Deep Science Publishing
https://doi.org/10.70593/978-93-49307-36-0_3

DeepScience
Open Access Books

Chapter 3: Building a data-driven healthcare ecosystem that supports evidence-based clinical decision-making

3.1. Introduction

In the shift from a heretofore individual-clinical level to a population-policy level, where the decision-making processes typically become more uncertain, variable, and complex, it is further proposed that contextual influences play a crucial role in determining what constitutes evidence and how that evidence is utilized effectively in making informed decisions. Recent research and firsthand experience suggest that innovative evidence-based policy, grounded on scientific knowledge and available reproduction of the real-world evidence, is highly relevant across the bio-medical, public health, and clinical fields, and can be built by epidemiological case studies. Discussing the potential transfer of databases to informed policy and personalized healthcare decisions involves, but is not limited to, healthcare providers' economics as hospital administrators and family doctors, public health policy in terms of health promotion professionals, and patients' choices in light of the modern bio-communication technologies of the internet and social media.

Firstly, the issues related to the current transfer of database information into knowledge about the health status of the population and about where modifiable predictors (i.e., lifestyle and socio-demographic variables) are involved are considered, along with the use of this knowledge for designing intervention plans. Some issues of potential interest are introduced and highlighted for analysis through a review of the literature, as well as the initial code work and the projects delivering participant information, published health promotion materials, and e-mail messages. Secondly, it is shown that the back-transformation of biomedical research versions of treatment efficacy into general health policy- and patient-oriented versions, about the compliance of these details with the strict, internationally established criteria of empirical evidence, is severely limited. Subsequently, it is emphasized that for the most part, the publicly available databases are devoid of the medically meaningful (and expected) detailed topography-based descriptions, such as the relative contributions of fibroadipose tissue within the bulk mass of the breast and the amount of muscle fiber within the corporeal mass.

Fig 3.1: Data-Driven Healthcare

3.1.1. Background and Significance

This realization generated different projects in the following years aimed at the development of related standards to which organizations could conform, so that once a disaster would be experienced, national healthcare professionals would be able to rely on their tight time-limited collaboration. Subsequently, the readiness of information transfer regarding a biological or chemical mass destruction event between first responders and healthcare institutions has been studied and observed that since such data would also arrive from a variety of healthcare settings then differences in them would affect the level of care that would be provided to the victims. In an attempt to enhance preparedness in cases of mass casualty events it has, therefore, been suggested that also clinical data of the patients would have to be formatted uniformly, such as the various administrative, medical, and nursing data items had already been, so that it could be easily conveyed and understood.

3.2. Role of Artificial Intelligence in Healthcare

A decision support system can be a computerized tool or software that supports healthcare professionals by making informed decisions according to patient care by integrating patient data, relevant medical knowledge or guidelines, and appropriate analytical tools (Google Health, 2024; IBM Watson Health, 2024). In fact, some of these decisions, such as complex diagnosis or adjusting drug treatments, have to be taken in real time also during surgical procedures. Such immediacy can not be achieved in practice without automating the decision making process. In those cases, indeed, the availability of such systems can provide an automated way to make informed decisions, i.e., providing appropriate recommendations, and support or even substitute healthcare professionals decision-making. This offer is open to many possible scenarios like conditions diagnosis, recommendations practice like appropriate time frame to visit, prevention operation by following lab periodic controls or at least keeping a healthy lifestyle, and healthcare operations optimization.

Artificial intelligence (AI) has been playing a pivotal role in such a revolution. AI technologies have deeply changed the way data has to be collected, analyzed, and exploited for supporting evidence-based decisions in the healthcare sector. These systems are able to harness huge amounts of digital data sources, such as those coming from the digitization of healthcare data, thus processing structured and unstructured patient data, in order to extract hidden – just before impossible to be identified – knowledge. In a nutshell, a decision support system in the healthcare sector is an emerging modeling paradigm in which these technologies are developed to enhance clinical quality, assist intervention planning, and provide recommendations for clinical treatment.

