VARCHAR(10)	Unique identifier for adverse events with alphanumeric characters, maximum length 10.
	ParticipantID (FK)	VARCHAR(10)	Associated participant identifier with alphanumeric characters, maximum length 10.
	TrialID (FK)	VARCHAR(10)	Associated trial identifier with alphanumeric characters, maximum length 10.
	DateReported	DATE	Stores date when the event was reported.
	Description	TEXT	Longer descriptions of adverse events.

	Severity	VARCHAR(10)	Represents severity level of adverse event.
	ActionTaken	TEXT	Longer descriptions of actions taken in response to adverse events.
	Outcome	TEXT	Longer descriptions of outcomes of adverse events.
Statistical Analysis	AnalysisID (PK)	VARCHAR(10)	Unique identifier for analysis with alphanumeric characters, maximum length 10.
	TrialID (FK)	VARCHAR(10)	Associated trial identifier with alphanumeric characters, maximum length 10.
	AnalysisType	VARCHAR(50)	Type of statistical analysis performed with a maximum length of 50 characters.
	StartDate	DATE	Stores start date of the analysis period.
	EndDate	DATE	Stores end date of the analysis period.
	MainFindings	TEXT	Longer descriptions of main findings from the analysis.
	Conclusion	TEXT	Longer descriptions of conclusions drawn from the analysis.

This table provides a structured overview of the recommended data types for each attribute in the database tables. Adjustments can be made based on specific database requirements and constraints.

Full DBMS Physical ERD:



Normalization

The tables mentioned above adhere to normalization principles. Take, for instance, the Participant Enrollment table, where each entry includes an EnrollmentID (which is the primary key) and a few foreign keys, along with other attributes. Take, for instance, the foreign key -

ParticipantID. By utilizing this key to reference the Participant table, comprehensive details such as first name, last name, and date of birth can be retrieved. However, data redundancy is avoided within the Participant Enrollment table as it solely contains necessary linking information, maintaining normalization. Similarly, we maintain normalization across all tables in the database.

Normalization of a database involves organizing the tables and their relationships to increase efficiency and minimize redundancy. The process involves dividing large tables into smaller, well-structured tables and defining relationships between them. Normalization typically proceeds through several normal forms (1NF, 2NF, 3NF, BCNF, etc.).

After confirming that each table adheres to normalization rules, we can confidently say that our database is normalized to a certain normal form.

Minimum Entity Check

After normalization, We have more than 8 entities in our DBMS physical ERD to support the minimal complexity requirements for this project.

CONSTRAINTS APPLIED:

Participant Table Constraints:

ParticipantID is the primary key.
ParticipantType indicates whether the participant is an Adult or a Minor.

Trial Design Table Constraints:

TrialID is the primary key.
PrincipalInvestigatorID is a foreign key referencing PrincipalInvestigatorID in the PrincipalInvestigator table.

Participant Enrollment Table Constraints:

EnrollmentID is the primary key.
ParticipantID is a foreign key referencing ParticipantID in the Participant table.
TrialID is a foreign key referencing TrialID in the TrialDesign table.

Principal Investigator Table Constraints:

PrincipalInvestigatorID is the primary key.

Conditions ID Table Constraints:

ConditionID is the primary key.
ParticipantID is a foreign key referencing ParticipantID in the Participant table.

Adverse Events Table Constraints:

EventID is the primary key.
ParticipantID is a foreign key referencing ParticipantID in the Participant table.
TrialID is a foreign key referencing TrialID in the TrialDesign table.
Mild, Moderate, and Severe Adverse Event Tables Constraints:
EventID in each specialized adverse event table is the primary key, and it is a foreign key referencing EventID in the AdverseEvents table.

Statistical Analysis Table Constraints:

AnalysisID is the primary key.
TrialID is a foreign key referencing TrialID in the TrialDesign table.

Additional Constraints to be added:

Participant Table Constraints:

ParticipantID is NOT NULL and UNIQUE.
ParticipantFirstName and ParticipantLastName are NOT NULL.
Dob, Gender, and BMI can be NULL.
ParticipantType is NOT NULL.

Trial Design Table Constraints:

TrialID is NOT NULL and UNIQUE.
TrialName, Objective, Phase, StartDate, and EndDate are NOT NULL.
PrincipalInvestigatorID is NOT NULL.

Participant Enrollment Table Constraints:

EnrollmentID is NOT NULL and UNIQUE.
ParticipantID, TrialID, ConsentDate, EligibilityStatus, and EnrollmentDate are NOT NULL.
WithdrawalDate can be NULL.

