Law, Ethics, and Security Plan of Database Management System (DBMS)
for
Westlake Research Hospital Clinical Trial

Charles Adkins

JUNE 5, 2020

Analysis of Organization
Westlake Research Hospital

Organizational and Scenario Background

Westlake Research Hospital is conducting an 18-month, double blind study on the effectiveness of a new depression treatment drug. The test study will involve two supervisors, 20 doctors and approximately 400 patients. Half the patients will receive the new drug and the other half will receive traditional Prozac. The term 'double blind' refers to the fact that neither the patients nor the doctors will know who receives which drug, as it is only known by the two observing supervisors who administer the drug and monitor the study.

The patients will have checkups and interviews conducted with their assigned doctor twice a month. The checkups and interviews will gauge the patients personal behaviors, mental health and incurring side effects, if any. Specific data-viewing restrictions will in place, of which include the following: Doctors are only permitted to view their assigned patients information and enter any relevant results and notes. Patients are permitted to view only their own information. Supervisors are granted full access to the information of the study. In addition to the data access restrictions, there are study restrictions which assign the patient to a single doctor for the duration of the study.

Challenges, Requirements and Limitations

The current process of collecting and storing patient data is lengthy and vulnerable. This process initiates with the office secretary emailing each doctor a table consisting of their patients and respective appointment information. The doctor prints the table out and gathers each

patients file from a storage place and documents relevant information prior to the patient arriving. The doctor visits with the patient and proceeds to document the appointment findings to the patients file and returns it to storage.

Due to the complexities of the study in terms organization, storage and security, a database is required to effectively establish and maintain an appropriate level of integrity and validity. The database will not only solve these aforementioned issues, but it will also provide a viable means for reflection and reproducibility upon completion of the study. The ideal conclusion of the study involves results that suggest the new drug is more effective than treatment carried out by Prozac and in order to obtain a proliferating acceptance of such findings, several peer-reviewed studies will need to subsequently take place with similar findings. The nature of the database will provide other institutions and personnel with centralized, reference data to conduct similar tests, for the purpose of comparison.

The requirements of the database include establishing a centralized entity that houses all relevant study data and provides the necessary security and authorization specifications to maintain the standards of confidentiality and the aspects of a double-blind study. In reference to possible, subsequent peer-reviews and clone-studies, HIPAA Privacy Regulations and Data Use Agreements, will play a key role in upholding patient privacy and data security. “A Data Use Agreement is needed when a researcher wants to share PHI in the form of a Limited Data Set (defined as a data set that contains no identifiers other than certain "indirect identifiers") with someone not otherwise involved in the research protocol, such as a colleague at another institution” (Johns Hopkins Medicine, 2015). The utility of the system may be harnessed in

future endeavors, given that it adequately provides the institution and respective study stakeholders with final deliverables that meet their expectations.

Organizational Impact

As previously mentioned, secretarial duties have been effected by the needs of the study, which has hindered their ability to efficiently attended to other administrative duties within their respective department. Given that each patient visits the hospital twice a month and that there are ~400 patients, secretaries have found themselves endlessly emailing patient and appointment information to doctors, rather than having the ability to optimize the function of a database and client application interface.

Doctors have also felt the strenuous nature of the administrative requirements warranted by the study. Without the database, doctors have been spending resources and time locating and reviewing relevant information from disparate sources, rather than conducting such effort within the confines of a centralized and lean place of such information. With the database, doctors will also be able to quickly enter, edit and review information within the system with ease and have the ability to analyze and compare such information in an efficient manner.

Sponsors of the study, which include the Westlake Research Hospital and pharmaceutical companies, are also impacted by the requirements of the study and the need of a database system. Both stakeholders are financially invested in the research and expect the study to produce results that can help them pursue new lines of treatments, of which will be based upon the results of

studies, much like this one. The results must be accurate, unbiased and objectively conclusive in the way that they characterize the effects of the drug.

Patients are also impacted by the need for a database. They have found themselves in search of emotional and cognitive relief, by which they have placed such hope within the study and the abilities engineered to be contained within the new drug itself. Patients rely not only the drug, but also their doctor ability to concisely document their progress throughout the study. The validity of the study relies on both the patients ability to communicate their observations and experiences, but also the doctors ability to translate that information into the formatted guidelines outlined in the study requirements.

The database will be the place in which all stakeholders will be able to clearly identify, review, understand and trust the information from the study, with the ultimate goal being to produce results and necessary support to move forward in developing and implementing treatment options that provide excellent care to patients.

Analysis and Design

Conceptual Model

Westlake Research Hospital requires a database that can help manage, organize and maintain the data collected during their clinical trial. The trial is, as previously stated, a double-blind study aimed towards examining the results of a new depression treatment drug. Specific authorization protocols will need to be in place within the architecture of the database and the

associated application, to maintain the levels of integrity and security the trial needs, of which will be developed during a later part of the project.

Conceptually, the database has several elements, through which was communicated and understood by way of interviews and business discussions with stakeholders. These required elements or entities, will store valuable data that will be utilized to analyze the results of the study and as such, must be constructed in a sophisticated manner, with appropriate attributes and relationships to make the retrieval of such data accurate and effective. Attributes and data types are discussed in more detail in the logical and physical models. At the highest level, a conceptual model (**APPENDIX A**) has been developed, which contains the following entities:

Patients,	**Appointments,**	**Test Results,**
Addresses, Doctors,	**Recommendations,**	**Symptoms, and**
Drugs, Supervisors,	**Notes, Medications,**	**Conditions.**

These entities cover the basis of all subject lines of data that will be collected throughout the lifecycle of the clinical trial. These entities have specific relationships between themselves, as noted by the respective arrows in the model diagram. The arrows indicate that the primary entity has associated data with the end-point entity. This covers the nature of their relationship, but the dynamic nature of the relationship is explained in more detail in the logical and physical models, in terms of relationship types.

For example, to clarify conceptual model relationships, take the ‘Patients’ entity (centered in the diagram). The ‘Patients’ entity has ‘Appointments’ (bottom-center of diagram), thus the arrow displays the relationship ‘Patients’, being the primary entity, has with ‘Appointments’. A more complex relationship can be seen with ‘Patients’, ‘Notes’, ‘Doctors’ and ‘Recommendations’. Both ‘Patients’ and ‘Doctors’ have a relationship with ‘Notes’ and ‘Recommendations’, as a Doctor writes notes about the Patient, makes recommendations based upon those notes and associated test results, while at the same time, the Patient has access to review their personal Notes and Recommendations. The conceptual model arrows communicate these relationships in the way that they connect their respective primary entities and the associated end-point entities. These relationships, along the remaining relationships amongst the other entities, has been established based upon the business rules of the project requirements. The business rules, as they are defined as the restrictions or constraints, imposed by the organization, on certain aspects of a database, based on the ways they perceive and intend to use the data (Hernandez, 2013), help develop the design of the database and how the entities relate to each other.