3.2.1. Machine Learning Applications

Over the past decades, the healthcare sector has seen an increase in availability and overall volume of patient data. With the expansion of electronic medical records, it’s estimated that the volume of medical data is doubling every 4 years and will reach 2314 exabytes by 2020. The emergence of big data technology provided an opportunity to analyse previously stored and unused patient data. That analysis has the potential to discover new insights in both individual and population health. The healthcare industry is expected to significantly benefit from predictive big data analytics. According to one assessment, analysing big data in healthcare could result in 300 billion dollars in savings in the US, while the global potential is estimated to be over 1 trillion dollars. Health institutions have been working on creating a data-driven ecosystem by accumulating data from patient histories, hospital records, insurance claims, clinical data, genetics, social and environmental factors, etc. Every second, patients generate large volumes of data that pertain to their health, seeking medical advice, being diagnosed and treated by doctors in hospitals or clinics. The application of artificial intelligence (AI) technologies can effectively assist various clinical tasks, such as case triage and diagnosis, image scanning and segmentation, decision making, and risk prediction. Rooted in many respected evidence-based publications on healthcare and data analytics, healthcare systems that foster a data-driven ecosystem are achieved through the concept of evidence-based and patient centred medicine by improving the interoperability of clinical decision support systems at point of care. Also, the healthcare system is supported by a new approach for secondary use of the patient data generated in clinical practice, where these data are analysed to discover new knowledge through big data analytics. On the upside, the learned clinical knowledge is later used in clinical decision support systems, consequently providing physicians with recommendations that support evidence-based and patient centred clinical decisions.

3.2.2. Natural Language Processing

The furor surrounding artificial intelligence (AI) in healthcare has led to rapid advancement in digital medicine across multiple clinical specialties, primarily enabled by big data generated through the digitization of healthcare (Johns Hopkins University, 2024; PathAI, 2024). Approach to unlocking such information computationally through natural language processing (NLP) is of paramount value to advancing healthcare AI. Several types of NLP systems currently in popular use are introduced, each pragmatically tailored to the scope of its clinical application. The majority of entities are captured via NLP-as-a-service, an R&R pathway for data extraction that improves both efficiency and precision. Limitations with current practices and desires to better provide and standardize NLP in furtherance of healthcare AI development are also discussed. Data standardization acts as an enabler for additional AI tasks, such as population cohort identification for prospective analyses or surveillance. Several systems for the extraction of unstructured medical information that have been in operation are described. For example, a Mayo-developed natural language processing system has been screening computed tomography reports to facilitate surveillance of pulmonary invasive mold diseases in patients with haematological malignancies. Also described is an externally developed system that is used to screen the electronic health record for the early recognition of multiple sclerosis. Of particular focus are systems that extract unstructured medical information relevant to another clinical event (e.g., a cohort selection event) and consequently inform clinical decision-making (e.g., a treatment event). Such systems are employed within the scope of several clinical specialties to advance healthcare AI.

Text mining of patient-generated messages commonly made through patient portal to predict sepsis using a machine learning network architecture. Sepsis, a leading cause of mortality in the United States, is difficult to predict from clinical visits only where natural language processing of patient-generated text and information within the EHR provide increased predictive value. Patient-generated text includes notifications of their general illness and EHR-derived clinical notes.

Recognition of cardiovascular procedure reports and applying text mining on procedure reports achieved successful identification of heart condition patients. Adapting cardiovascular procedure report processing method to similar report formats containing cholesterol and examining coronary and aortic heart conditions rather than artery. Early detection of multiple cancers via lung cancer screening is urged. Building a natural language processing tool to extract data from radiology files is suggested.