Principal Investigator Table Constraints:

PrincipalInvestigatorID is NOT NULL and UNIQUE.
FirstName, LastName, Email, and Phone are NOT NULL.

Conditions ID Table Constraints:

ConditionID is NOT NULL and UNIQUE.
ParticipantID is NOT NULL.
ConditionName is NOT NULL.

Adverse Events Table Constraints:

EventID is NOT NULL and UNIQUE.
ParticipantID and TrialID are NOT NULL.
DateReported, Description, and Severity are NOT NULL.
Mild, Moderate, and Severe
EventID in each specialized adverse event table is NOT NULL and UNIQUE.

Statistical Analysis Table Constraints:

AnalysisID is NOT NULL and UNIQUE.
TrialID is NOT NULL.
AnalysisType, StartDate, and EndDate are NOT NULL.

Indexing Databases:

Identifying Columns Needing Indexes for the Database:

Below are the primary keys using indexes for the database and a list of indexed foreign keys as per the requirement of the Clinical trial Database.

Primary Keys Using Indexes:
1. ParticipantID in the Participant table
2. TrialID in the TrialDesign table
3. EnrollmentID in the ParticipantEnrollment table
4. PrincipalInvestigatorID in the PrincipalInvestigator table
5. ConditionID in the ConditionsID table
6. EventID in the AdverseEvents table
7. AnalysisID in the StatisticalAnalysis table

All the foreign keys requiring an Index are listed in the table below. The table provides details identifying each foreign key, column name, its uniqueness, and description.

Table Name	Column Name	Uniqueness	Description
ParticipantEnrollment	ParticipantID	No	Foreign key referencing ParticipantID in the Participant table. This foreign key allows linking participant enrollment information to the respective participant details. It's not unique because multiple enrollments can be associated with the same participant.
ParticipantEnrollment	TrialID	No	Foreign key referencing TrialID in the TrialDesign table. This foreign key establishes a relationship between participant enrollment and trial details. It's not unique because multiple enrollments can be associated with the same trial.
ConditionsID	ParticipantID	No	Foreign key referencing ParticipantID in the Participant table. This foreign key links condition information to the respective participant details. It's not unique because

			multiple conditions can be associated with the same participant.
AdverseEvents	ParticipantID	No	Foreign key referencing ParticipantID in the Participant table. This foreign key establishes a relationship between adverse events and the participant involved. It's not unique because multiple adverse events can be associated with the same participant.
AdverseEvents	TrialID	No	Foreign key referencing TrialID in the TrialDesign table. This foreign key links adverse events to the respective trial details. It's not unique because multiple adverse events can be associated with the same trial.
MildAdverseEvent	EventID	Yes	Foreign key referencing EventID in the AdverseEvents table. This foreign key establishes a

			relationship between specialized mild adverse events and their parent adverse events. EventID is unique within the MildAdverseEvent table as it serves as the primary key.
ModerateAdverseEvent	EventID	Yes	Foreign key referencing EventID in the AdverseEvents table. This foreign key establishes a relationship between specialized moderate adverse events and their parent adverse events. EventID is unique within the ModerateAdverseEvent table as it serves as the primary key.
SevereAdverseEvent	EventID	Yes	Foreign key referencing EventID in the AdverseEvents table. This foreign key establishes a relationship between specialized severe adverse events and their parent adverse events. EventID is unique within the SevereAdverseEvent table as it serves as

			the primary key.
StatisticalAnalysis	TrialID	No	Foreign key referencing TrialID in the TrialDesign table. This foreign key establishes a relationship between statistical analyses and the respective trial details. It's not unique because multiple analyses can be associated with the same trial.

In the provided clinical trials database schema, indexing is employed to expedite data retrieval operations and enhance overall query performance. By indexing key columns such as ParticipantID, TrialID, and other identifiers across various tables, the database engine can swiftly locate and retrieve specific records based on these criteria. For instance, indexing the ParticipantID column in the Participant table enables rapid access to participant details, facilitating efficient participant management and enrollment processes. Similarly, indexing TrialID in tables such as TrialDesign and ParticipantEnrollment accelerates queries related to trial details and enrollment status, streamlining trial management activities.

The use of indexing for this database not only improves query performance but also enhances system responsiveness and user experience. By minimizing the time required for data retrieval operations, indexing ensures that researchers, clinicians, and administrators can access critical information swiftly, leading to faster decision-making and more efficient workflows in clinical trial management. Moreover, indexing optimizes resource utilization by reducing the load on hardware resources such as CPU and memory, ensuring optimal performance even during peak usage periods. Overall, indexing is crucial for a clinical trial database and serves as a vital optimization technique, enabling seamless access to trial data and supporting informed decision-making processes.

