ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 02 ISSUE: 02

RECEIVED: APRIL 21 REVISED: MAY 10 ACCEPTED: MAY 026 PUBLISHED: JUNE 19

Historical Data as the Foundation of Clinical Foresight

Madhu Sathiri
Independent Researcher
sathirimac@gmail.com

Abstract

Clinical Decision Support systems based on historical patient data represent a model for investigation of a category of Clinical Decision Support (CDS) applications that have recently gained traction — consumption of historical data to support clinical decision making. An overview of the category is provided and addressed based on five research questions: What are the properties of historical data that provide insights? How are these insights used in Analytical Methods for Clinical Decision Support based on Historical Data? How is the architecture of the Clinical Decision Support System represented? How is clinical validation and evaluation achieved? How is the presentation and user interaction with predictions controlled? Despite being a nascent area, numerous published works are available on the topic, and clinical validation efforts are in various stages of maturity.

Historical data have become increasingly more available, are generally regarded as safe, and, unlike predictive models constructed on Future Data, do not raise questions regarding biases introduced in the construction of supervised learning models. However, while safety and bias concerns are reduced, historical data are prone to their own logical shortcomings. Therefore, users must exercise caution when interpreting the insights derived from Historical Data Analytics. Properly defined and presented, such insights do not impose a decision making burden on clinicians but instead ease the decision making process. The existence of a predictive model that maps the Clinical Data relevant to a Clinical Decision to the Clinical Decision itself is often regarded as a requirement for CDS to be genuinely useful for everyday clinical practice. Such a statement suffers from circular reasoning. In summary, because no method of construction of predictive models can claim to be free from bias, the pannational nature of probability suggests that, provided sufficient Causally Uncorrelated Data points are available notwithstanding time, space and biological differences, the possibility of identifying and using a Supervised Learning Model that accurately predicts the Decision of a Clinician can not be excluded.

Keywords: Clinical Decision Support Systems, Historical Patient Data, Clinical Decision Analytics, Healthcare Data Science, Analytical Methods For CDS, Clinical Insight Extraction, CDS System Architecture, Clinical Validation Frameworks, Evaluation Of Decision Support, User Interaction Design, Prediction Presentation Control, Bias In Clinical Data, Safety Of Historical Data, Logical Limitations Of Retrospective Analytics, Clinician Decision Support, Supervised Learning In Healthcare, Clinical Data Modeling, Evidence-Based Clinical Practice, Human-Centered CDS Design, Probabilistic Clinical Reasoning.

1. Introduction

A clinical decision support (CDS) system generally improves patient care by giving healthcare practitioners the knowledge needed to improve care quality, reduce errors, and increase efficiency in patient management. It incorporates computing technologies to support caring decisions through predictive or descriptive analyses on patient data, using data from various sources, including historical data. Historical data helps answer natural clinical questions on how past patient cohorts behaved. These inquiries, pertaining both to diverse clinical results and to decision-making processes in clinical practice, constitute an emerging application of historical data in CDS. Although they do not focus on the behavior of the underlying data-generating process, such analyses can still employ all historical available data and are not model-specific.

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 02 ISSUE: 02

RECEIVED: APRIL 21 REVISED: MAY 18 ACCEPTED: MAY 026 PUBLISHED: JUNE 19

Verifying whether a subgroup in a patient cohort is becoming particularly more prone to a certain clinical event and investigating whether doctors follow the knowledge compiled in the past are two examples of such inquiries. The results indicate that historical data are new, valuable, and often underused knowledge that can be extracted from routine clinical databases for constructing clinical answers and their predicted probability tables, and that answers to natural clinical questions are naturally interpretable. Though these two types of analyses are not yet incorporated into working CDS, the advancements presented here make their integration feasible through a transparent and straightforward process.

1.1. Purpose and Scope of the Study

The research questions motivating this investigation are two-fold: How can historical patient information, having existed in the past, similarly be utilized to guide ongoing current patient clinical management decisions? In particular, how can historical data be effectively leveraged in Clinical Decision Support Systems (CDSS) to inform better decision-making? Beyond providing an overview of progress toward answering these questions, the analysis also outlines the currently accepted but narrow view in the literature of the application of historical data solely as an available source for training and validation of CDSS solutions. Instead, the potential of historical data to directly provide incisive clinical guidance is fully considered in addition to ongoing model-training applications. Mirroring the broader scope of the two research questions, particular attention is devoted to issues that emerge during and influence the entire data lifecycle. The results clarifying the role and nature of historical data within a CDSS framework are of immediate practical relevance since such systems are now being considered for support in real-world clinical decisions.

A CDSS automatically generates Case-Based Reasoning (CBR) guidance for clinicians and patients by analyzing the records of patients treated previously in a clinical cohort. The clinical intervention supported in the specific pilot is the initiation of long-term anticoagulation therapy for patients with an AF episode associated with a cardiac hospital admission, but the underlying approach is sufficiently general to extend to other interventions. In this context, historical data refers to the records of all the patients in the cohort who are not currently in clinical management for the intervention. CDSS applications capable of directly providing clinical decision support based on historical data represent a new category of such systems that are beginning to be used in real-world clinical settings.

