A Study on the Role of Medical Record Data in Supporting Clinical Research at City Based Hospital Kolkata

Sohini Pal
The Neotia University

OBJECTIVE

i) To explore the significance of medical record data in clinical research.
ii) To assess the role of medical records in clinical research and decision-making.
iii) To examine the challenges in using medical records for research purposes.

RESEARCH METHODOLOGY

The methodology adopted for this project involves a combination of qualitative research, secondary data analysis, and literature review to understand and evaluate the role of medical record data in clinical research. The steps followed are detailed below:

1. Research Design
- Type of Study: Descriptive and analytical.
- Approach: Qualitative and secondary data-based.

Scope: Focused on exploring the utilization of medical record data (both electronic and manual) in clinical research and identifying associated benefits, challenges, and best practices.

2. Data Collection
- Source of Data

The data for this study is collected from the Medical Records Department of the selected hospital. These records include:

Document Analysis:
Analysis of existing medical record templates, clinical trial reports, and guidelines on the use of electronic health records (EHRs) in research.

Observational Study:
- What are the clinical researches done in this hospital?
- The results found in "NIKSHAY" portal
- The results found in the official portal of "Clinical Establishment Act"

3. Data Analysis Plans
The data analysis plan outlines the methods and strategies used to interpret and extract insights from the data collected during the study. Since this project involves qualitative (from interviews, document reviews) data, a mixed-methods analysis approach has been adopted.

➢ Qualitative Data Analysis

Qualitative data from interviews, open-ended survey responses, and document reviews will be analyzed using thematic analysis.

a. Transcription and Organization
- Interview recordings and written responses will be transcribed and compiled.
- Responses will be categorized under broad themes such as:
  i) Data accessibility
  ii) Perceived benefits
  iii) Ethical concerns
  iv) Technical challenges

b. Thematic Coding
- Common patterns and keywords will be highlighted manually.
- Themes will be grouped and interpreted to identify:
  i) Recurring challenges
  ii) Best practices
  iii) Suggestions for improvement

c. Analysis of Secondary Data
Secondary data from EMRs, policy documents, and literature reviews will be analyzed by:
- Content Analysis: Reviewing hospital protocols, data standards, and ethical guidelines to understand organizational practices.

d. Presentation of Results
- Findings will be presented in tables, graphs, and thematic summaries.

- Each research objective will have a corresponding analysis segment in the results chapter.
- Key insights will be aligned with the project's aims to ensure coherence.

➢ Limitations of Data Analysis

- Access to real patient records may be restricted, affecting case-based validation.

❖ Ethical Considerations:

1. Patient Confidentiality and Privacy:-

i) All medical records used for clinical research must be handled with strict confidentiality.

ii) Personal identifiers (e.g., name, address, phone number) should be removed or anonymized to protect patient identity.

iii) Researchers must comply with data protection laws such as the *Health Insurance Portability and Accountability Act (HIPAA)or applicable national privacy regulations.

2. Informed Consent

i) Whenever possible, informed consent should be obtained from patients whose data is being used.

ii) In some cases, waivers of consent may be permitted (e.g., for retrospective data studies), but only with ethical committee approval.

3. Data Security and Storage

i) Medical record data must be stored securely, using encrypted databases or password-protected systems.

ii) Access to the data should be restricted to authorized personnel only.

4.Beneficence and Non-Maleficence

i) Research should be designed to benefit public health and improve medical knowledge while avoiding harm to individuals or communities whose data is used.

ii) Misuse or misinterpretation of data should be avoided as it can lead to stigmatization or discrimination.

5. Transparency and Accountability

i) Researchers must clearly disclose how the data will be used, stored, and shared.

ii) Accountability mechanisms should be in place to address any breaches of ethical or legal obligations.

DATA ANALYSIS AND INTERPRETATION

PRESENTATION OF DATA:

i) "Dengue" cases over the years:

Year	Dengue Cases
2021	44585
2022	193245
2023	233251
2024	31464

Table: 01

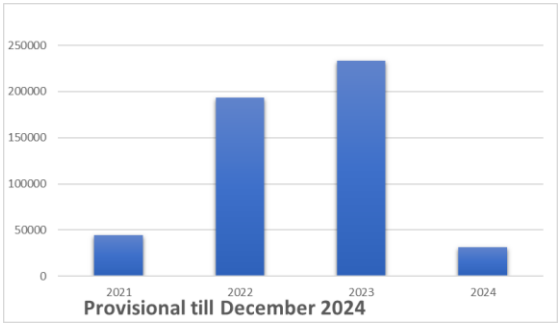

Figure: 02

The bar chart illustrates the number of Dengue cases reported annually from 2021 to 2024, with 2024 data being provisional.

Interpretation:

- Rising Trend (2021–2023): A significant upward trend in dengue incidence was seen from 2021 to 2023.
- Sharp Decline (2024): A notable decline in 2024 is seen, but being provisional, this figure should be interpreted cautiously.
- Public Health Implications:
  - The 2022–2023 surge likely stressed healthcare resources.
  - The drop in 2024 is encouraging but warrants verification and continuous surveillance.

ii) Trend analysis of TUBERCULOSIS incidence:

Year	TB Notification
2022	2386418
2023	2508061
2024	2618613
2025	1535732

Table: 02

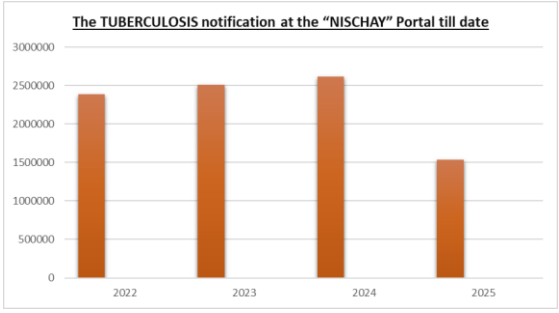

Figure: 03

NI-KSHAY- (Ni=End, Kshay=TB) is the web enabled patient management system for TB control under the National Tuberculosis Elimination Programme (NTEP).