Stored procedures

```sql
-- Stored procedure for adding data to Participant table
CREATE or alter PROCEDURE AddParticipant
    @ParticipantID VARCHAR(10),
    @ParticipantFirstName VARCHAR(50),
    @ParticipantLastName VARCHAR(50),
    @Dob DATE,
    @Gender VARCHAR(10),
    @MedicalConditions TEXT,
    @MedicationHistory TEXT,
    @Allergies TEXT,
    @BMI DECIMAL(5,2),
    @ParticipantType VARCHAR(10)
AS
BEGIN
    INSERT INTO Participant (ParticipantID, ParticipantFirstName, ParticipantLastName, Dob, Gender, MedicalConditions, MedicationHistory, Allergies, BMI, ParticipantTy
    VALUES (@ParticipantID, @ParticipantFirstName, @ParticipantLastName, @Dob, @Gender, @MedicalConditions, @MedicationHistory, @Allergies, @BMI, @ParticipantType);
END;
GO

-- Stored procedure for adding data to AdultParticipant table
CREATE or alter PROCEDURE AddAdultParticipant
    @ParticipantID VARCHAR(10),
    @ConsentStatus VARCHAR(20)
AS
BEGIN
    INSERT INTO AdultParticipant (ParticipantID, ConsentStatus)
    VALUES (@ParticipantID, @ConsentStatus);
END;
GO
```

```sql
-- Stored procedure for adding data to ParticipantEnrollment table
CREATE or alter PROCEDURE AddParticipantEnrollment
    @EnrollmentID VARCHAR(10),
    @ParticipantID VARCHAR(10),
    @TrialID VARCHAR(10),
    @ConsentDate DATE,
    @EligibilityStatus VARCHAR(20),
    @EnrollmentDate DATE,
    @WithdrawalDate DATE,
    @ReasonForWithdrawal TEXT
AS
BEGIN
    INSERT INTO ParticipantEnrollment (EnrollmentID, ParticipantID, TrialID, ConsentDate, EligibilityStatus, EnrollmentDate, WithdrawalDate, ReasonForWithdrawal)
    VALUES (@EnrollmentID, @ParticipantID, @TrialID, @ConsentDate, @EligibilityStatus, @EnrollmentDate, @WithdrawalDate, @ReasonForWithdrawal);
END;
GO

-- Stored procedure for adding data to PrincipalInvestigator table
CREATE or alter PROCEDURE AddPrincipalInvestigator
    @PrincipalInvestigatorID INT,
    @FirstName VARCHAR(50),
    @LastName VARCHAR(50),
    @Email VARCHAR(50),
    @Phone VARCHAR(20),
    @Institution VARCHAR(255),
    @Department VARCHAR(255),
    @Specialty VARCHAR(255),
    @ExperienceYears INT
AS
BEGIN
    INSERT INTO PrincipalInvestigator (PrincipalInvestigatorID, FirstName, LastName, Email, Phone, Institution, Department, Specialty, ExperienceYears)
    VALUES (@PrincipalInvestigatorID, @FirstName, @LastName, @Email, @Phone, @Institution, @Department, @Specialty, @ExperienceYears);
END;
GO
```

Tables with added data

```sql
-- Select statements for all tables

-- Participant table
SELECT * FROM Participant;

-- AdultParticipant table
SELECT * FROM AdultParticipant;

-- MinorParticipant table
SELECT * FROM MinorParticipant;

-- TrialDesign table
SELECT * FROM TrialDesign;
```

100 %

Results | Messages

	ParticipantID	ParticipantFirstName	ParticipantLastName	Dob	Gender	MedicalConditions	MedicationHistory	Allergies	BMI	ParticipantType
1	P001	John	Doe	1990-05-15	Male	None	None	None	24.50	Adult
2	P002	Jane	Smith	1995-08-20	Female	Asthma	None	Peanuts	22.30	Adult
3	P003	Michael	Johnson	2005-03-10	Male	ADHD	Ritalin	None	18.80	Minor
4	P004	Emily	Brown	2000-11-30	Female	None	None	Pollen	20.10	Minor
5	P005	David	Wilson	1985-07-25	Male	Diabetes	Insulin	Penicillin	28.90	Adult

	ParticipantID	ConsentStatus
1	P001	Consented
2	P002	Consented
3	P005	Consented

	ParticipantID	GuardianName	GuardianContactInfo
1	P003	Sarah Johnson	123-456-7890
2	P004	Jennifer Brown	987-654-3210