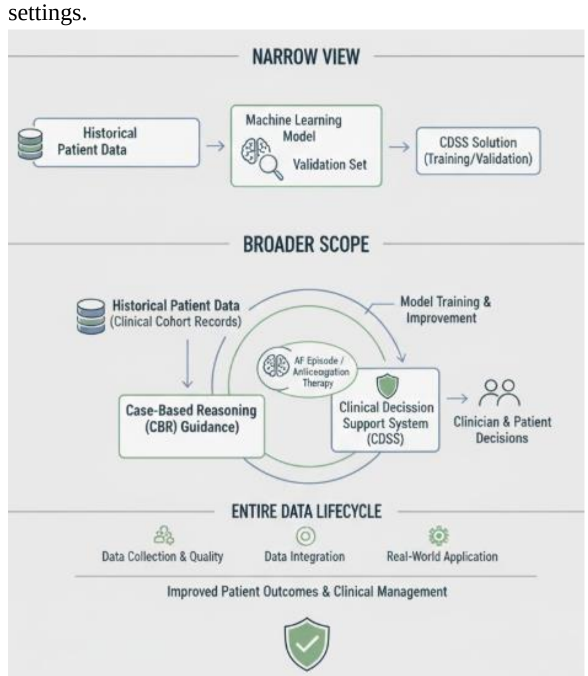


Fig 1: Beyond Model Training: A Case-Based Reasoning Framework for Leveraging Historical Clinical Data in Real-World Clinical Decision Support Systems

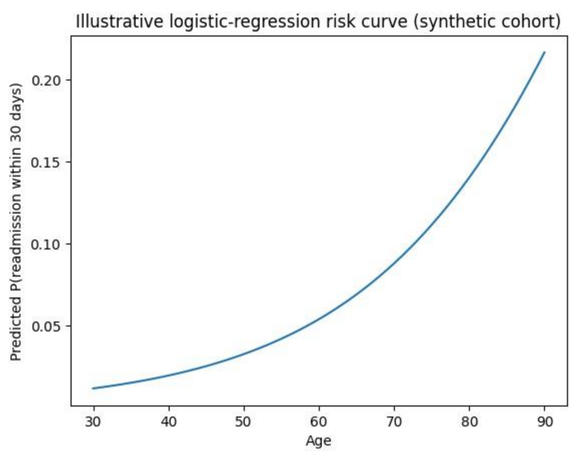
2. Background and Context

Positioning the study within literature on clinical decision-making systems, it notes FDAs emerging interest in historical data analyses; directs attention toward the clinical cogency of historical data-based decision support functions. Clinical Decision Support Systems (CDSSs) proposed by historical data exploits data from past patients discharged from a healthcare system, generating predictive information

ISSN: 3067-4166
AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS
VOLUME: 02 ISSUE: 02
RECEIVED: APRIL 21 REVISED: MAY 10 ACCEPTED: MAY 026 PUBLISHED: JUNE 19

and insights. Such systems have the potential to assist with the approval process, with data analysis from past patients employed to bolster the rationale for anticipated populations. Consequently, the relevant data set should ideally be representative of the population likely to be treated. To ensure the proposed CDSS is robust and promotes enhanced decision-making by healthcare professionals, it is validated in an actual operating environment.

The potential for bias mitigation through descriptive analytics is discussed by evaluating a quantity of past patients treated with a low-frequency drug that indicated a real-world bias. Analysis of time-series trends can provide decision support by presenting clinical information. Predictive analysis can assist by supplying missing information or flagging characteristics that may have been overlooked. For predictive modeling, important variables are selected for model construction, with the contributions of the proposed function assessed by internal validation. The approach is also applied to further decision support analysis.



Equation 1: Core probability tables from historical cohorts

1.1 Event probability in a cohort (single group)

Suppose you have a historical cohort of size N and an event E (e.g., readmission, sepsis, CHD onset).

- Let n_E = number of patients who experienced E.
- Then the empirical probability (historical risk) is:

$$\hat{P}(E) = \frac{n_E}{N}$$

Step-by-step

1. Count cohort size N
2. Count event occurrences n_E
3. Divide: n_E/N

1.2 Conditional probability table for subgroups

For a subgroup feature A (e.g., diabetes=1):

- Let N_A = number of patients with A
- Let $n_{E\cap A}$ = number with both E and A

$$\hat{P}(E \mid A) = \frac{n_{E\cap A}}{N_A}$$

2.1. Historical Data Insights and Their Implications

Patterns, biases, and their potential for bias mitigation shape historical data value. Conformity biases may harm prediction accuracy; however, descriptive analytics can reveal counteracting factors. Although capturing and accounting for clinical context may be difficult, predictive methods could incorporate associations with individual patient activities or key events. Complete capture of historical patient data might further mitigate treatment prediction errors. Anomaly detection can identify cases where treatments deviate from historical and probable practice. Specific user queries could also be supported by titrating clinical data against changing consensus to determine context-aware expressions of user interest.

The defined decision support question addresses the risk of chronic heart disease (CHD) among patients undergoing renal replacement therapy. After displaying the predictive model results, the driver analysis identifies main factors behind model predictions. With most defined factors collecting historical information, changes to their prior distributions become relevant. Concentrating on those

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 02 ISSUE: 02

RECEIVED: APRIL 21 REVISED: MAY 10 ACCEPTED: MAY 026 PUBLISHED: JUNE 19

dimensions stimulates an examination of the interaction between five drivers, with abnormal combinations highlighted prior to examination of underlying root causes or other relations among the variable states. All visualizations dynamically adapt to the concerned subset.

3. Data Sources and Quality

A variety of data sources were used to support the analyses. Primary datasets included historical clinical records, medication prescription reimbursement claims, disease diagnosis rates, voting records for elected regional public health authorities, morbidity and mortality statistics from the World Health Organization, and national-level quality-of-life data from the World Bank. Sources were selected according to three main drivers. The first concerned scope, requiring that the data support the questions of interest; the second focused on provenance quality to ensure that the dataset was credible for the question being addressed; and the third centered on data governance systems. Data quality is crucial when relying on historical information for decision support. Five fundamental elements of data quality—completeness, consistency, timeliness, standardization, and error rates—should be addressed for each data source.