In this project, there are several business rules and they are as follows:

- Each Patient will be assigned to a single Doctor
- Each Patient enters their own medical information
- Each Patient will visit their respective doctor twice a month
- Patients are authorized to view only their medical information, test results and the notes and recommendations made by their doctor
- Doctors are authorized to view only their patients medical information, test results, notes and recommendations
- Supervisors are authorized to view all information contained in the database
- Each Doctor will have multiple Patients
- Each Doctor will document each Patients visit information, test results, notes and recommendations
- Each Patient will be administered one drug, of which type is known only by the supervisor researchers

Logical Model

The logical model displays the same entities and relationships as the conceptual model, but includes more detail about each entity (**APPENDIX B**). These details include entity attributes, attribute data types and the specific nature surrounding the relationships each entity has with other entities. The logical model also displays foreign keys, that which link one table to another, for instances that involve one-to-one or one-to-many relationships. For instance, take

the ‘Patients’ entity table. Patients now has several attributes: PatientID, FirstName, LastName, BirthDate, etc. These attributes are also accompanied by specified data types: INT, VARCHAR(45), VARCHAR(45), DATE, etc., respectively. These attributes are line items that are to be filled with appropriate data, within the confines of their respective data types. For example, FirstName is the Patients first name, which will be input with variable characters (letters), up to a limit of 45. Foreign keys, as briefly mentioned, are also displayed in the Patients entity table, such as AddressID, DoctorID, DrugID. These foreign keys directly link a specific patient to specific observations with those entity tables. The Patient is linked to a specific Address in the Address table, a specific Doctor (limited to one) in the Doctors table and a specific Drug (limited to one), in the Drugs entity table.

In the case of many-to-many relationships, linking tables have been incorporated into the logical model, which were not displayed in the conceptual model. These linking tables are established by utilizing the primary key attributes (in the example of ‘Patients’ and ‘Appointments’, these primary keys would be ‘PatientID’ and ‘Date’, respectively) of the related entities, as the primary attributes in the linking table itself. Again, the linking table between ‘Patients’ and ‘Appointments’ would be an entity called ‘Patients Appointments’, with its’ attributes being ‘PatientID’ and ‘Date’. In this way, the linking table associates records from each entity table and helps ensure that there will be no issues when adding, updating or deleting data (Hernandez, 2013).

Physical Model

The physical model is the last stage in the data modeling phase of the data design process. At this point we have developed a sound layout for the database, included necessary entities and established appropriate relationships between the entities based upon the business rules, technical requirements and project objectives. In the physical model (**APPENDIX C**), these entities have been transformed into tables (**APPENDIX D**) and are now fully capable of receiving and storing data for the purpose of testing. The physical model, as the logical model did, showcases the entities (now tables), their respective relationships with other tables in the form of foreign keys and linking tables and the attributes of each table has been transformed into columns. With these new modifications, the physical model can be given sample data to test the integrity of the design and queried to establish whether or not the architecture meets the technical needs and business rules required. This physical model was developed utilizing Microsoft SQL Server Management Studio (SSMS). In this particular RDBMS formatting, the model showcases a different syntax for communicating table relationships, as opposed to the logical model, which was done in MySQL Workbench. In the physical model, the tables are connected by key symbols and infinity symbols. The infinity symbol represents the “many” side of the relationship, while the key represents the “one” side of the relationship.

DBMS Research and Recommendation

Database Management Systems

When developing a database, one must consider what database management system (DBMS) will be utilized to carry out the functions of the database. There are several different DBMS products available in todays' market, all of which complete similar tasks, but with slight variations in their respective areas of execution. DBMS programs are designed to perform database operations of which include creating new database structures, storing data, providing tools for data retrieval, administration actions, report generation, information queries and establishing and enforcing security protocols (Learn Computer Science, 2019).

It's important to select a DBMS that fits the technical requirements of the data and meets the needs of the organization that intends to utilize the database. These technical requirements may include differences in database architecture, as well as the nature of the data contained within it. More specifically, there are two types of DBMSs, relational and non-relational. The differences between the two types involve their execution of data retrieval, distribution and processing, also referred to as SQL and NoSQL, respectively (Altexsoft, 2019). In a relational or SQL DBMS, the data appears as tables of rows and columns that displays its framework and dependencies clearly, which makes traversing the database a systematic operation. A non-relational or NoSQL DBMS, relies upon a document-oriented structure, where various data types are collected on to documents rather than structure theme-based tables, which makes querying simple and scalability easier (Altexsoft, 2019).

In the Westlake Research Hospital Clinical Trial Use Case, the database prototype has been developed utilizing Microsoft SQL Server Management Studio, a relational DBMS, where the database appears as tables with rows and columns. We will explore several other DBMS products available in the following sections by addressing their similarities and differences, as well as their applicability, strengths and weaknesses. After analyzing these products, a selection will be made for the Westlake Research Hospital Clinical Trial database and the conclusion will include the justification for said selection and any additional software and/or hardware necessary to complete the project.

Relational Database Management Systems (RDBMS)

As previously mentioned, Relational or SQL DBMSs are different from their counterpart, non-relational DBMS, in their structuring and retrieval execution and as such must be analyzed to assess their applicability towards ones specific database and organizational needs. Examples of SQL DBMS products include MySQL, Oracle and MSSQL.

MySQL

MySQL is an open-source DBMS solution that is highly versatile in system compatibility and provides several benefits such as free installation, simple syntax and gentle learning curve (Altexsoft, 2019).

Strengths

MySQL is low risk investment that provides the functionality that a simple, small-to-medium sized database would require, of which can be managed and operated by novice users.

Weaknesses

On the downside, MySQL does lack scalability, as SQL databases do require extensive and complex engineering efforts to expand and in addition to this, MySQL requires its users to pay for support and the SQL standards upheld by MySQL differ from traditional protocols, which could present an issue if the project ever needs to be transported to other systems (Altexsoft, 2019).

If the database and the requirements for a DBMS remain simple and the expectation is that the entire project remains static throughout its lifecycle, MySQL may be a great option for the Westlake Research Hospital Clinical Trial.

Oracle

Strengths

The Oracle DBMS is a robust, highly scalable and sophisticated system. The system operates with vertical scalability features, of which reduces power and software requirements, licensing costs and maintains a centralized system base (Dihuni, 2020).

Weaknesses

However, depending on the expectations surrounding the data volume, velocity, variety, veracity and its' anticipated user base, the centralized nature of its vertical structure can lead to hardware failures and upgrade restrictions, while at the same time, the implementation of this system is costly (Dihuni, 2020).

The typical application of Oracle DBMSs include large enterprise endeavors where traffic is large, scale is expansive and costs are given quick attention. Given these examples, it appears

that Oracle DBMS, although commendable in its technology and reliability, may be outside the realm of plausible for Westlake Research Hospital.

MSSQL

Microsoft SQL Server, as it's also known, is a popular DBMS for a number of reasons that makes implementation and management simple and reliable. This DBMS is both cloud-based and local server operated and is compatible across Linux and Windows-based systems (Arsenault, 2017).

Strengths

Benefits of utilizing MSSQL include its' processing speed, performance throttling capabilities, which adjusts to incoming/outgoing loads, and its' unique ability to track changes made to the data and the "dynamic data masking, which ensures that only authorized individuals will see sensitive data" (Arsenault, 2017).

Weaknesses

Although the system does have several impressive features, it does require significant computing resources and large enterprise pricing may be a deterrence for small organizations.

MSSQL has the reliability, the performance measurements and security protocols that match specific needs for the Westlake Research Hospital Clinical Trial, which makes it a great candidate for selection.