	TrialID	TrialName	Objective	Phase	StartDate	EndDate	PrincipalInvestigatorID	Sponsor
1	T001	COVID-19 Vaccine Trial	Evaluate the efficacy of a new COVID-19 vaccine.	Phase	2023-01-15	2024-06-30	101	ABC Pharmaceuticals
2	T002	Pain Management Study	Assess the effectiveness of a new pain relief me...	Phase	2024-03-10	2025-12-31	102	XYZ Biotech

```sql
-- TrialDesign table
SELECT * FROM TrialDesign;

-- InterventionalTrial table
SELECT * FROM InterventionalTrial;

-- ObservationalTrial table
SELECT * FROM ObservationalTrial;

-- ParticipantEnrollment table
SELECT * FROM ParticipantEnrollment;

-- PrincipalInvestigator table
SELECT * FROM PrincipalInvestigator;
```

100 %

Results | Messages

	TrialID	TrialName	Objective	Phase	StartDate	EndDate	PrincipalInvestigatorID	Sponsor
1	T001	COVID-19 Vaccine Trial	Evaluate the efficacy of a new COVID-19 vaccine.	Phase	2023-01-15	2024-06-30	101	ABC Pharmaceuticals
2	T002	Pain Management Study	Assess the effectiveness of a new pain relief me...	Phase	2024-03-10	2025-12-31	102	XYZ Biotech

	TrialID	InterventionType
1	T002	Medication

	TrialID	ObservationMethod
1	T001	Longitudinal Study

	EnrollmentID	ParticipantID	TrialID	ConsentDate	EligibilityStatus	EnrollmentDate	WithdrawalDate	ReasonForWithdrawal
1	E001	P001	T001	2023-01-20	Eligible	2023-01-25	NULL	NULL
2	E002	P002	T001	2023-01-22	Ineligible	2023-01-25	2023-02-10	High BMI
3	E003	P003	T002	2024-03-15	Eligible	2024-03-20	NULL	NULL

	PrincipalInvestigatorID	FirstName	LastName	Email	Phone	Institution	Department	Specialty	ExperienceYears
1	101	Dr. Robert	Smith	robert.smith@example.com	123-456-7890	University Hospital	Cardiology	Cardiologist	15
2	102	Dr. Lisa	Johnson	lisa.johnson@example.com	987-654-3210	City Clinic	Neurology	Neurologist	12

```sql
-- ConditionsID table
SELECT * FROM ConditionsID;

-- AdverseEvents table
SELECT * FROM AdverseEvents;

-- MildAdverseEvent table
SELECT * FROM MildAdverseEvent;

-- ModerateAdverseEvent table
SELECT * FROM ModerateAdverseEvent;

-- SevereAdverseEvent table
SELECT * FROM SevereAdverseEvent;
```

100 %

Results Messages

	PrincipalInvestigatorID	FirstName	LastName	Email	Phone	Institution	Department	Specialty	ExperienceYears
1	101	Dr. Robert	Smith	robert.smith@example.com	123-456-7890	University Hospital	Cardiology	Cardiologist	15
2	102	Dr. Lisa	Johnson	lisa.johnson@example.com	987-654-3210	City Clinic	Neurology	Neurologist	12

	ConditionID	ParticipantID	ConditionName	Allergies	MedicationHistory
1	C001	P002	Asthma	Pollen	None
2	C002	P005	Diabetes	None	Insulin

	EventID	ParticipantID	TrialID	DateReported	Description	Severity	ActionTaken	Outcome
1	AE001	P001	T001	2023-02-02	Fever after vaccination	Mild	Rest and hydration	Resolved on its own
2	AE002	P002	T001	2023-02-05	Allergic reaction to vaccine	Severe	Hospitalization and medication	Recovered after treatment
3	AE003	P003	T002	2024-04-02	Dizziness after medication	Moderate	Reduced dosage	Resolved with dosage adjustment

	EventID
1	AE001

	EventID
1	AE003

	EventID
1	AE002

```sql
-- StatisticalAnalysis table
SELECT * FROM StatisticalAnalysis;
```

100 %

Results Messages

	AnalysisID	TrialID	AnalysisType	StartDate	EndDate	MainFindings	Conclusion
1	SA001	T001	Descriptive Analysis	2024-07-01	2024-07-15	Increased antibody response observed.	Promising results for vaccine efficacy.
2	SA002	T002	Regression Analysis	2025-01-01	2025-01-15	Correlation between medication dosage and pain r...	Further studies recommended.