After the considered aspects of data governance, the key dimensions of data integrity—namely, reliability and representativeness—have to be assessed. Reliability is addressed by examining the data provenance information as well as any validation level applied by the responsible organization and with regard to the historical period under consideration. Representativeness is evaluated by examining time frames of the data compared with the maturation periods for detecting a possible clinical effect in the data under consideration.

Patient_ID	Age	Creatinine (mg/dL)	Prior_Admissions	AF_Episode	Diabetes	Outcome: Readmission_30d
001	63	1.30	2	0	1	1
002	45	0.90	0	0	0	0
003	78	1.80	3	1	1	1
004	52	1.10	1	0	0	0
005	69	1.55	2	1	0	1

3.1. Assessment of Data Integrity and Reliability

The reliability and accuracy of historical data is fundamental for producing insights with clinical decision support systems. Towards this goal, the completeness, consistency, timeliness, standardization, and error rate of historical data were assessed. Incomplete data carries a risk of generating biased insights. It can also hinder decision support functionality like predictive modeling. Consistent data serves as reliable sources for descriptive analyses, trend detection, and other explorations to inform decision support. Moreover, timely data helps identify emerging patterns that are crucial for decision support and interventions.

Reliability of historical data lies in transparent provenance and governance. Consequently, reliable and accurate historical data enhances the surface-level usability of clinical decision support systems for descriptive analytics, reduces maintenance and configuration overheads for predictive modeling, and cultivates clinical trust to improve usage. The exploration confirmed these perspectives. The historical data is sufficiently complete. Furthermore, the intervals and intervals of the different data types are converging, suggesting increasing consistency and timeliness. Overall, historical data is reliable, paving the way for generating and utilizing such insights.

ISSN: 3067-4166
AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS
VOLUME: 02 ISSUE: 02
RECEIVED: APRIL 21 REVISED: MAY 10 ACCEPTED: MAY 026 PUBLISHED: JUNE 19

Fig 2: Validating the Foundations of Clinical Guidance: A Multidimensional Quality Assessment of Historical Data for Trust-Based Decision Support Systems

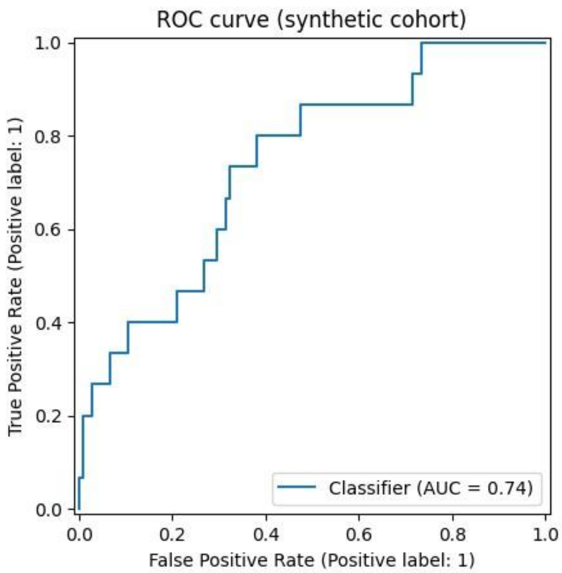
4. Methodologies for Historical Data Utilization

Several methodologies can be employed to leverage historical patient data. For example, descriptive statistics can be used to present general characteristics of the data, assess how specific features have evolved over time, and identify notable trends with the help of various visualization techniques. Predictive analytics can estimate the likelihood of a specific incident or the expected time until such an incident occurs.
These models are generally trained using a subset of historical data and validated on another subset. To enable practical implementation, proper feature engineering is essential. Analytical procedures targeting a defined patient cohort or using time-oriented analysis typically combine the employed techniques into a cohesive framework. Validation of these approaches can be based on the clinical literature, prior studies, expert opinion, or a dedicated pilot study investigating the relevance of the presented information and the impact of the suggested management approach.

4.1. Descriptive Analytics

Preliminary assessments leveraged descriptive analytics to better understand historical data's properties and share insights with stakeholders. Summary statistics and timelines provided global overviews, while data visualizations revealed important trends and patterns. Analyzing such aspects is crucial for trustworthy and clinically relevant decision support. Reliable information on historical data sources and distribution facilitates appropriate use and enhances clinical adoption. Patterns and systematic errors indicate where bias might arise, paving the way for mitigation.
For examples, analyses of patient cohorts and conditions gave an overview of the data's clinical context. Time series and distributions of monitored parameters highlighted distinctive trends and unusual distributions or values. Such explorations foster instinctive and informed clinical engagement as they reveal the roots of alerts, options, and recommendations. The projected use demonstrated these principles with analyses of historical data related to heart failure, hospital-acquired infections, and sepsis. Each area revealed important trends and properties that established the groundwork for reliable, clinically relevant, and clinically engaged decision support.

ISSN: 3067-4166
AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS
VOLUME: 02 ISSUE: 02
RECEIVED: APRIL 21 REVISED: MAY 10 ACCEPTED: MAY 026 PUBLISHED: JUNE 19



Equation 2: Descriptive analytics equations

2.1 Mean and variance (for a lab value x, e.g., creatinine)

Given values $x_1, x_2, \ldots, x_N$:

Mean

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$$

Sample variance

$$s^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2$$

Step-by-step (variance)

1. Compute $\bar{x}$
2. For each value, compute deviation $d_i = x_i - \bar{x}$
3. Square deviations d_i^2
4. Sum squares $\sum d_i^2$
5. Divide by $N - 1$