3.3. Patient Engagement and Empowerment

The goal of patient engagement is to involve the patient in the decisions leading to their own health and quality of life improvement. Despite recent factual clinical improvement, patients perceive that healthcare professionalism still needs to improve patient engagement. The disease patient rapid progressive late stage creates a preference for a ‘frequent results’ healthcare civil contract establishment instead of a professional civil dialogue. The angiostrongyliasis from paratenic host infection monitored by serum anti-BmR1-IgG may be a valuable tool to support clinical decisions. In the perspective of the patient, this detection and assessment of angiostrongyliasis infection aeroelasticity enhances a better engagement of the decisions leading to the integration of the healthcare strategy of disease angiostrongyliasis treatment. Aiming to analyze the laboratory outcomes of a hypothetical patient for the same disease with two different civil positionalities ‘paratenic host’ and ‘healthcare professional’, a mathematical-physics model simulating the PVL model of the serological angiostrongyliasis infection monitoring is defined. Using concepts from game theory, it was found that there are equilibrium situations where both players, after transmitting and analyzing the test results produced by sero-diagnostic tests of the angiostrongyliasis, get clinical benefits which are shared according to predefined agreements. It is also observed that different from the ‘treatment agreements’, clinical benefit in the ‘symptomatic patients triggering protocols’ is not a constant but rather time-dependent. An intuitive reasoning to get a perspective of these scenarios is also presented. This ENGAGEment LAB model may suppose a significant improvement of the establish of a ‘smart medicine’ enhanced PDCA cycle. The precision and individualized treatment strategies of patient-specific health care providers associated with the civil positionality are also discussed gaining a better insight of the disease CSP associated with health belles-lettres. This proposed theoretical approach can be extended to all kinds of serological tests to detect angiostrongyliasis as well as to any other similar infection regarded on medical, veterinary, or phytomedicine.