Non-Relational Database Management Systems

As previously discussed, non-relational DBMSs, or NoSQL DBMS, have different applications, disadvantages and advantages than SQL DBMS. There are several NoSQL products available worth investigation, but we will focus on MongoDB and Cassandra.

MongoDB

Strengths

MongoDB is an open-source system that operates with both structured and unstructured data (Altexsoft, 2019), meaning that it can utilize both formatted data that is organized for ease of search criteria and unformatted data, which requires less prerequisite, organizational work, but does take more processing time. Given its processing versatility, MongoDB is also equipped with database drivers, which connects the databases with the applications and allows for simultaneous processing of various data types by way of internal caching (Altexsoft, 2019), which is quite useful when the database is being repeatedly "pinged" by its users over a short period of time.

Weaknesses

Given the non-relational nature of MongoDB, vast computer memory requirements and decreased performance seem to be downfalls for the system (Altexsoft, 2019). This requirement for extensive memory is due to the document-oriented nature of non-relational systems, as it does not rely on table structures but rather all values being dispersed across the system for speed of access. In addition to this, heightened data security protocols are available only in the commercial edition, which does increase its cost.

MongoDB has been seen to be an effective DBMS choice for real-time data integration and analytics platforms, such as product catalogs (Altexsoft, 2019), but for the Westlake Research Hospital Clinical trial, it seems to not fit the bill due to instability in security and lack of efficacy in this particular scenario.

Cassandra

Cassandra is a unique DBMS which provides several features that make it a great choice for applications that need scalability, high performance and availability.

Strengths

Cassandra utilizes a linear-scalability approach (The Apache Software Foundation, 2016), meaning performance is consistent with data loads, more specifically, smaller data sizes requires less query time vs. larger data requires more query time (Shimoni, 2017), which supports the systems overall reliability. Another benefit of Cassandra is its fault-tolerant node structure, where nodes contain replications of data, which prepares and adjusts to any instance of node failure (The Apache Software Foundation, 2016), without the worry of losing value.

Weaknesses

Downsides to the system are its slow reading power and its need for additional software and hardware, which make it the ideal system for data centers, where security are key and input velocity is high and output velocity is less of a focus (Altexsoft, 2019). With these characteristics, although Cassandra makes for an interesting discussion and is innovative in its ability, it appears to be out of focus for the Westlake Research Hospital Clinical Trial database.

Recommendation

MSSQL is the recommended DBMS solution that meets the needs of the Westlake Research Hospital Clinical Trial database. MSSQL offers a tier functionality option, which increases in dimensionality and required expertise as you climb, which includes the Express version, Developers option, and larger project editions such as the Web, Standard and Enterprise options (Altexsoft, 2019). Given the nature of the clinical trial, the Express or Developer options may be more suitable as Express is perfect for entry-level design and small, server applications, while the Developers provides testing options and Enterprise functions, without the production license (Altexsoft, 2019). The clinical trial is an internal event with no requirement to deploy, which, in concern of cost, would eliminate the more expensive options. Additionally, in the event of system complications or development/implementation roadblocks, MSSQL has rich documentation and community assistance (Pijacek, 2019), which is a helpful advantage depending on the skill level and familiarity of its users.

Hardware and Software

Based upon the included documentation, support and internal resources provided by MSSQL, there are virtually no hardware or software additions required, beyond MSSQL itself. MSSQL is a Microsoft product, but the 2017 version is Linux supported and has several helpful distributions that elevate the need to transform any servers licenses. The learning curve for MSSQL is steep and will take time for its users to become familiar with its interface, but with time, it will become second nature.

Enterprise Data Model

Enterprise Data Model Overview

To further assist Westlake Research Hospital in their current endeavor of developing and deploying a database for their clinical trial, an Enterprise Data Model has also been developed. The Enterprise Data Model (EDM) is designed with domain specific characteristics that places emphasis on unbiased integration. The EDM enables the identification of shareable and/or redundant data across functional and organizational boundaries, which unites, formalizes and represents the things important to an organization, such as data quality, consistency, accuracy and cross-functional extensibility (Kendie, 2005). The EDM, given its' domain specificity, yet lack of system/application bias, organizes itself to be implemented in various circumstances which require the specifications that the EDM provides. "The model can be thought of much like an architectural blueprint is to a building; providing a means of visualization, as well as a framework supporting planning, building and implementation of data systems" (Kendie, 2005).

In the case for reoccurring needs of data systems, the EDM serves as a cistern of sorts, which provides the framework for such events, eliminating the need to redevelop such architecture, due to its deployable readiness. The design and development of the EDM is conducted in an iterative process, by which each level of creation advances in specificity and individual purpose.

EDM Development Levels

The design and development of the EDM is built in three levels, each designed to produce its own unique and detailed deliverable, which provides its users with insight as to how it was accomplished and offers the flexibility to tailor its construction to their needs. The three levels include the Enterprise Subject Area Model (ESAM), Enterprise Conceptual Model (ECM) and Enterprise Logical Model (ELM), also known as the Enterprise Conceptual Entity Model (ECEM) (West, 2011).

Enterprise Subject Area Model (ESAM)

The ESAM captures key concepts of the organization and transforms them into subject areas that will be the foundational entities for the model design. These subjects take on basic geometric shapes and display how they interact with another, or their basis for being established, as these, in theory, will not change (Smith, 2011). The ESAM for Westlake Research Hospital is included in **Appendix E**. The ESAM lays out the foundational structure and the subjects that are critical to the current clinical trial implementation, as well as any future trial events. This model will be the data system blueprint for all subsequent trial studies.

Enterprise Conceptual Model (ECM)

The ECM establishes the identified subjects areas by incorporating high-level subject concept details and displays the nature of their relationships. The details regarding the purpose of each subject concept are confirmed in the ECM by conceptualizing the subjects with data requirements and associations (Kendie, 2005). The details of these subject concepts is conducted

in an agile approach, as it may need tailoring, due to the needs of the users and the integration of particular data. To ensure these needs are met and appropriately addressed, review sessions can be conducted, which verify and approve the development. From these sessions, documentation is created, describing enterprise overlap, conflicts, and data integration issues or concerns (Kendie, 2005). The prototype of the ECM is included in **Appendix F**.

Enterprise Conceptual Entity Model (ECEM)

The ECEM is the third level in the EDM process, as it concludes the design and development of the model and displays the most detail required by the system. In this level of development, the constraints and rules governing the system play a vital role in the functionality of the system, of which are properly communicated and presented in the form of the ECEM. The deeper level of detail included in the model, based upon organizational requirements, the more useful and sustainable the model will prove to be. The entities within the model represent the business lines and the data requirements for the organization and are "independent of technology and implementation concerns (Kendie, 2005). The included aspect of designating entity primary keys reflect the business/operating rules that uphold the structural environment of the system itself. "A key validates business rules; as entity concepts are related and keys are inherited, they must continue to work correctly" (Kendie, 2005). The relationships portrayed in the model also reflect business rules as they reflect the interdependency of the entity network. Overall, the "ECEM provides a data architectural framework for the organization's data designs and subsequent data stores, in support of data quality, scalability and integration" (Kendie, 2005). In

this way, the model can be used time and time again, establishing a trusted framework for system implementation and event support. The ECEM is included in **Appendix G**.