Triggers

```sql
-- Create the ParticipantChange table
CREATE TABLE ParticipantChange (
    ParticipantChangeID DECIMAL(12) NOT NULL PRIMARY KEY,
    ParticipantID VARCHAR(10) NOT NULL,
    FieldName VARCHAR(100) NOT NULL,
    OldValue TEXT,
    NewValue TEXT,
    ChangeDate DATETIME NOT NULL,
    FOREIGN KEY (ParticipantID) REFERENCES Participant(ParticipantID)
);
```

```sql
-- Create trigger to track changes in the Participant table
CREATE TRIGGER ParticipantChangeTrigger
ON Participant
AFTER UPDATE
AS
BEGIN
    DECLARE @ParticipantID VARCHAR(10);
    SET @ParticipantID = (SELECT ParticipantID FROM INSERTED);

    IF (UPDATE(ParticipantFirstName)) -- Check if ParticipantFirstName is updated
    BEGIN
        INSERT INTO ParticipantChange (ParticipantChangeID, ParticipantID, FieldName, OldValue, NewValue, ChangeDate)
        SELECT
            ISNULL((SELECT MAX(ParticipantChangeID)+1 FROM ParticipantChange), 1),
            @ParticipantID,
            'ParticipantFirstName',
            (SELECT ParticipantFirstName FROM DELETED),
            (SELECT ParticipantFirstName FROM INSERTED),
            GETDATE();
    END;

    IF (UPDATE(ParticipantLastName)) -- Check if ParticipantLastName is updated
    BEGIN
        INSERT INTO ParticipantChange (ParticipantChangeID, ParticipantID, FieldName, OldValue, NewValue, ChangeDate)
        SELECT
            ISNULL((SELECT MAX(ParticipantChangeID)+1 FROM ParticipantChange), 1),
            @ParticipantID,
            'ParticipantLastName',
            (SELECT ParticipantLastName FROM DELETED),
            (SELECT ParticipantLastName FROM INSERTED),
            GETDATE();
    END;
```

Results before trigger execution

```sql
select * from Participant;
```

100 %

Results | Messages

	ParticipantID	ParticipantFirstName	ParticipantLastName	Dob	Gender	MedicalConditions	MedicationHistory	Allergies	BMI	ParticipantType
1	P001	John	Doe	1990-05-15	Male	None	None	None	24.50	Adult
2	P002	Jane	Smith	1995-08-20	Female	Asthma	None	Peanuts	22.30	Adult
3	P003	Michael	Johnson	2005-03-10	Male	ADHD	Ritalin	None	18.80	Minor
4	P004	Emily	Brown	2000-11-30	Female	None	None	Pollen	20.10	Minor
5	P005	David	Wilson	1985-07-25	Male	Diabetes	Insulin	Penicillin	28.90	Adult

```
select * from ParticipantChange;
```