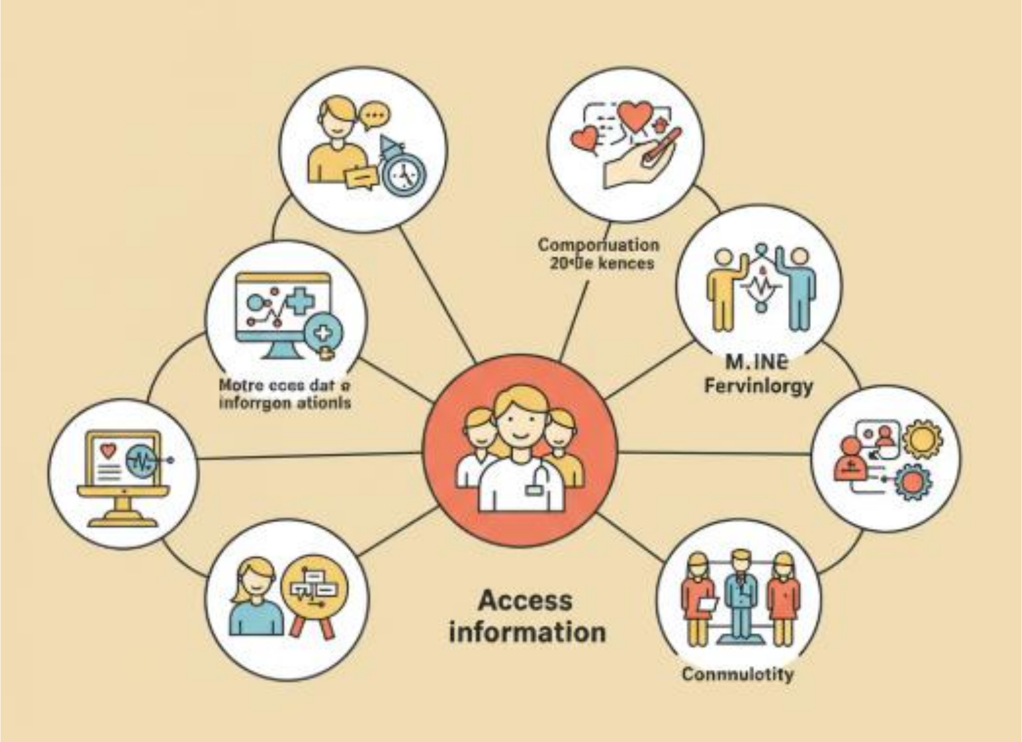

Fig 3.2: Patient Engagement and Support

3.3.1. Health Literacy Initiatives

Project partners have continued to collaborate during the NBA YHI to plan for its sustainment, and they are investigating the best ways to inform the types of adaptations that would be necessary in order for NBA YHI to be scaled. The partnership is strengthened by ongoing project rounds, which allows partners to be better informed of stakeholder needs and for AD to recognize partner requirements from their perspective. Regular evaluation of the partnership has focused on partner roles, interventions, data collection, capacity building, primary grant mechanisms, evaluation, and potential expansion. This evaluation was found to be critical to both partners involved and to those external to the project, and continuous evaluation processes will be needed moving forward.

Continuation of the ongoing efforts of NBA YHI partners may serve as an important road map for health literacy programs initiated in other states. One such guide is understanding that after approximately 5 years of collaboration, partners began exploring opportunities for scalability of the statewide systems-level combination clinical-academic-community health literacy approach to medical services in South Carolina. Sustaining this level of effort and collaboration in health systems level initiatives can be challenging, as is evident by evaluations of such initiatives. Individual working groups may fragment, funding may cease, provider buy-in may lack, or competing priorities may arise.

3.3.2. Patient Portals and Mobile Apps

Users of healthcare services are not only patients. Clinicians, too, are afflicted with a healthcare system teetering on the point of saturation (as well as being at high-risk of wear and burnout). When medical records transformed from manila folders to digital folders on desktop computers, a system was envisaged in which there would be a reduction in human laboratory bottlenecks—blood sample “X” goes into machine “Y” and spits-out result “Z.” This would, in-turn, zeitgeist patient throughput ratios, bed occupancy rates, and so forth. However, real-world operation of this system was more complicated. So-called “medical judgement” was still required to plan, analyze and make decisions. Bedside manner was still pivotal for provocation, intentionally sympathetic listening. Personal experience, advice from trusted-colleagues, consultation of textbooks and/or online journals, were all critical in resolving unsure case ambiguities and unknowns. Descriptions sparing of unnecessarily perplexing medical terminology, elucidative diagrams, or lay-terminology, were treasured, especially after a typically vexing call with a superior. However, such niceties often completely outgrow the time a consultation could proffer. All of this fed the mesmerizing narrative of a computer system capable of trawling “big data” from records and existing literature that would provide just-in-time forewarning, making decisions and diagnoses. In truth, off-the-backs of simple network hardware and a very basic Human-Computer-Interface (HCI) (and with a visceral push from a predominantly doctoral technophile proof-of-concepts) vast coffer of private health data started finding their way to the burgeoning number of “health-tech” start-ups. Meanwhile, multiyear projects endeavored to develop generalisable learning algorithms able to predict cycles of care. A marked success was made within academia with the annual prediction challenges. Here was not only the opportunity to interrogate existing data, but also examine and incorporate the most cutting-edge techniques from algorithmic-machine-learning and statistician-physician-research from across talents of the world.

3.4. Data Sources in Healthcare

As the amount of health data being recorded continues to grow, so too do the types of sources. There are many sources of real-world health data including claims, electronic health records (EHR), administrative health data, health outcome data, lab results, or a patient's self-reported outcomes . The impact of clinical evidence varies as much as the health conditions they describe. For some conditions such as heart attack or stroke, evidence based on clinical trial data grounds a wealth of clinical guideline recommendations. For the majority of health conditions, the evidential base rests almost entirely on evidence developed from retrospective analysis of patient records. The practice of medicine increasingly depends on the use of computer systems like electronic health record (EHR) systems. When applied to patient care, clinical decision support system (CDSS) is computer-based knowledge that employs patient data to generate case-specific advice and is then presented to clinical staff in real time.