Enterprise Data Model Operating Rules

Given the nature of the database, Westlake Research Hospital will utilize the system for all of its clinical trial data needs and as such, the operating rules established therein surround the aspects of trial requirements. The database must collect and maintain sensitive data regarding patient history, current ailments and conditions, as well as trial test results. This establishes a baseline level of understanding of the patients, used for comparative analysis at the conclusion of the trial. The patients are observed by the doctors and are done so in periodic fashion, established by appointments, of which such interactions and observations are maintained in a log of notes and recommendations. The doctors and researchers involved are key stakeholders in the trials as they are front-line observers and data collectors, maintain the trustworthiness of the trial. The trials themselves are based upon scientific notions, rooted in research, surrounding new lines of medications. These medications should be linked to the specific trial and a record of such should be maintained for trial integrity and reproducibility.

Enterprise Data Model Rule Reflection

In this particular trial, there are numerical constraints placed upon the participants, such as each patient should be attended to by only one doctor and the doctor can have no more than twenty patients. Furthermore, the nature of the study, being double-blind, the details regarding the medication being researched and administered, is shielded in mystery for both the patients

and doctors, to enforce unbiased trial events. The procedures that maintain entity security and privacy, may differ in subsequent trials, which is why the model refrains from establishing such concrete restraints within the system. These restrictions can be developed and tailored to the parameters of each trial, rather than hard-coding it into the model itself. Given that each trial may involve different researched medications (drugs), the trial event should also be maintained and given relations to the appropriate medication (drugs). These operating rules are requested by Westlake Research Hospital and are accurately reflected in the design and development of the EDM.

Law, Ethics, and Security

Standards

Given the sensitive nature of the data being collected and maintained within the database, more specifically, Personally Identifiable Information (PII) & Protected Health Information (PHI), it is a critical that relevant legal and ethical standards be appropriately considered and implemented to enforce protective measures to uphold the integrity of the data and to respect the privacy of the patients. To appropriately enforce relevant measures to protect said data, we must understand the roles of those who will be interacting with the data and the ethical responsibilities of those particular actors. Such responsibilities include being transparent regarding how the data will be maintained, processed and used, in addition to the limitations of said use (Debussche & César, 2019). From a legal standpoint, the data must adhere to specific laws and regulations,

later discussed, of which follow and include aspects of consent, disclosure and reporting (Datical, 2020).

Legal Compliance

Best Practices for Legal Compliance

Designing the database, declaring the specifics surrounding data use and the procedures and policies regarding storage, are all elements that must be considered for establishing system best practices, which ensure legal compliance. Designing the database and declaring the specifics surrounding the data use are established within the scope of the project itself and its purpose.

Design

The design follows specific business and system requirements that enable Westlake Research Hospital to collect, maintain, protect and analyze patient data for the purpose of the clinical trial research. The data necessary to conduct such research has been included in the design, displayed in the form of tables and their respective rows and columns, all of which has been discussed and understood by all participants and system actors. Each stakeholder has an understanding of their role, responsibilities and limitations, all of which reflect the business requirements and overarching organizational value system and legality adherence policies.

Data Use

The patients should be given a Notice of Privacy Practices (NPP), which informs the patients of the uses and disclosures of their PHI, as well as their rights to access and amend their medical information (OfficeSafe, 2016), while still remaining within the agreed upon terms and

conditions outlined in the structuring of the clinical trial requirements. The HIPAA Security Rule requires organizations to implement administrative, physical, and technical safeguards to ensure PHI is stored, transmitted and received in a safe and secure manner (OfficeSafe, 2016). The Westlake Research Hospital undoubtedly has these concepts in place throughout the appropriate departments of the organization, but it must also be adhered to and maintained within the scope of the clinical trial arena, where elements of concern are relevant. Established procedures and policies surrounding data access and use should be clearly outlined and enforced, as they are declared in the business requirements.

Storage

Additionally, security policies that cover encryption, decryption and backup planning should be in place, as well as data storage and disposal policies and procedures, which establish data shelf-life and protocols for such disposal (OfficeSpace, 2016). When the data is no longer needed or has an organizational function, it must follow the set of established procedures to ensure it is properly destroyed.

Regulations and Legal Compliance

The most important reason an organization must meet and reside within database compliance requirements is to ultimately remain in business. This can be achieved by understanding the requirements and regulations, adhering to them, avoiding monetary penalties issued due to non-complaint actions, all of which maintain their reputation as a respected data steward whom prevents exposing data assets to attackers.

Specifically, Westlake Research Hospital must understand, practice and adhere to the policies and procedures outlined in Title II of the Health Insurance Portability and Accountability Act of 1996 (HIPAA). "Title II of HIPAA establishes policies and procedures for safeguarding the privacy and security of individually identifiable health information to control fraud and abuse in the health care system" (Datical, 2020). The patient data collected and utilized in the clinical trial must be kept confidential and maintained within the realm of the research arena. At the end of the trial, when being utilized for peer-review and publication, the PII must be protected against identity theft and the PHI must be protected within the confines of the HIPAA Title II policies and procedures. Following these policies and procedures include aspects of consent and disclosure. Specifically, organizations under HIPAA may not use or disclosed PHI without the written consent of the patient, with exceptions in specific situations that include, (1) third-party disclosures authorized by the patient, (2) for treatment, payment or general healthcare operations, and (3) if the individual has the opportunity to agree or object to a disclosure (OfficeSafe, 2016). These control measures protect not only the patient, but the organization from liability concerns surrounding breeching of such regulatory laws. To help administer these safeguards, HIPAA requires that every practice designate a HIPAA Security and Privacy Officer, whom lead the implementation and training of HIPAA requirements (OfficeSafe, 2016).

Ethical Practices

Ethics as a global term, involves meditating on ones moral aptitude and carrying out actions that reflect principles which are just and trustworthy (BBC, 2014).

Design

In terms of database design, ethical considerations can be taken to into account in order to uphold just and trustworthy principles. For example, the structuring and access displays of the database should incorporate ethical aspects, as to present its users with accurate and transparent information, within the constraints of the views established by the system. The data that which has been established to be accessible by certain members, should indeed be accessible to them members, in full. Withholding or restricting viewable information or manipulating the retrieval efforts of its users, beyond the scope of the aforementioned viewing constraints, would be unethical.

Data Use and Storage

The ethics behind HIPAA ensure that a persons data be well-protected, as well as confidential. The agreement established between an organization and its patients, clearly identifies and communicates the use and storage of the data be utilized and how it is being leveraged for a specific task. This agreement is rooted in just and trustworthy notions that deal with privacy, access, disclosure, implementation and disposal (UK Essays, 2017). The agreed upon terms will be laid out, discussed in full detail and agreed upon, by way of review and contract signature, prior to the initiation of the clinical trial. This establishes a tracking history and record by which the work to be done, was transparently discussed and agreed upon. These best practices will likely change depending on the industry in which your company resides.

Security Needs of Solution

The clinical trial database management system requires security measures that are able to maintain the privacy and confidentiality of its users data and to deter and defend the value contained within from malicious attacks. The integrity of the clinical trial rests on the ability to secure patient and trial data, while also adequately documenting the process, of which will support trial conclusions during review and any iterations of reproduction. Specifically, in terms of deterrence, implementing perimeter security and defenses such as firewalls and intrusion detection systems/intrusion prevention systems (IDS/IPS) would provide a barrier against attacks and offer peace of mind (UK Essays, 2017).