100 %

Results | Messages

ParticipantChangeID	ParticipantID	FieldName	OldValue	NewValue	ChangeDate

Results after trigger execution

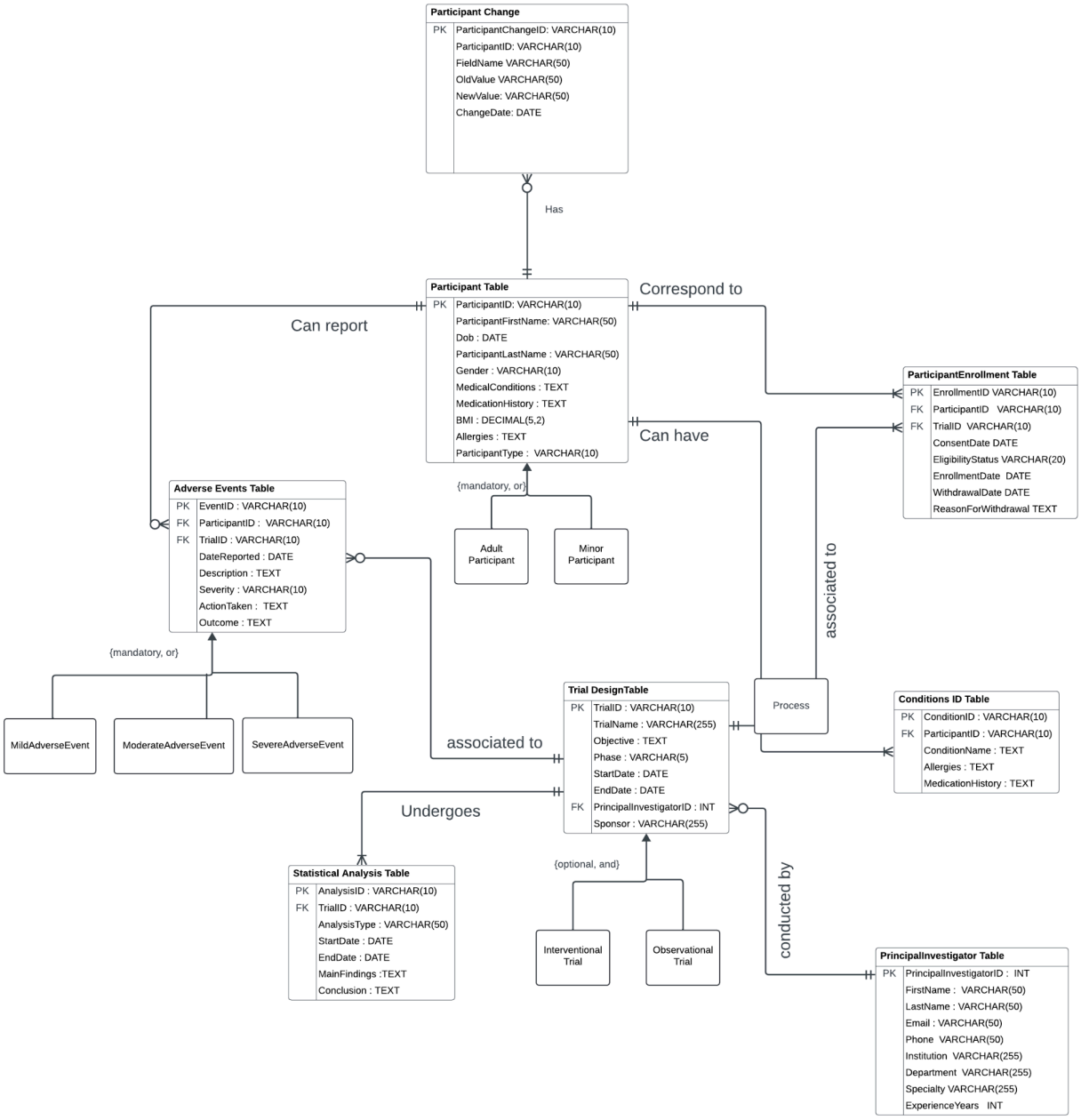
```
UPDATE PARTICIPANT
SET ParticipantFirstName= 'Jon'
where ParticipantID = 'P001';


UPDATE PARTICIPANT
SET ParticipantLastName= 'Snow'
where ParticipantID = 'P001';

select * from ParticipantChange;
```

100 %

Results | Messages

	ParticipantChangeID	ParticipantID	FieldName	OldValue	NewValue	ChangeDate
1	1	P001	ParticipantFirstName	John	Jon	2024-04-19 22:06:10.253
2	2	P001	ParticipantLastName	Doe	Snow	2024-04-19 22:06:37.997

Conceptual Clinical Trials ERD with Participant Change Table (History)



Leveraging the Database for Clinical Trial Insights

A well-structured clinical trial database can provide crucial insights that drive research efficiency, improve participant management, and enhance trial outcomes. Below are examples of a few queries and the valuable information they uncover.

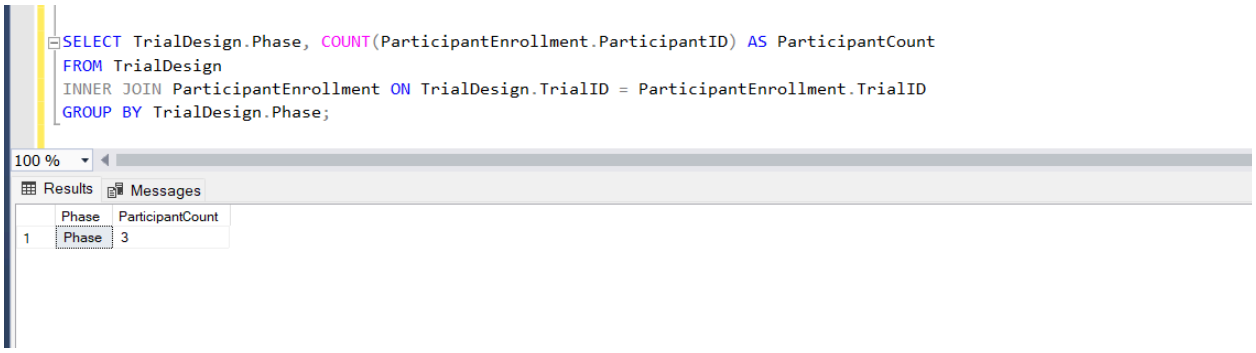
These queries enable data-driven decision-making and help clinical researchers manage trials effectively, ensuring participant safety and successful outcomes.