Interpretation:

❖ Steady Increase (2022–2024):

- There was a gradual rise in TB notifications from 2022 to 2024, indicating improved case detection or a possible increase in incidence or reporting efficiency.

❖ Sharp Decline (2025):

- The sudden drop in 2025 suggests one or more of the following:
  - Actual reduction in TB incidence due to successful intervention programs.
  - Under-reporting or incomplete data (since 2025 might be a partial year or ongoing).
  - Possible technical or reporting delays in the NISCHAY portal data for the year 2025.

iii) Malaria Cases and Deaths:

Year	Malaria Cases (Green)	Malaria Deaths (Blue)
2022	~580	~370
2023	~460	~290
2024	~640	~390
2025	~230	~110

Table:

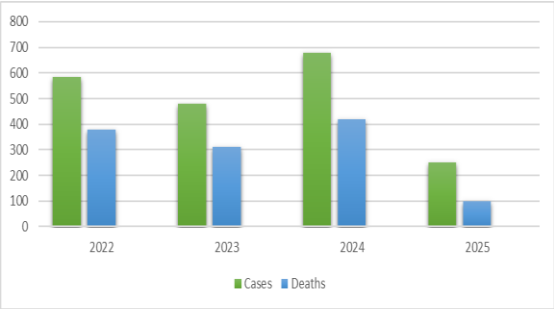

Figure:04

Interpretation:

- 2022–2023:
  - Both cases and deaths decreased slightly from 2022 to 2023.
  - This could indicate a temporary improvement in prevention, diagnosis, or treatment efforts.
- 2024:
  - A spike in both malaria cases and deaths occurred, reaching the highest level in the 4-year span.
  - Possible reasons:
    - Seasonal outbreak or failure in vector control.
    - Gaps in healthcare service delivery or drug resistance.
- 2025:
  - A sharp decline in both malaria cases and deaths.
  - This significant drop may suggest:
    - Effective implementation of malaria control programs.
    - Improved public health awareness.
    - Or it could indicate incomplete data for 2025 if the year is still ongoing.

DISCUSSION AND RECOMMENDATION

∝ Discussion:

The use of medical record data in clinical research has transformed the way studies are conducted, offering both efficiency and depth in understanding patient outcomes. This project highlights several key points:

i) Data Richness and Availability: Electronic Medical Records (EMRs) and paper-based records hold a wealth of longitudinal patient information including diagnoses, treatments, outcomes, and lab reports. This real-world data provides a foundation for observational studies, cohort analyses, and retrospective research.

ii) Improved Patient Selection: EMR systems help researchers quickly identify eligible participants for clinical trials based on demographic, clinical, or genetic criteria, reducing recruitment time and cost.

iii) Monitoring and Outcome Assessment: Continuous tracking of patient data enables researchers to monitor treatment progress and side effects, and evaluate long-term outcomes more accurately than self-reported methods.

iv) Challenges in Data Accuracy and Completeness: One significant limitation is the variability in documentation practices, incomplete records, and errors in data entry. These inconsistencies can lead to bias or misinterpretation of results.

v) Data Privacy and Ethical Concerns: The use of medical record data must be balanced with strict adherence to patient confidentiality and data protection regulations (e.g., HIPAA, GDPR). Anonymization and secure storage are critical.

vi) Interoperability and Standardization Issues: Many healthcare institutionsuse different

systems, which limits the ability to aggregate data across sources. Lack of standard formats (e.g., SNOMED CT, LOINC) can hinder large-scale or multicenter studies.

vii) Potential for Bias: Since medical records are generated during routine care and not for research, the data may be influenced by socioeconomic or institutional factors, leading to possible bias.

∝ Recommendation:

Based on the findings and challenges identified, the following recommendations are proposed:

1.Ensure Ethical Compliance: Enforce robust ethical practices including informed consent (when necessary), data anonymization, and secure access protocols to protect patient privacy.

2. Promote Training and Awareness: Train healthcare professionals on the importance of accurate documentation and the research potential of clinical data.

3.Encourage Integration of AI and NLP: Invest in advanced tools like natural language processing (NLP) and artificial intelligence to extract useful insights from unstructured data such as clinical notes.

4.Foster Collaboration: Encourage collaboration between hospitals, research institutions, and regulatory bodies to create unified health data repositories for multicenter research.

5. Support Policy Development: Advocate for clear national and institutional policies regarding the secondary use of medical data for research to avoid legal and ethical ambiguities.

CONCLUSION

Medical record data is an invaluable resource for clinical research. It enables a shift from traditional, resource-intensive trials to data-driven, efficient, and real-world research paradigms. While challenges remain in data quality, standardization, and ethics, continued investment in “technology, infrastructure, and policy frameworks” can unlock the full potential of medical record data. When responsibly and intelligently used, it will accelerate innovation, improve patient care, and strengthen evidence-based medical practice.