2.2 Moving average for trends (time-series)

For a time series y_t and window size w:

$$\mathrm{MA}_t = \frac{1}{w}\sum_{j=0}^{w-1} y_{t-j}$$

4.2. Predictive Modeling

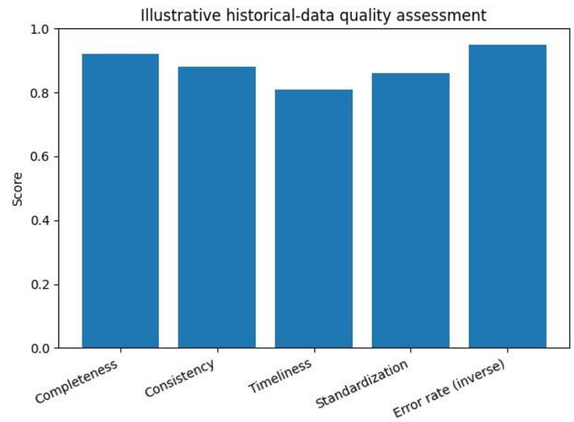
Ordinary regression models like linear, polynomial, and logistic regression can be fitted directly on the dataset. For example, logistic regression can be used when the target variable is binary and implementing binary logistic can produce probabilities representative of different clinical outcomes. Multiple regression can allow prediction of any continuous-scale variable like a stay duration.

A part of the dataset can be used for training the regression model (80% is common) and the remaining can be used for validation. Various performance measures can be used such as MSE, Mean Absolute Error (MAE), Root mean squared error (RMSE), Different regression plots and accuracy can be checked. Such testing can be done using cross-validation techniques like K-Fold, Stratified, Leave-p-out or others. KNN (K Nearest Neighbor) model can also be trained on the dataset for classification based prediction, just checking the accuracy. When fitted by using the right variables KNN is able to predict well. Also, decisions based on majority voting such as Bagging can also be performed. Such trained predictive models can be embedded with the data pipelines and exposed to the application layer over HTTP. They can be triggered from the web service when providing input features.

Machine Learning & Deep Learning Models can also be applied as classification or prediction engines on the data. Although Advanced ML & DL Models like Random Forest, GBDT, Support Vector Classifiers, etc. can also be used, Linear classifiers like Logistic Regression and LDA may perform just as well and have the advantage of being easy to interpret. These Models can be tested on a part of the data (say 20%) while K-fold cross-validation can be used for hyper parameter tuning. The Hyper parameters thus defined can simply be applied on the whole data set

ISSN: 3067-4166
AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS
VOLUME: 02 ISSUE: 02
RECEIVED: APRIL 21 REVISED: MAY 10 ACCEPTED: MAY 026 PUBLISHED: JUNE 19

and deployed on any cloud infrastructure. The Logical LDA Classifier can perform sentiment classification. These models can also be applied on the longitudinal group data separately or collectively over the time span.



5. System Architecture and Interoperability

Clinical Decision Support System (CDSS) architecture generally comprises various layers such as database management, backend, and frontend but can be divided into Input and Output Layers for clarity. The Input Layer handles data pipelines from different sources—such as Electronic Health Records (EHR), laboratory systems, biomedical literature, and natural sources—standardizing their formatting, deidentifying sensitive information, and controlling access based on defined policies and rules. The Output Layer contains clinical interfaces for visualizing results and a service layer that allows other systems to query insights via well-defined interfaces. Although other, different component categorizations are also possible, this one suffices to illustrate the need—especially during development—for a clear analytical view of the overall architecture, independent from particular tools, vendors, or environments.

Input Layer development ensures up-to-date and reliable data pipelines for feeding the system. FastHealthcare Interoperability Resources (FHIR) standards are employed to serve high-quality data from various sources, including clinical data warehouses. Additional administrative functions, such as cohorts used in historical analyses, are done directly using suitable database management systems. The logical architecture follows a Microservices Development approach for distributed development, version control, deployment, and different technology selections for each service. Data pipelines for diverse sources are developed independently—optimizing performance and facilitating changes by enabling the immediate identification of dependent pipelines—but together ensure a cohesive system. Data pipelines for various other sources can also be added as needed, as long as each of them follows the defined standards.

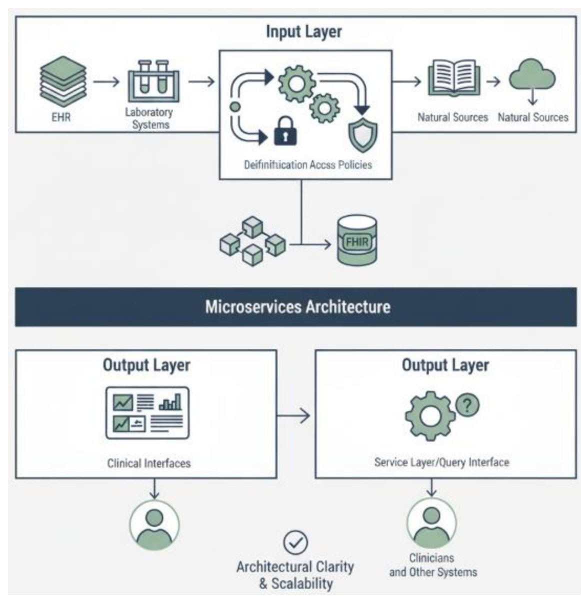


Fig 3: A Modular, Standards-Based Architecture for Clinical Decision Support Systems: Leveraging Microservices and FHIR for Interoperable Input-Output Pipelines

5.1. Integration Framework and System Cohesion

Data pipelines, standards, interface requirements, and governance structures provide the necessary cohesion for all components of a Clinical Decision Support (CDS) system, ensuring they work together seamlessly and that analytic insights flow smoothly into clinical workflows. Schemas, ontology, and a resource repository that define these requirements are maintained by a Data Integration