Question 1: How many participants are enrolled in each trial phase?

This query helps to understand the distribution of participants across different phases of trials, which can indicate the focus areas of research.

SQL Query:

```
SELECT TrialDesign.Phase, COUNT(ParticipantEnrollment.ParticipantID) AS ParticipantCount
FROM TrialDesign
INNER JOIN ParticipantEnrollment ON TrialDesign.TrialID = ParticipantEnrollment.TrialID
GROUP BY TrialDesign.Phase;
```



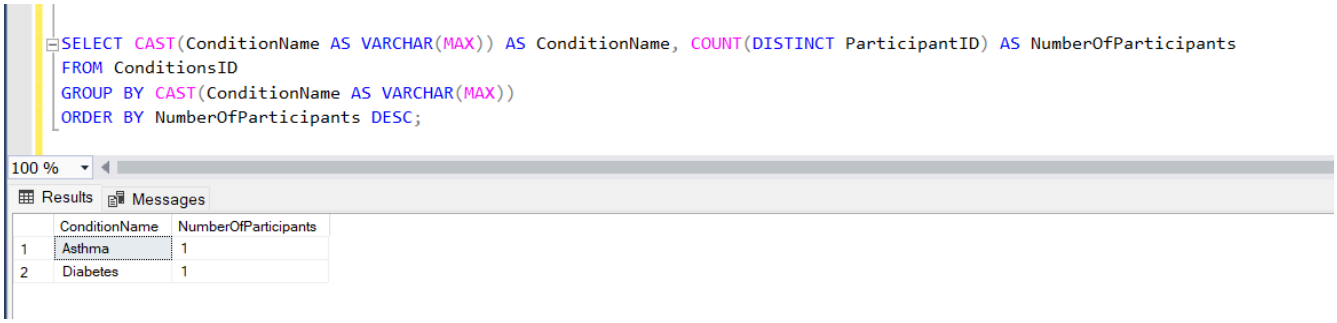
Question 2: What are the most common medical conditions among participants?

This query can reveal the prevalent medical conditions within the study population, aiding in targeted research and resource allocation.

SQL Query:

```
SELECT CAST(ConditionName AS VARCHAR(MAX)) AS ConditionName,
COUNT(DISTINCT ParticipantID) AS NumberOfParticipants
FROM ConditionsID
```

GROUP BY CAST(ConditionName AS VARCHAR(MAX))
ORDER BY NumberOfParticipants DESC;

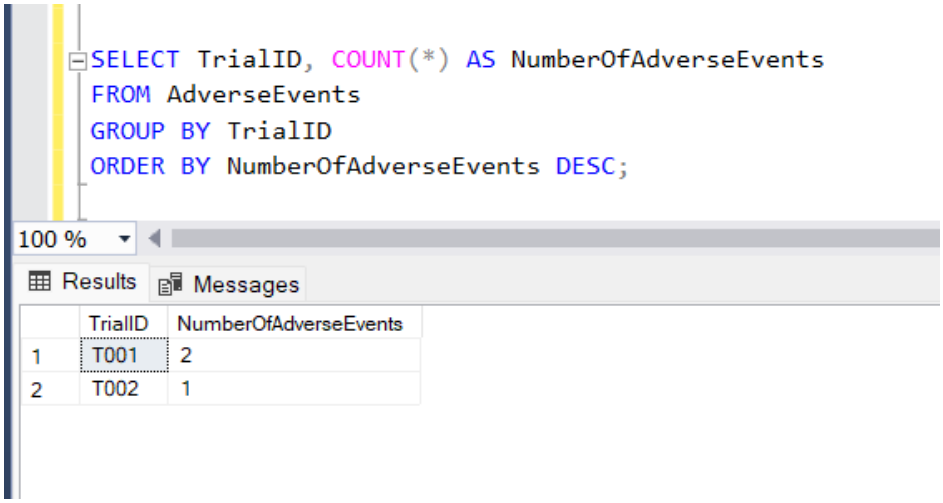


Question 3: Which trials have the highest number of adverse events reported?

Understanding which trials have higher instances of adverse events can help prioritize safety reviews and adjustments in trial protocols.