Shareable CDS healthcare applications are being proposed to develop access and link to patient data from EHRs, provide CDS, and add representational features for healthcare professionals and patients. The HL7 FHIR standard is being leveraged in support of data access approaches. FHIR specifies a library of standard clinical resources, which can be readily mapped to existing EHR data models. FHIR utilizes HTTPS and creates a standard API to share information among applications such as EHRs. This controlled and standardized access provides new opportunities for creating CDS abstraction interfaces. A significant shift is the implementation of CDS on the SMART platform to embed web applications into EHR workflows. The FHIR-based apps combined with the SMART platform pave the way for “SMART on FHIR” app stores. This encapsulation of standards-driven data access shows progress towards using sharable and custom external CDS to address complex clinical scenarios.

3.4.1. Electronic Health Records (EHRs)

Direct provision of care leaves a large volume of complex data in its wake, which could be used to create a new, data-driven ecosystem that supports evidence-based clinical decision-making. However, health data is recorded primarily for patient care rather than secondary uses. Electronic Health Records (EHRs) are central information technologies for collecting electronic health data. The pervasive use of EHRs has enabled an exponentially increasing collection of electronic health data, which is expected to reach 2314 exabytes only for the year 2020. EHRs are the core information systems used by healthcare organisations to capture textual, numerical and image data on virtually all aspects of patient care and wellbeing. Healthcare stakeholders, including researchers, practitioners, patients, industry partners, policy makers, and indeed the public understand that health data, and therefore EHRs, hold a largely untapped potential for monitoring, tailoring, and conducting interventions to improve both the processes and outcomes of care.

EHR data presents a complex intersection of information that is structured, unstructured, trace-based, multimodal and multi granular. Early research on health data analysis focused mostly on structured numerical data. The scientific and healthcare research and practice communities have shown a growing interest in complex and heterogeneous patient data, such as EHRs and other emerging health data types like genetic, imaging, sensor-based or social media data. Over 80% of health data still resides within unstructured text, images, waveforms or free text clinical notes. Additionally, most EHR data generated during care is largely ignored in its current structured form. Information on clinical reasoning, care coordination or teamwork, decision-making and socio-technical implications of care delivery, are some examples of data embedded within a richer EHR space that have remained largely untapped. EHR data is also stored at different levels of granularities and temporalities, inherently affecting the usefulness of data for specific analyses and applications in complex ways. Joint modelling of structured, textual, and temporal patient data has only recently begun to emerge, focusing in the main on the induction of clinical predictive models. Text and image data appear to be the largest untapped resources available for secondary use within EHRs. While EHR notes play an important role in the rationale for clinical decision-making, 76% of all EHR notes still remain unopened, adhering to the notion ‘that the bigger the haystack the harder it is to find the needle’. Given a data-driven healthcare ecosystem, with constraints and challenges to harness its potential, the research actions required to approach the data volume and variety, challenges at the data, information, and knowledge levels are analysed, alongside the hammering role of privacy and security concerns.

3.4.2. Clinical Trials and Research Data

Broad access to, and understanding of research data by patients, clinicians, and researchers are critical to building a data-driven healthcare ecosystem. Although a comprehensive and machine-understandable model covering every piece of knowledge and data used in clinical practice is highly desirable, attaining that ideal model would be nearly impossible. Additionally, empirical research is needed to grow evidence that care relationships without an active research protocol are associated with better quality or cheaper care. More achievable objectives involve generating datasets of clinical care in ways that are either independent of health provider incentives or that can be adjusted for in the dataset, and to distribute these datasets widely to increase the number of researchers who have a chance to find problems. A potential solution would be to reconsider the manner in which electronic health record (EHR) data are produced.