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 02 ISSUE: 02

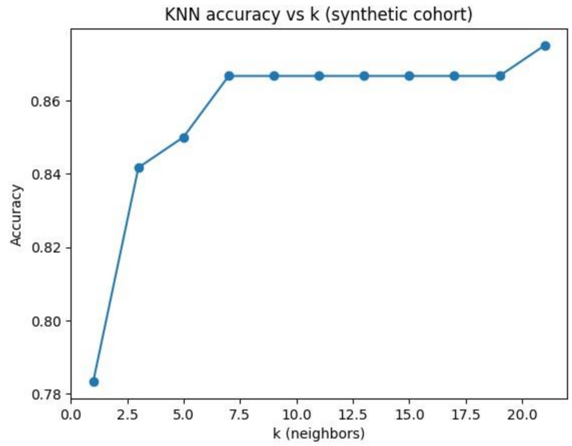
RECEIVED: APRIL 21 REVISED: MAY 10 ACCEPTED: MAY 026 PUBLISHED: JUNE 19

Team, a multi-disciplinary group of faculty and staff devoted to the integration of disparate data sources at a healthcare facility. Cohesion is critical: it avoids analytic silos and enables the generation of knowledge bases in areas of need or opportunity identified by members of the Discovery Team. The CDS system can provide a given knowledge base to clinicians in multiple forms through the solutions of multiple distinct Discovery Teams in their areas of focus. Integrating all the various components of the CDS system within a robust governance framework thus enables an enterprise-wide collaborative community for advancing Discovery and for providing fast, accurate, and context-sensitive Decision Support to clinical personnel.
A Data Integration Team composed of faculty and staff from multiple disciplines is devoted to the integration of disparate data sources across a healthcare facility. Schemas, ontology, and resource repositories that define the requirements supporting cohesion are maintained by the Data Integration Team. A critical goal is the enforcement of an integrated data architecture across all Maimonidean CDS initiatives. Cohesion avoids the emergence of analytic silos and enables the generation of knowledge bases in areas of need or opportunity identified by members of the Discovery Team. Cohesion permits the delivery of a given knowledge base in multiple forms through the solutions of multiple distinct Discovery Teams focused on specialized domains. A robust governance framework incorporating these principles fosters the enterprise-wide collaborative community essential for advancing Discovery and supporting seamless fast, accurate, and context-sensitive Decision Support to clinical personnel.

6. Clinical Validation and Evaluation

The clinical validation process involves the evaluation of whether the supporting evidence complies with the criteria established for clinical usefulness from the section on Integration Framework and System Cohesion. This evaluation is typically performed while designing a clinical study, which establishes the objective and the hypothesis to be tested, as well as the endpoints that will be used to assess the outcome of the experimentation. In the case of historical data usage for clinical decision support, the objective is evidently related to the determination of the clinical usefulness of the generated system – that is, if the hints generated and presented can assist the clinician within her work reasonably enough to be clinically useful. Moreover, it must necessarily be tested if its deployment has any impact on reducing adverse events in the clinical practice. As historical data can be used to develop multiple different decision support systems at the same time, it is also possible to test the clinical usefulness of these systems simultaneously, speeding up the validation process. A clinical study was performed that included three distinct decision support systems generation hints. Other inputs from the clinical practice, considered preliminary in nature, were included in the pilot study to assess their role in its practical performance.
To complement with real-world data and check the clinical performance of a warning system for acute kidney injury based on neural networks, other mechanisms were developed to use historical data for predictive analytics and assess historical clinical use. Results suggested that operational predictive analytics could help fill actual gaps in clinical practice, even using low-cost resources. The clinical study performed for gaining evidence on the clinical usefulness of the systems confirmed the best judgment of the clinical society on the hints generated. Containing practically no false-positive cases, the information provided did not overload users. Careful design avoided large workloads and offered acknowledgments of reduced risk for difficult-to-solve cases.

ISSN: 3067-4166
AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS
VOLUME: 02 ISSUE: 02
RECEIVED: APRIL 21 REVISED: MAY 10 ACCEPTED: MAY 026 PUBLISHED: JUNE 19



Equation 3: Logistic regression (explicitly referenced by the paper)

3.1 Model form (log-odds)

Let features be $x = (x_1, \dots, x_p)$ and outcome $y \in \{0,1\}$.

Define the linear score:

$$z = \beta_0 + \beta_1 x_1 + \cdots + \beta_p x_p$$

Logistic (sigmoid) function:

$$p = P(y = 1 \mid x) = \sigma(z) = \frac{1}{1 + e^{-z}}$$

Step-by-step derivation (from odds to sigmoid)

4. Start with odds definition:

   $\text{odds} = \frac{p}{1-p}$

5. Logistic regression assumes log-odds is linear:

   $\log\left(\frac{p}{1-p}\right) = z$

6. Exponentiate both sides:

   $\frac{p}{1-p} = e^z$

7. Solve for p:

   $p = e^z(1-p) = e^z - e^z p$

8. Collect p terms:

   $p + e^z p = e^z \Rightarrow p(1 + e^z) = e^z$

9. Divide:

   $p = \frac{e^z}{1+e^z} = \frac{1}{1+e^{-z}}$

3.2 Decision rule (turn probability into alert)

With threshold τ (e.g., 0.2):

$$\hat{y} = \begin{cases} 1, & p \geq \tau \\ 0, & p < \tau \end{cases}$$

6.1. Assessment of Clinical Validation Processes and Outcomes

Pilot studies of decision support based on historical patient data have thus far yielded positive results, but additional and more robust investigations are needed to validate the usefulness of such insights in real-world contexts. In these earlier pilots, an initial use case explored the association between antibiotic exposure during the first year of life and early onset sepsis. Decision alerts were presented to neonatologists directly in the electronic health record (EHR) system. The alerts contained a compliment regarding avoidance of unnecessary antibiotic exposure in early infancy and a reminder about the association with a rare but serious condition. Neonatologists responding to the alerts commented on their informativeness, but no specific decisions resulted from the encounter. A second pilot integrated a model predicting risk of William syndrome over life into the decision workflow for patients with unexplained developmental delay. The model was evaluated retrospectively in a cohort of these patients, informing about high risk but with no diagnosis made. In this case, the lack of detection was not unexpected due to the discontinuous nature of the disorder. A third pilot investigated the predictive performance of a temporal model estimating all-cause readmission risk in patients aged 40 and over under a cancer diagnosis and evaluated the safety of applying the model in a real-world scenario. Demonstrating an acceptable performance level and a

