# Expanded Review of EHR and Research Database Data Quality to Identify Common Errors

Donia Shawn, Dr. Iris Zachary, Dr. Uzma Khan, Dr. Magda Esebua

Health Management and Informatics
School of Medicine
University of Missouri Health

School of Medicine
University of Missouri Health

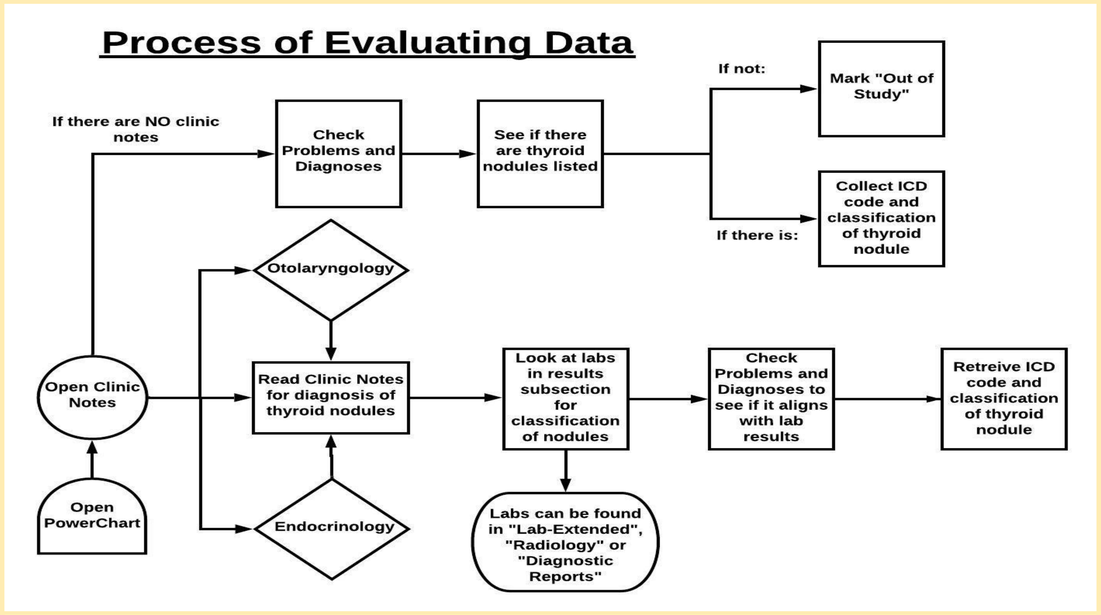
Health
University of Missouri

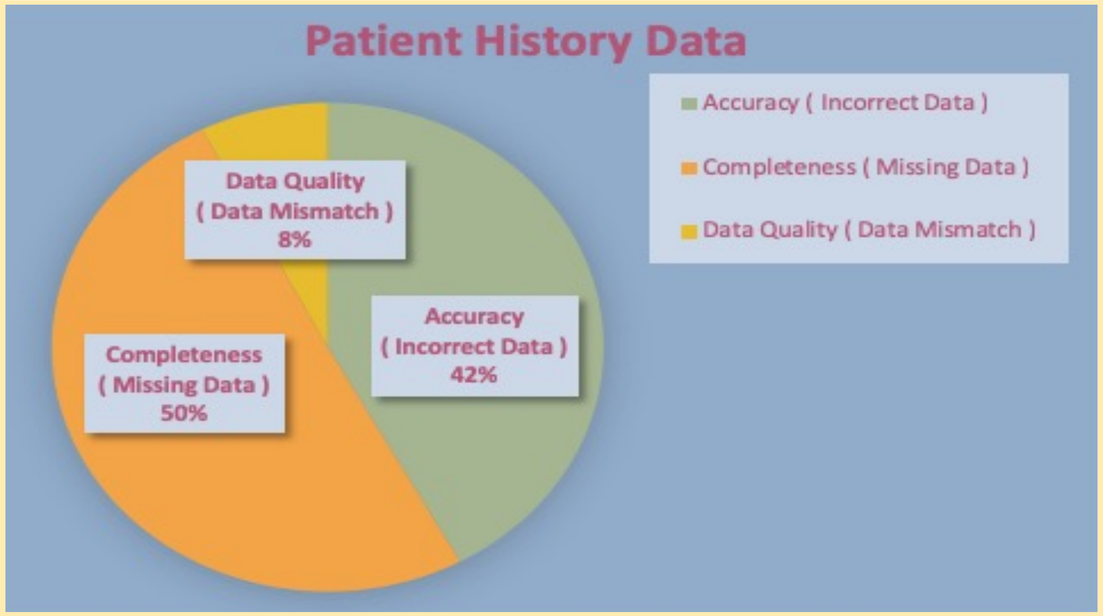
Honors College
University of Missouri

**Objective:**
To analyze MU-TNED ( MU – Thyroid Nodule Electronic Database) for accuracy, completeness, and data quality for the variables: age, gender, nature of thyroid nodules and BMI.

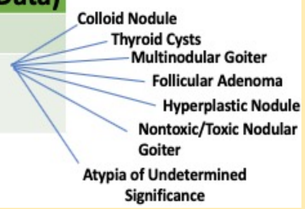
**Introduction:**
There is a need for quality in clinical databases that include complete, accurate, and validated data to perform population-based health assessments. MU TNED was designed to focus on patients with thyroid nodules to enable review of epidemiological and clinical data, focusing on quality control and quality improvement, and assess long-term follow-up of patients. The goal of this study is to evaluate MU-TNED for age, gender, nature of thyroid nodules and BMI variables for completeness and accuracy.

**Methodology:**
In this study, an electronic health record (EHR), PowerChart is used to retrieve patient data from the University of Missouri Health Care (UMHC) over an eleven-year time period, 2008-2019. MU-TNED is a REDCap research electronic data capture repository that enables automatic data transfer from PowerChart to MUTNED. MUTNED includes 9,907 patient records.




| | Data Quality (Data Mismatch) | Completeness (Missing Data) | Accuracy (Incorrect Data) |
|---|---|---|---|
| BMI | 17 | 114 | 0 |
| Classification of Thyroid Nodules | 0 | 35 | 95 |




**Results:**
As the initial step, we analyzed accuracy (incorrect data), completeness (missing data), data quality (data mismatch), from the set of 9,907 patients. We found that:

- Age was complete for 100% of patients
- Gender data was complete for 100% of patients

However, data for BMI was found to be incomplete or mismatch. In addition, manual review for nature of thyroid nodules resulted in improvement of the specificity for the existing data categories, benign or malignant.

**Conclusion:**
It is important to review data collections for data quality including, accuracy, completeness and validation to ensure consistent use of data. Good quality and valid data is necessary to support studies that can provide accurate population-based assessments. Such studies are needed to enhance health records and improve patient care and wellness.

**Benefits:**
Improving the healthcare databases, not only will benefit doctors when accessing history of a patient, but it will allow patients to view their health records with ease. By evaluating MU-TNED database for completeness and accuracy, it will allow us to work on the next step the study which is to find if there is an association of BMI and TSH with the nature of thyroid nodules, benign or malignant and correlation among genders.