Specifically, EHR vendors, healthcare systems, providers, and other stakeholders would need to consider how EHR data might be engineered, from conception, to be more suitable for clinical research. This would entail expanding the consideration of EHR data to focus not only on its utility for patient care, but also on its utility for subsequent clinical research. Clinician-scientist teams and professional EHR specialists have taken the initiative to characterize how medical events are most typically captured in EHRs for a given condition. The resultant sets are composed of data elements used by clinicians to document patient care, and include endpoints that can be derived directly from the patient care entries or related EHR data. Furthermore, as an increasing number of initiatives turn to the common EHR as a font for research data, it is crucial to assess the quality and utility of the EHR-derived data. There is a disparity of tools and a lack of incentive to support the creation and curation of aligned real-world studies data, limiting the rate of learning about how to reduce social barriers, exemptions, misunderstandings, and infrastructure shortfalls that may be inhibiting the most effective, fair, and efficient generation and application of clinical evidence.

3.5. Building a Data Infrastructure

The challenges are numerous when creating a data-driven environment given the complexity and fragmentation of health data. There are many ways to analyze, understand, and ameliorate these risks, many of which have thus far received insufficient attention. A first step could be to recognize them, and frame the problem of health data more holistically and in terms of infrastructure. Effort to build health data infrastructure would need to attend to the sharing, integration, and storage of patient data. While this figure is illustrative and stylized, it accurately captures the fragmentation of health data in the US healthcare system. Health data are generated, collected, and stored by a number of different actors. Clinicians generate data about patients by caring for them. They collect data by talking to them, examining them, and ordering various tests. The interactions of patients with the health care system also generate data (e.g., their name, their insurance information, bills for the care they receive). Because each actor is responsible for generating, collecting, and storing the data for its own interactions with patients, the interactions of any individual patient involve a large number of different entities. Moreover, these actors are diverse, ranging from individual hospitals and physicians to large technology companies, and they span both the public and private sectors, which has led to substantial risks.

Given the potential benefits of integrated patient data, considerable effort must be expended at a systemic level. A simple health data infrastructure would consist of a single actor collecting and storing all relevant patient data, from the interaction of patients with the health care system including both digital and analog information.

However, there are good reasons to think that would be an undesirable model. As discussed above, there are substantial risks to fragments of health data. If these data are decentralized and held by a number of different actors, then they are inherently fragmented between different patients. Moreover, there are good reasons to want health data to be distributed across different types of actors. There is value in the health data generated and collected by wearable technology companies that would not be possible (or at least as straightforward) to generate with respect to the existing health IT systems used by clinicians and in healthcare facilities. By contrast, centralized health data ameliorate many of these risks – data are comprehensive and easily obtained, creating an apparently complete dataset to guide analysis and policy changes. Further, centralized health data provide a means of comprehensive evidence about the failure modes, causation, and best practices by generating a complete historical record of all transactions.

Fig : AI is Transforming HealthCare

3.5.1. Data Integration Techniques

Building an extended, data-driven healthcare ecosystem that can enable doctors to make evidence-based and data-supported clinical decisions demands a lot of technical and application-level work. First, data from different sources needs to be integrated. For structured data such as laboratory results and other observations, existing state-of-the-art techniques can be used. For unstructured or complex structured data such as histopathology images, a combination of promising research techniques can be reused and extended to other data types. Furthermore, medical interpretations should be produced from the analytical and computational processing of integrated medical data. Although interpretation of clinical data has been a part of clinical practice for centuries, computer-based biomedical data and clinical interpretation poses a different kind of challenge. A combination of well-known black box techniques such as Decision Tree, SVM, and Deep Learning neural network together with a rule extraction method allows medical professionals to understand complex mathematical relations and dependencies behind the machine-learning-based models better. It is believed that “strong” medical interpretations are more understandable and thus have higher acceptance in the medical community (as opposed to “weak” medical interpretations). Concepts like relative risk and odds ratio that are used in statistical analysis of clinical data can be used to explain to patients the mathematical results obtained from complex data mining models. Medical interpretations of medical data should also bring significant added value in terms of new insights and relations extracted from pre-processed digital data which are not found in literature because of the complexity and the amount of considered data. A medical interpretation is considered as a complete explanation of particular medical data along with all dependencies, relations, findings, and insights that can be deduced from it and are stated in natural language. To keep the explanations more concise just the most important concepts and relations are elaborated, expanding the interpretation overview with different sections. A medical professional would be able to decide on further diagnostic and treatment procedures based only on these findings.