ISSN: 3067-4166
AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS
VOLUME: 02 ISSUE: 02
RECEIVED: APRIL 21 REVISED: MAY 10 ACCEPTED: MAY 026 PUBLISHED: JUNE 19

negated negative impact of alerting clinicians about low-risk patients mitigated risks for patients receiving alerts. Real-world experience can help to expose clinical utility limitations and effective mitigation measures in large-scale systems. In the context of using historical patient data for decision support, the focus appears particularly informative. The presence of diverse patient cohorts and clinical settings in published works establishes a foundation for both positive and negative findings related to specific prediction tasks and goes beyond isolated case reports.

7. Decision Support Presentation and User Interaction

A crucial aspect of any clinical decision support system (CDSS) is how it presents its insights to the end user, and the manner in which users interact with it. Two features of such insights are especially important: their clarity—so that users can easily understand the information—and the degree to which they create user trust. The user experience (UX) for engagement with the CDSS is also critical for ensuring that clinicians use it frequently, that they use it in the appropriate contexts, and that the information provided influences their decision-making.

An analysis of interaction patterns among treating physicians with the CDSS for a real-world deep vein thrombosis (DVT) dataset showed that presenting structured historical insight at the exact moment when the physician needed to make a diagnosis had the biggest effect on decision-making. Because the CDSS alerts the user specifically when the historical summary data is most relevant, it is particularly context-aware. In the context of a clinical pilot study, the CDSS was found to enrich clinical decision-making in a complex multimorbidity setting, and it did so without creating additional cognitive load for the treating physician. The insights therefore complemented—not replaced—the expert decision-making process.

When designing user interaction presents a considerable UX challenge. It is important to ensure that physicians find the CDSS clear, easy to use, and trustworthy. The goals of the interaction design and UX analysis are therefore to minimize the cognitive load required to understand and analyze the insights presented. The discussion covers several UX principles that could be considered during interaction design and engine development.

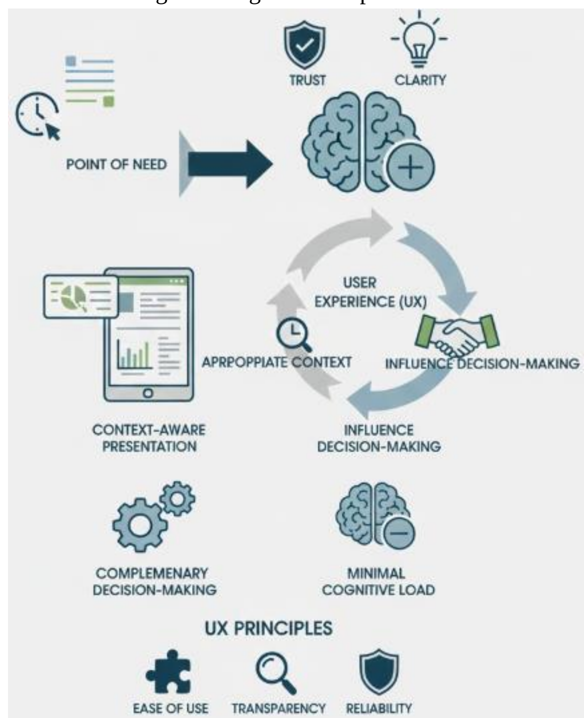


Fig 4: Context-Aware Clinical Decision Support: Optimizing User Experience and Cognitive Load for Trust-Based Human-AI Collaboration in Multimorbidity Management

7.1. User Engagement and Interaction Dynamics

Understanding how clinicians engage with clinical decision support (CDS) systems and insights generated from historical patient data is crucial in assessing their impact on clinical practice. Evidence on patterns of user interaction with these systems and their implications for decision-making processes, however, is inherently limited, as CDS systems typically yield only a set of recommendations without retrospective end-user feedback for evaluation. Moreover, potential effects on decision-making processes remain largely unaddressed in the literature. Two hospitals participated in an analysis of the interaction of intensive care unit (ICU) clinicians with insights generated from historical patient data over six months. Evaluation of the

ISSN: 3067-4166
AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS
VOLUME: 02 ISSUE: 02
RECEIVED: APRIL 21 REVISED: MAY 10 ACCEPTED: MAY 026 PUBLISHED: JUNE 19

frequency of checks engaged and clinical context surrounding the interactions provides a detailed understanding of their user engagement with and contextual usage of the generated insights, as well as potential factors influencing decision-making beyond aiding direct patient management.

Among a cohort of 1,533 ICU admissions from four years of data, the CDI system iteratively generated descriptive, predictive, and risk-profiling insights. These insights summarize the expressive power of the historical data and take a "Iceberg" approach to decision support. It is also confirmed that clinicians actively engaged the system checking the insights when supervising ICU patients beyond the active patient management stage. The role of clinical context should also be further assessed in shaping the user engagement and the interaction dynamics between clinicians and decision support systems. Further experiments are thus warranted to deepen the understanding of the impact of such insights on clinical practice, decision-making processes, and patient outcomes, as well as the user-centred aspects influencing these patterns.

