Based on week three I would say my design is a relational database design with one-to-many relationships. When there is a one-to-many relationship between two entities A and B, entity A occurrence is related to zero, one or more occurrences of entity B, and each occurrence of entity B is related to one occurrence of entity A. For instance, if we add homeowners to our home database, then there will be a one-to-many relationship between an owner and a home. At any time, a person can own zero, one, or more homes, and a home belongs to zero or one owner. (Johnston, 2014) Based on the scenario I chose and the table I provided in the previous discussion post, there are one-to-many relationships.

I have proposed seven tables for the database. If we look at the patient table, patient\_id is the primary key and it is related to the Adverse Event table where patient\_id is the foreign key. Similarly, tables like Home Health, Medication Administration Error, Inpatient Areas, and Outpatient Areas have the field patient\_id as a foreign key in those Tables whereas patient\_id is the primary key in the Patient table.  This shows one-to-many relationships.

**Reducing the risk of poor data integrity and eliminating duplicate data:**

Database data needs to be complete, organized, and accurate so that the data can be retrieved in the desired format when needed. Poor data integrity results in inaccurate and inconsistent data.

In duplicate data, data remains the same throughout the database. Every given customer order should have the same name, address, and phone number, typed the same way. Every order for a single title should have the same title, typed the same way. It will work when the duplicated data is consistent throughout the database but as the database grows it is not easy to maintain this type of consistency and this will be the time when a problem arises because many business-oriented database software are case sensitive, in that it considers upper- and lowercase letters to be different characters. In addition, no one is a perfect typist. A difference in capitalization, or even a single mistyped letter, will cause database software to consider two values to be distinct. (Harrington, 2016)

**How to reduce the risk of poor data integrity and eliminate duplicate data**

 Poor data integrity can be reduced through rigorous testing after we collect data or get any type of data to make sure that it is valid. There are many different types of data validation, but their goal is the same. There must be sufficient cleansing to improve the quality. Removing duplicate data and taking regular backups can be another way to reduce the risk of data integrity because these data can complicate a database and, if this data is sensitive, there is an increased risk of valuable data being stolen or leaked. Regular backups will ensure that we are protected against any ransomware attacks or data loss incidents. An audit trail can help us determine if any changes need to be made and if changes are made to either the data itself or the permissions surrounding the data, we can take action immediately to mitigate any potential damages that may arise. (Robinson, 2020)

According to Unleash Team, to reduce the risk of duplicate data a unique identifier can be used to check if a record already exists. Regularly cleaning the data to identify and remove duplicate records and using a data entry form with dropdown lists or autocomplete fields can help prevent duplicates by showing available options that match what the user is trying to enter. Lastly implementing data validation rules to check for any duplicates can reduce data duplication.

**References:**

How do you avoid having duplicate data in a database? – unleash community. Unleash. (2022, March 8).

[https://www.unleash.so/a/community/knowledge-management/how-do-you-avoid-having-duplicate-data-in-a-databaseLinks to an external site.](https://www.unleash.so/a/community/knowledge-management/how-do-you-avoid-having-duplicate-data-in-a-database)

Harrington, J. L. (2016). The Database Environment. Why good design matters (4th ed., pp. 49). essay, Morgan Kaufmann.

Johnston, T. (2014). One-to-many relationship. One-to-Many Relationship - an overview | ScienceDirect Topics.

[https://www.sciencedirect.com/topics/computer-science/one-to-many-relationshipLinks to an external site.](https://www.sciencedirect.com/topics/computer-science/one-to-many-relationship)

Robinson, P. (2020, June 17). What is data integrity and how do Reduce Data Integrity Risk. Lepide Blog: A Guide to IT Security, Compliance and IT Operations.

[https://www.lepide.com/blog/how-you-can-improve-data-integrity-in-3-easy-steps/Links to an external site.](https://www.lepide.com/blog/how-you-can-improve-data-integrity-in-3-easy-steps/)

**Response1:**

Hello Samuel

Thank you for the post. Both primary and foreign keys are crucial in relational databases because as per Byju’s primary key focuses on the uniqueness of the table and assures the value in the specific column is unique whereas foreign key is used to build a relationship between the two tables.

Data integrity is also a critically important aspect of systems that process or store data because it protects against data loss and data leaks. Maintaining data integrity over time and across formats is a continual process involving various processes, rules, and standards. Incomplete and inaccurate data can lead to bad decisions costing significant time and effort. If there is a loss of sensitive data, it might end up in criminal hands resulting in negative impacts. There are a wide variety of threats to data integrity. While most people imagine malicious hackers as the main threat, the majority of root causes are internal and unintentional, such as errors in data collection, inconsistencies across formats, and human error. According to Qlick, educating leaders, establishing a robust data governance framework, and investing in the right tools can be some measures in reducing the risk of poor data integrity.

**References:**

BYJU’S. (2023, March 29). Difference between primary key and foreign key. BYJUS.

[https://byjus.com/gate/difference-between-primary-key-and-foreign-key/#:~:text=1-,A%20primary%20key%20generally%20focuses%20on%20the%20uniqueness%20of%20the,relationship%20between%20the%20two%20tables.&text=The%20table%20allows%20only%20one%20primary%20keyLinks to an external site.](https://byjus.com/gate/difference-between-primary-key-and-foreign-key/#:~:text=1-,A%20primary%20key%20generally%20focuses%20on%20the%20uniqueness%20of%20the,relationship%20between%20the%20two%20tables.&text=The%20table%20allows%20only%20one%20primary%20key).

What is Data Integrity? why you need it & best practices. Qlik. (2023).

[https://www.qlik.com/us/data-management/data-integrityLinks to an external site.](https://www.qlik.com/us/data-management/data-integrity)

**Response2:**

Hello Bernice

Thank you for the post. Data organization is a crucial part of data science that helps make data easy to understand, analyze,  visualize, and protect from unauthorized access. Database administrators categorize and store data in a relational database that can then be queried and filtered to extract information for reports. Relational databases can be extended, and they don’t rely on physical organization. Following the original database creation, a new data category can be added without having to modify the existing applications via a relational database.  (Lutkevich & Biscobing, 2021)

I agree with you that, it is crucial to prioritize a well-defined relationship structure when improving the design because according to Fortinet.com, organizations are becoming more reliant on data integration and the ability to accurately interpret information to predict consumer behavior, assess market activity, and mitigate potential data security risk.  For better healthcare data management, consistent datatypes, and enforcing limitations are important in ensuring data integrity and enhancing the outcomes.

**References:**

Lutkevich, B., & Biscobing, J. (2021, June 24). What is a relational database?. Data Management.

[https://www.techtarget.com/searchdatamanagement/definition/relational-databaseLinks to an external site.](https://www.techtarget.com/searchdatamanagement/definition/relational-database)

What is Data Integrity? why is it important?. Fortinet. (2021).

[https://www.fortinet.com/resources/cyberglossary/data-integrityLinks to an external site.](https://www.fortinet.com/resources/cyberglossary/data-integrity)