Routinely assessing these systems and patrolling areas for vulnerabilities will ensure the system is maintaining coverage and such assessment logs can be useful in times of review or audit. Utilizing encryption/decryption keys and monitoring and auditing user authentications can also add a layer of security that protects the data and its users (UK Essays, 2017). Verifying user roles and access authorizations will be critical throughout the clinical trial, as the nature of the study requires there to be restrictions in data access for particular users, specifically doctors and patients. A primary outcome of database security is the effective limitation of access to your data. Access controls authenticate legitimate users and applications, limiting what they can access in your database. Access includes designing and granting appropriate user attributes and roles and limiting administrative privileges (Looker Data Science, 2020).

Clinical Trial Database Compared to Hospital Database

The requirements of the clinical trial database management system are simple and clearly defined. The overall purpose of the database is to collect, store and maintain specific patient and doctor data for analysis and reflection regarding the effects of a new depression drug. The number of users has been explicitly numbered and the details surrounding authorizations and restrictions are outlined in the project requirements and reflected in the design and development of the system. The database is subject to scheduled changes and periodic reviews, specifically during appointments and supervising researcher reviews. The nature of the data, although sensitive in nature, is predominantly uniform, as the complexity, as previously mentioned, is simple, while the variety and volume is internally governed. Given these dimensions and methods of utility, the nature of this particular database requires less dynamic security measures and protocols that other databases that may be found in other departments of the hospital.

Database security safeguards defend against a myriad of security threats and can help protect your enterprise from deployment failure, excessive privileges, privilege abuse, platform vulnerabilities, unmanaged sensitive data, backup data exposure, weak authentication and database inject attacks (Looker Data Science, 2020). These aspects are aspects that can also be areas of concern for the clinical trial database, but in terms of the entire hospital, these aspects are quite large obstacles to approach and accomplish. The nature of data being collected throughout the entire hospital includes a variety of data types with vastly different levels of sensitivity, being taken in at a higher rate of velocity and volume, with potentially more immediate-granting value (Dillon, 2020). The hospital is responsible for all patient, employee, organizational and partnership data. There are several dimensions of financial and payment data

assets being stored in the enterprise database as well. The hospital is responsible for safeguarding the data, protecting it from both internal and external threats, properly storing and utilizing it, as well as training all of its stewards on proper handling and dissemination protocols. In comparison to the requirements of the clinical trial database, the hospital requirements are evermore increasingly complex and robust, all of which warrants a more comprehensive security plan.

Database Security Plan

Given the nature of the data and the established requirements of the clinical trial, protecting the database management system and the data within it is of the utmost importance to maintain not only the privacy of the patients and other stakeholders, but also to uphold the integrity of the study. There are several steps that should be emphasized and practiced to ensure such measures are efficiently carried out. They include, but are not limited to:

- Ensure log-in credentials are strong in complexity, to deter unauthorized access
- Enforce authentication and viewing constraints (stored procedures) for specified roles and ensure they are being routinely verified
- Enforce encryptions to all data
- Keep all relevant software and protocols updated and current
- Maintain a secure storage of sensitive data (e.g., use strong passwords, install firewalls, intrusion prevention and intrusion detection systems) (UK Essays, 2017).

- Established the Disaster Recovery Plan and disseminate the information to all internal stakeholders (Conger, 2014).
- Backups (database, log files, system applications) should be conducted every 12 hours and stored off-site
- Routine inventory of backup assets should be verified monthly
- Document all database changes, updates, log reports, etc. for integrity and reproducibility purposes

References

Altexsoft. (2019). Comparing Database Management Systems: MySQL, PostgreSQL, MSSQL Server, MongoDB, Elasticsearch and others. Retrieved May 7, 2020 from https://www.altexsoft.com/blog/business/comparing-database-management-systems-mysql-postgresql-mssql-server-mongodb-elasticsearch-and-others/

Arsenault, C. (2017). The Pros and Cons of 8 Popular Databases. Retrieved May 5, 2020 from https://www.keycdn.com/blog/popular-databases

BBC. (2014). Ethics: A General Introduction. Retrieved May 29, 2020 from http://www.bbc.co.uk/ethics/introduction/intro_1.shtml

Blaha, M. (2017). Enterprise Data Modeling. Dataversity. Retrieved May 23, 2020 from https://www.dataversity.net/enterprise-data-modeling/#

Conger, S. (2014). Hands-On Database. Requirements and Business Rules. (2nd ed.). Pearson. Retrieved April 18, 2020 from https://mbsdirect.vitalsource.com/#/books/9780133927078/cfi/6/22!/4/2/2/2/2@0:0

Datical. (2020). Database Compliance & Security: What You Need to Know. Retrieved May 29, 2020 from https://www.liquibase.com/blog/database-compliance-security

Debussche, J. & César, J. (2019). Data-related legal, ethical and social issues. Bird&Bird. Retrieved May 29, 2020 from https://www.twobirds.com/~/media/pdfs/eu-data-economy-legal-ethical--social-issues.pdf?la=en&hash=5E348656C2F87363DD3695F8F06C99E53CEA6B21

Dihuni. (2020). Database Scalability: Vertical Scaling vs Horizontal Scaling. Retrieved May 9, 2020 from https://www.dihuni.com/2020/04/10/database-scalability-vertical-scaling-vs-horizontal-scaling/

Dillion, K. (2020). Big Data & the 5Vs. Retrieved June 2, 2020 from https://coservit.com/servicenav/en/big-data-the-5vs/

Hernandez, M. (2013). Database Design for Mere Mortals. A Hands-on Guide to Relational Database Design. (3rd ed.). Addison-Wesley. Retrieved April 15, 2020 from https://mbsdirect.vitalsource.com/#/books/9780133122275/cfi/6/22!/4/2@0:0

Johns Hopkins Medicine. (2015). Office of Human Subjects Research - Institutional Review Board. HIPAA and Research. Retrieved April 16, 2020 from https://www.hopkinsmedicine.org/institutional_review_board/hipaa_research/faq_research.html

Kendie, N. (2005). The Enterprise Data Model. Retrieved May 20, 2020 from https://tdan.com/the-enterprise-data-model/5205

Learn Computer Science. (2019). What is Database Management System? | What is DBMS? Retrieved May 6, 2020 from https://www.youtube.com/watch?v=lDpB9zF8LBw

Looker Data Sciences. (2020). Database Security. Retrieved June 2, 2020 from https://looker.com/definitions/database-security

OfficeSafe. (2016). HIPAA Training: What is required for HIPAA Compliance? Retrieved June 1, 2020 from https://www.youtube.com/watch?v=s9znUYvVO4A

Pijacek, R. (2019). Microsoft SQL Server Pros and Cons. Retrieved May 10, 2020 from https://learnsql.com/blog/microsoft-sql-server-pros-and-cons/

Shimoni, A. (2017). Linearly Scaling SQL - It's Not a Myth! Retrieved May 9, 2020 from https://sqream.com/linearly-scaling-sql-its-not-a-myth/

Sluiter, S. (2018). Activity 4 Part 1 - How to do a conceptual, logical and physical design for a database. Retrieved April 28, 2020 from https://www.youtube.com/watch?v=RzbH-oumqpo