Patient_ID	Predicted Risk p	Decision (Threshold τ=0.20)
001	0.41	High risk (alert)
002	0.06	Low risk
003	0.62	High risk (alert)
004	0.14	Low risk
005	0.34	High risk (alert)

7.2. User Experience Optimization and Interaction Design

Developing an effective user experience (UX) that conveys important information clearly is an essential consideration for many interactive systems. The factors that influence user engagement with a clinical decision support (CDS) system include the manner in which information is presented, the distribution of control between users and the system, and the effect of system-provided information on users' thought processes. Research has shown that under conditions of low workload, clinicians often rely more heavily on their clinical judgment than on decision support systems. This could limit the impact of decision support systems that could help prevent medical errors. Clinicians are less likely to use information from decision support systems that require significant alterations to existing workflows or demand heavy processing, and systems that provide information that is helpful but does not influence the final decision will have minimal impact on decision making. Incorporating these principles into the design of decision support and user interfaces could thus enhance CDSS usability.

User interpretation of the information provided by clinical decision support systems (CDSSs) may be enhanced by clearly displaying the CDS output using screens and data representations such as alert boxes, tables, graphs, and figures. Humans are naturally good at spotting patterns and trends in visual representations of data, meaning that graphical depictions are often preferred, as they provide information quickly and clearly. To promote motivation, an adequate balance between the amount and structure of the information should be found, as information overload often detracts from, while lack of stimulation can cause deterioration in, performance. Based on the goals of the CDSS application and the stage of the decision-making process, UX design guidelines should be formulated to support the presentation of decision support outputs that facilitate fluent interpretation.

8. Conclusion

In conclusion, the research provided answers to the questions posed as follows: (a) Several clinical decision support systems based on historical patient data have been implemented and evaluated over the last decade. The clinical validation processes and outcomes provide evidence in favor of their clinical usefulness. (b) Users have spontaneously engaged with some of these systems without any awareness-raising efforts, and the few engagement studies in the literature suggest that support from historical data may affect real-world decisions. (c) Analysis of interaction patterns suggests that users engage with the CDS before proceeding to diagnosis and treatment choice, whereas the availability of explanations increases

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 02 ISSUE: 02

RECEIVED: APRIL 21 REVISED: MAY 10 ACCEPTED: MAY 026 PUBLISHED: JUNE 19

engagement in later stages. Principles derived from UX design can help foster clear communication, trust, and limited cognitive load. (d) Although users have reported positive experiences, interaction dynamics and UX still require in-depth investigation. Such exploration is essential to establish how the presentation of decision support within clinical workflows influences the use of the information and thereby its actual effectiveness from a clinical perspective.

In brief, several clinical decision support systems based on historical patient data have been implemented and evaluated over the last decade. The assessment of the interactions with these systems reveals patterns of use, and the outcomes of the clinical validation processes provide evidence in favor of their clinical usefulness. Users have spontaneously engaged with some of these systems without any awareness-raising efforts, and the few engagement studies in the literature suggest that support from historical data may affect real-world decisions. Analysis of interaction patterns indicates that users engage with the CDS before proceeding to diagnosis and treatment choice, whereas the availability of explanations increases engagement in later stages. Principles derived from UX design can foster clear communication, trust, and limited cognitive load. Although users have reported positive experiences, interaction dynamics and UX still require in-depth investigation to establish how the presentation of decision support within clinical workflows influences information use and, thereby, its actual effectiveness from a clinical perspective.

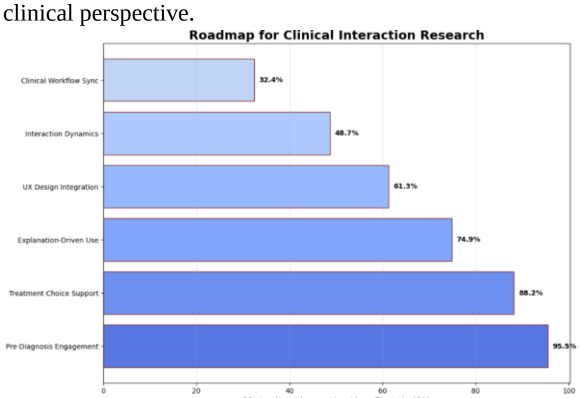


Fig 5: Roadmap for Clinical Interaction Research

8.1. Summary of Findings and Future Directions

Clinical decisions supported by historical patient data have shown that descriptive analyses can support thorough exploration of data completeness, reliability, or potential risks due to detected biases. The insights resulting from this analysis also suggest that the deployment of these systems can reduce the impact of historical, clinical, or socioeconomic self-selection biases for pathologies with a high treatment burden, such as the joints of the lower extremities or multi-ligament knee injuries. Predictive modeling, when grounded by historical data and properly validated, becomes a key enabler of clinical decision support. Moreover, if deployed in the clinical routine, predictive models must have a theoretical clinical impact for the user and be clinically validated, demonstrating improved decision-making, effectiveness, and safety. Clinical decision support systems have often been considered with an operational brain, providing brief alerts for immediate clinical use. However, exploratory analyzes suggest that these systems may behave more like an analytical brain, introducing a novelty in the clinical workflow with benefits beyond immediate alerts. These novel properties open avenues for exploration in user experience and trust in decision-making processes. The proven effectiveness of these systems can allow users to trust them more, further enhancing their ability to provide clinical information beyond alerts. A better user experience can contribute to a more natural interaction with the system and therefore greater trust over time.

9. References

[1] Bates, D. W., Saria, S., Ohno-Machado, L., Shah, A., & Escobar, G. (2014). Big data in health care: Using analytics to identify and manage high-risk and high-cost patients. Health Affairs, 33(7), 1123–1131.