SQL Query:

SELECT TrialID, COUNT(*) AS NumberOfAdverseEvents
FROM AdverseEvents
GROUP BY TrialID
ORDER BY NumberOfAdverseEvents DESC;

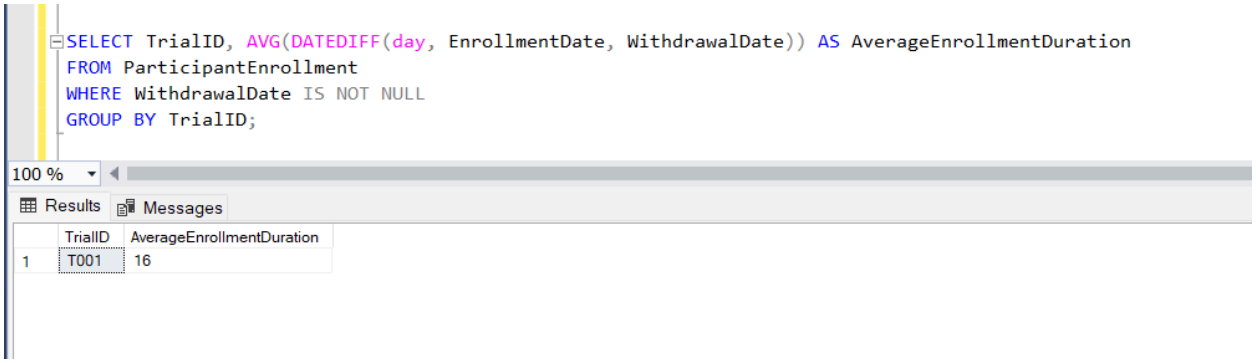


Question 4: How long do participants typically stay enrolled in a trial before withdrawing?

This query calculates the average duration of participation before withdrawal, which can be critical for understanding participant retention and trial feasibility.

SQL Query:

SELECT TrialID, AVG(DATEDIFF(day, EnrollmentDate, WithdrawalDate)) AS AverageEnrollmentDuration
FROM ParticipantEnrollment
WHERE WithdrawalDate IS NOT NULL
GROUP BY TrialID;

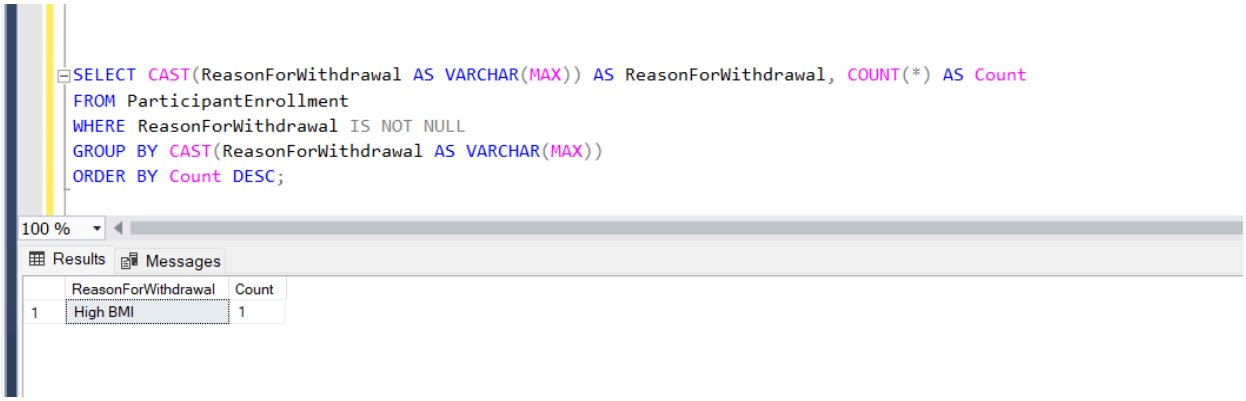


Question 5: What are the common reasons for withdrawal from the trials?

Identifying the reasons for participants' withdrawal can help improve trial designs and participant support strategies.

SQL Query:

SELECT CAST(ReasonForWithdrawal AS VARCHAR(MAX)) AS ReasonForWithdrawal, COUNT(*) AS Count
FROM ParticipantEnrollment
WHERE ReasonForWithdrawal IS NOT NULL
GROUP BY CAST(ReasonForWithdrawal AS VARCHAR(MAX))
ORDER BY Count DESC;



Appendix-A

Link to ERD (LucidChart)

https://lucid.app/lucidchart/2d9d33e7-2ea4-42b1-a4e3-3b5a9a863329/edit?invitationId=inv_82cdf84e-d2af-4faf-b51f-411d3f9ac4a2&page=0_0#