Smith, A. (2011). Enterprise Data Modeling. EW Solutions: Data Management University. Retrieved May 24, 2020 from https://www.ewsolutions.com/enterprise-data-modeling/

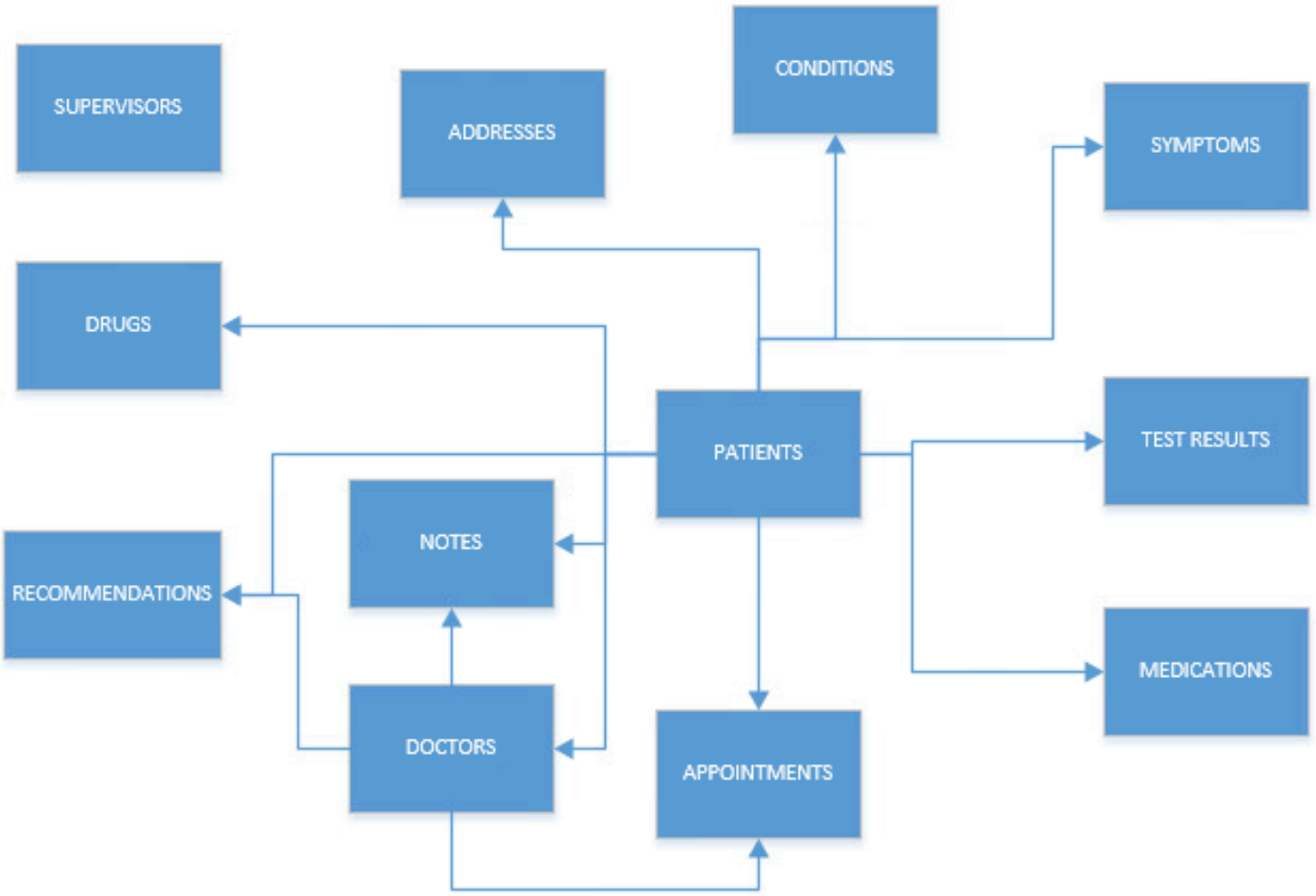
The Apache Software Foundation. (2016). Apache Cassandra. Retrieved May 10, 2020 from https://cassandra.apache.org/

UK Essays. (2017). Database Management: Law, Ethics and Security. Retrieved May 30, 2020 from https://www.ukessays.com/essays/computer-science/database-management-law-ethics-9552.php#:~:text=A%20combination%20of%20issues%20have,increase%20in%20circulation%20and%20excessive%2C

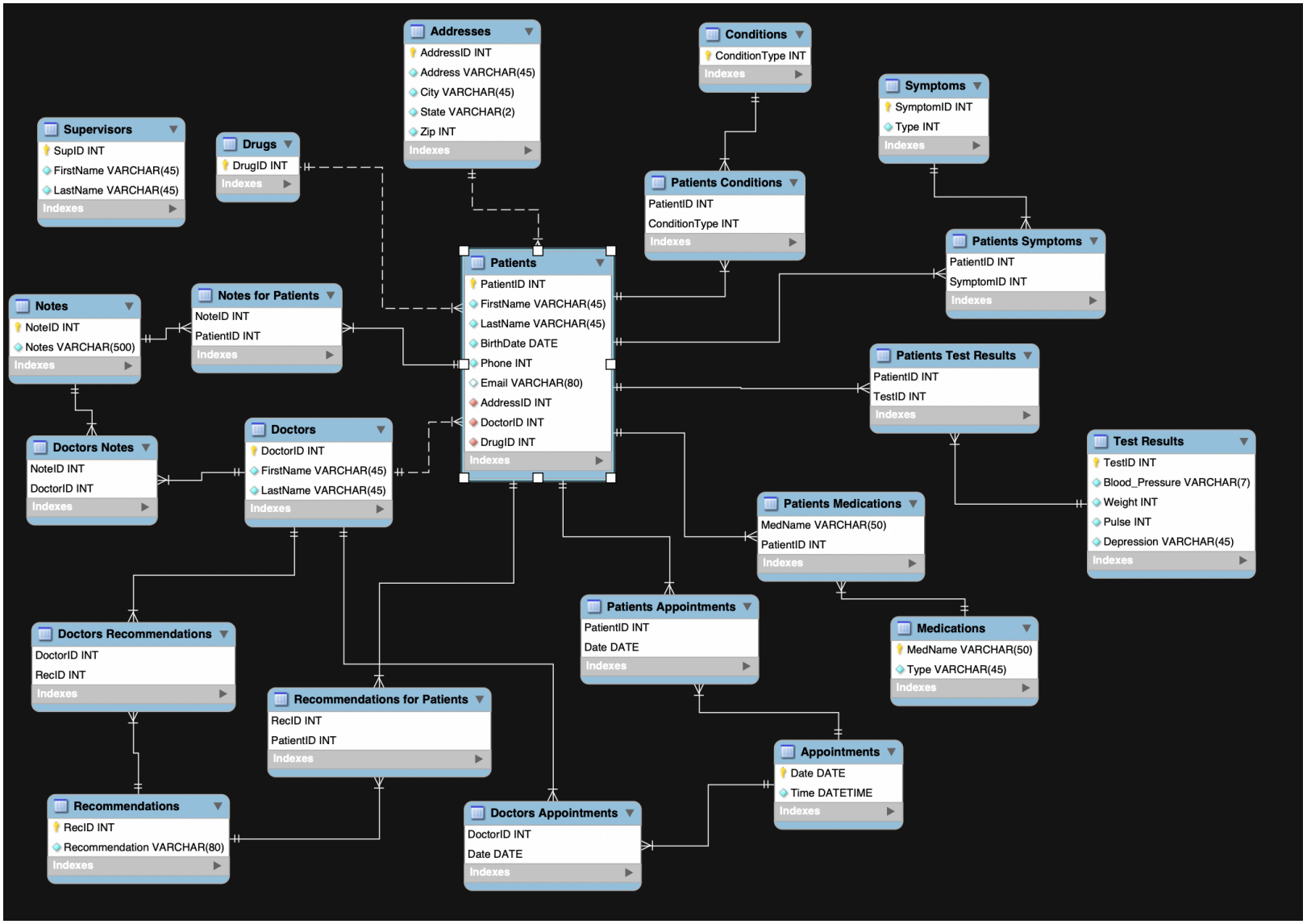
West, M. (2011). Developing High Quality Data Models. Data Models and Enterprise Architecture. Enterprise Data Model Retrieved May 22, 2020 from https://www.sciencedirect.com/topics/computer-science/enterprise-data-model