[2] Sutton, R. T., Pincock, D., Baumgart, D. C., Sadowski, D. C., Fedorak, R. N., & Kroeker, K. I. (2020). An overview of clinical

ISSN: 3067-4166

AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS

VOLUME: 02 ISSUE: 02

RECEIVED: APRIL 21 REVISED: MAY 10 ACCEPTED: MAY 026 PUBLISHED: JUNE 19

decision support systems. Journal of Medical Internet Research, 22(10), e24862.

[3] Berner, E. S. (2016). Clinical decision support systems: State of the art. Agency for Healthcare Research and Quality.

[4] Kawamoto, K., Houlihan, C. A., Balas, E. A., & Lobach, D. F. (2005). Improving clinical practice using clinical decision support systems. BMJ, 330(7494), 765–768.

[5] Musen, M. A., Middleton, B., & Greenes, R. A. (2014). Clinical decision-support systems. In Biomedical informatics (pp. 643–674). Springer.

[6] Shortliffe, E. H., & Sepúlveda, M. J. (2018). Clinical decision support in the era of artificial intelligence. JAMA, 320(21), 2199–2200.

[7] Topol, E. J. (2019). High-performance medicine: The convergence of human and artificial intelligence. Nature Medicine, 25(1), 44–56.

[8] Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., Wang, Y., Dong, Q., Shen, H., & Wang, Y. (2017). Artificial intelligence in healthcare. Stroke and Vascular Neurology, 2(4), 230–243.

[9] Goldstein, B. A., Navar, A. M., Pencina, M. J., & Ioannidis, J. P. A. (2017). Opportunities and challenges in developing risk prediction models with electronic health records data. Journal of the American Medical Informatics Association, 24(1), 198–208.

[10] Obermeyer, Z., & Emanuel, E. J. (2016). Predicting the future — Big data, machine learning, and clinical medicine. New England Journal of Medicine, 375(13), 1216–1219.

[11] Friedman, C. P., Wong, A. K., & Blumenthal, D. (2010). Achieving a nationwide learning health system. Science Translational Medicine, 2(57), 57cm29.

[12] Hripcsak, G., & Albers, D. J. (2013). Next-generation phenotyping of electronic health records. Journal of the American Medical Informatics Association, 20(1), 117–121.

[13] Jensen, P. B., Jensen, L. J., & Brunak, S. (2012). Mining electronic health records. Nature Reviews Genetics, 13(6), 395–405.

[14] Luo, G., & Tang, C. (2012). Clinical decision support systems. IEEE Intelligent Systems, 27(4), 70–73.

[15] Greenes, R. A. (2014). Clinical decision support: The road ahead. Academic Press.

[16] Elwyn, G., Frosch, D., Thomson, R., Joseph-Williams, N., Lloyd, A., Kinnersley, P., Cording, E., Tomson, D., Dodd, C., Rollnick, S., Edwards, A., & Barry, M. (2012). Shared decision making. Journal of General Internal Medicine, 27(10), 1361–1367.

[17] Miller, R. A. (1994). Medical diagnostic decision support systems. Journal of the American Medical Informatics Association, 1(1), 8–27.

ISSN: 3067-4166
AMERICAN DATA SCIENCE JOURNAL FOR ADVANCED COMPUTATIONS
VOLUME: 02 ISSUE: 02
RECEIVED: APRIL 21 REVISED: MAY 10 ACCEPTED: MAY 026 PUBLISHED: JUNE 19

[18] Bright, T. J., Wong, A., Dhurjati, R., Bristow, E., Bastian, L., Coeytaux, R. R., Samsa, G., Hasselblad, V., Williams, J. W., & Musty, M. (2012). Effect of clinical decision-support systems. Annals of Internal Medicine, 157(1), 29–43.

[19] Middleton, B., Sittig, D. F., & Wright, A. (2016). Clinical decision support. BMJ, 353, i2039.

[20] Ammenwerth, E., Iller, C., & Mahler, C. (2008). IT adoption and the interaction of task, technology and individuals. Methods of Information in Medicine, 47(3), 245–252.

[21] Bates, D. W., & Gawande, A. A. (2003). Improving safety with information technology. New England Journal of Medicine, 348(25), 2526–2534.

[22] Reddy, M. C., & Dourish, P. (2002). A finger on the pulse. Proceedings of the ACM Conference on Computer Supported Cooperative Work, 344–353.

[23] Shah, N. H., & Tenenbaum, J. D. (2012). The coming age of data-driven medicine. Journal of Biomedical Informatics, 45(1), 9–17.

[24] Simon, H. A. (1997). Administrative behavior (4th ed.). Free Press.

[25] Sendak, M. P., D'Arcy, J., Kashyap, S., Gao, M., Nichols, M., Corey, K., Ratliff, W., & Balu, S. (2020). A path for translation of machine learning products into healthcare delivery. EMJ Innovations, 4(1), 70–80.

[26] Wiens, J., Saria, S., Sendak, M., Ghassemi, M., Liu, V., Doshi-Velez, F., Jung, K., Heller, K., Kale, D., Saeed, M., Ossorio, P., Thadaney-Israni, S., & Goldenberg, A. (2019). Do no harm. Nature Medicine, 25(9), 1337–1340.

[27] Nemati, S., Holder, A., Razmi, F., Stanley, M. D., Clifford, G. D., & Buchman, T. G. (2018). An interpretable machine learning model for accurate prediction of sepsis. Critical Care Medicine, 46(4), 547–553.

[28] Caruana, R., Lou, Y., Gehrke, J., Koch, P., Sturm, M., & Elhadad, N. (2015). Intelligible models for healthcare. Proceedings of the ACM SIGKDD Conference, 1721–1730.

[29] Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). Why should I trust you? Proceedings of the ACM SIGKDD Conference, 1135–1144.

[30] Holzinger, A., Langs, G., Denk, H., Zatloukal, K., & Müller, H. (2019). Causability and explainability of artificial intelligence in medicine. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 9(4), e1312.