Appendices

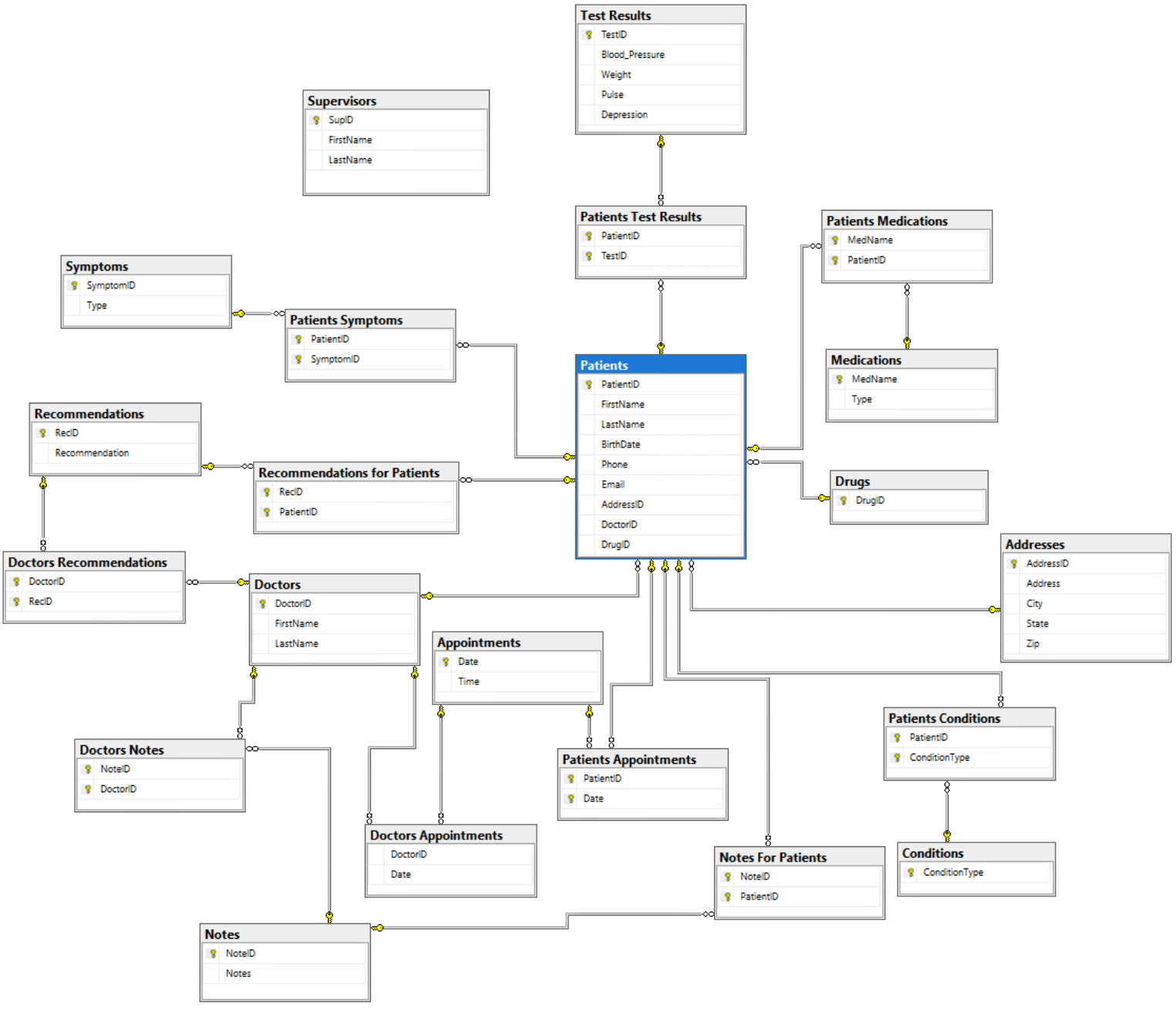
Appendix A: Conceptual Database Model



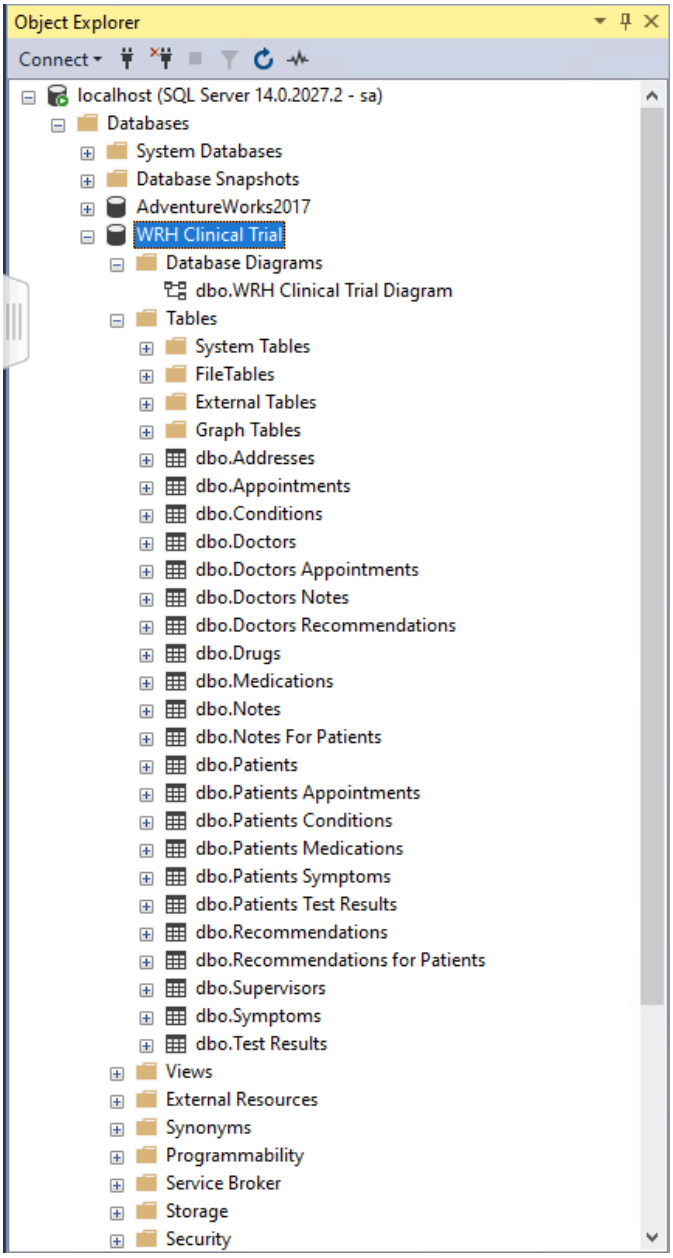
Appendix B: Logical Database Model



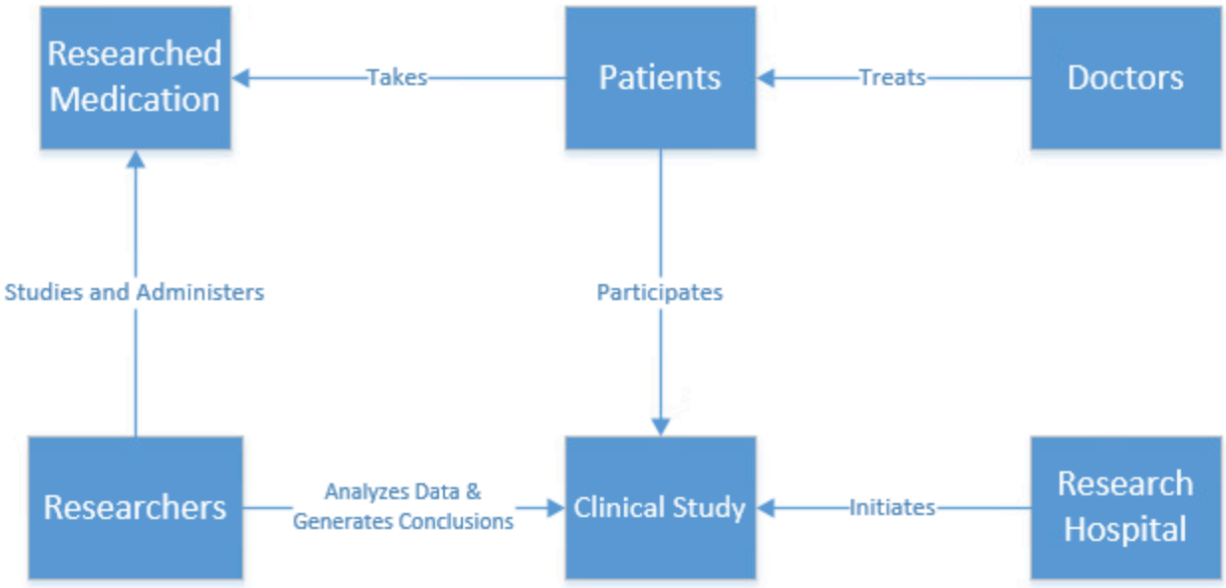
Appendix C: Physical Database Model



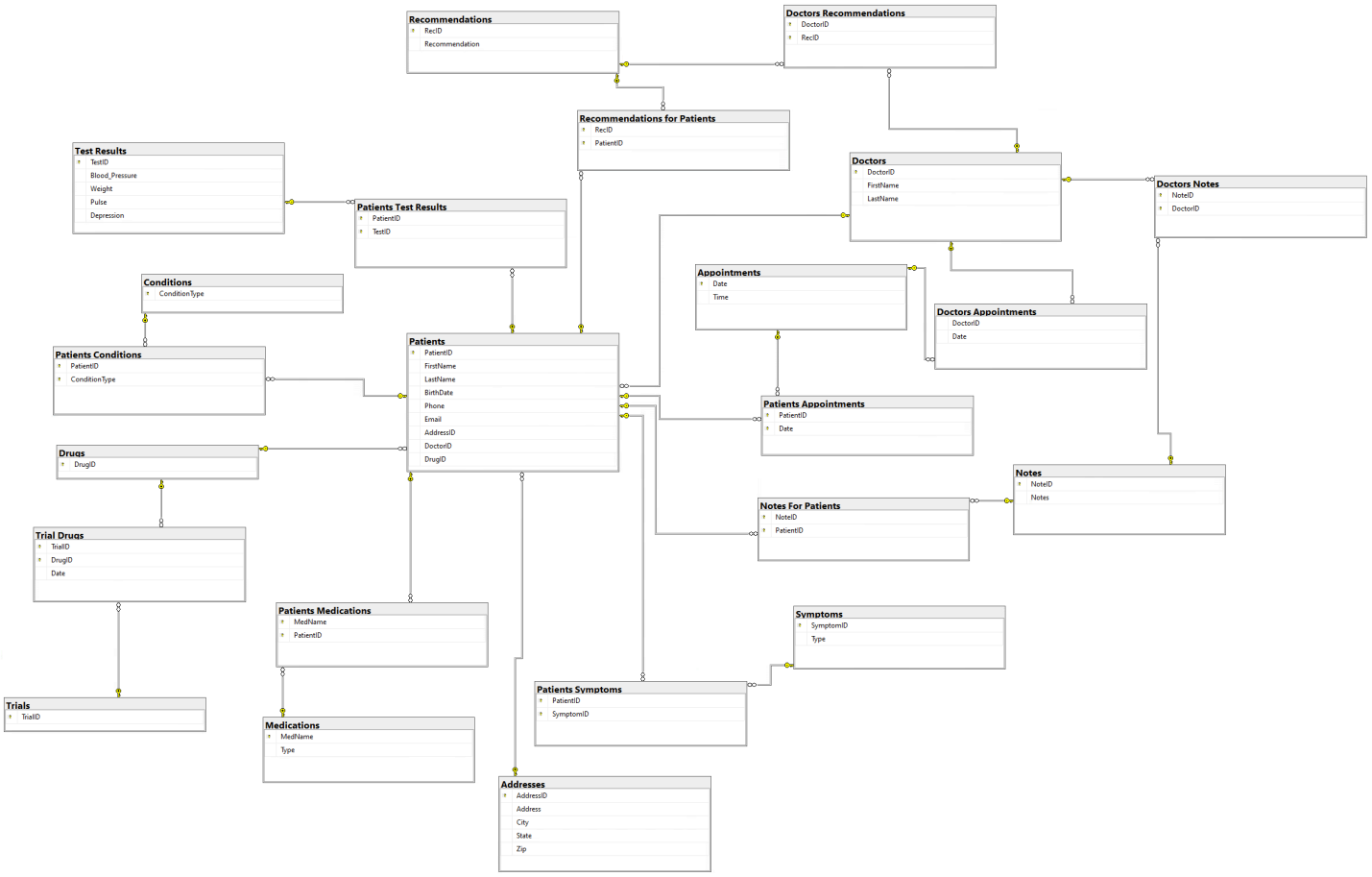
Appendix D: Physical Database Table List



Appendix E: Enterprise Subject Area Model (ESAM)



Appendix F: Enterprise Conceptual Model (ECM)



Appendix G: Enterprise Conceptual Entity Model (ECEM)