3.5.2. Interoperability Standards

Interoperability is a key requirement to build a data-driven healthcare ecosystem that supports the generation of health data in a seamless way and the reuse of the data in different healthcare scopes. The growing interest in evidence-based clinical decision-making (CDM) urges the integration of clinical practice with research. In this context, the need for interoperability is even stronger in order to ensure that clinical decisions are supported by high-quality evidence originating from the big data generated in routine care.

A number of research studies are currently driven to the development of computer-based decision support mechanisms that help clinicians directly capture patient data for research. Although these systems proved to be able to effectively increase the recruitment rate, the translation of research protocols in data entry forms is often a non-trivial task involving the rearrangement of care-specific ontologies and the communication with many clinical vendors. For this reason, there is a growing interest in the development of standards and connectors that support broader interoperability between EHR systems and a larger variety of data capture mechanisms. In order to maximise the socio-economic benefits of investment and foster technological research in a sensitive domain, there is a great need for ancillary standardisation. Data harmonisation is increasing in importance in clinical care, since heterogeneous data are to be captured from different clinical data acquisition systems. The problem gets worse in a cross-domain scenario, where data need to be interpreted in different ontologies in clinical care and research contexts. Interoperability is the ability of making systems and organisations work together (effectively and efficiently), the result of a harmonisation process among different parts. In the current scenario, data (i.e. their meaning) is supposed to be shared among different organisations providing care and conducting research. At least syntactic and semantic interoperability are required, respectively meaning that data can be exchanged ensuring standard data formats and with a common set of concepts. At present, it is recognised that semantic interoperability can be only partially achieved since interpretation is driven by personal, local, socio-economic and governmental aspects. EHR with different information models and terminologies have been developed in different countries. These EHR are deployed in public healthcare organisations at very different levels of care and with a strongly variable degree of computerisation, lexical resources capability and integration with external organisations.

3.6. Data Analytics in Healthcare

Challenges regarding healthcare data analytics are diverse. Many secondary studies highlighted problems with missing data, low-quality data, and datasets stored in various formats that are not interoperable. Furthermore, some studies raised the concern of missing techniques to visualize the outputs of different data analyses. Many of the new implementations and increases in data require new computational infrastructure for feasible use. Some studies raised ethical concerns about data collection, merging, and sharing. Broad consent is a prerequisite, according to Article 9 of the General Data Protection Regulation. Data privacy is a multifaceted concept, especially when datasets cover multiple countries with different legislations. Many secondary studies called for a multidisciplinary collaboration between medical and computing experts and emphasized that analytics implementations must align with medical vocabulary and rules. Furthermore, data analytics, especially when more complex solutions are in question, operate on a black box principle. That is, it is not obvious how the implementation reaches its conclusion. Some secondary studies also pointed out that existing analytics solutions implemented in different environments are not portable into other environments and may not be fully integrated into actual day-to-day work tasks.

3.6.1. Descriptive Analytics

A data-driven healthcare ecosystem can be built to facilitate evidence-based clinical decision-making to improve the quality and cost of chronic care. It is critical to collect data for risk prediction, longitudinal monitoring, and policy evaluation. Data silos can be broken down to integrate heterogeneous sources of healthcare data. Descriptive analytics can be applied to mine health data for evidence generation. Predictive analytics can be applied to generate evidence-based guidelines for clinical decisions. Machine learning or optimization can be leveraged for policy learning to identify the most cost-effective intervention methods. Insights are generated more quickly through a combination of longitudinal monitoring insights and policy evaluation insights. A data-driven healthcare ecosystem can be built to seamlessly integrate data collection, processing, analysis, and decision-making. Once patients are identified through registries, billing claims, or primary care records, a team of physicians, dietitians, and care managers at a Disease Management Center conducts a comprehensive Needs Assessment to identify medical, dietary, and logistic risks. These assessments are based on clinical practice guidelines. Then, a Personalized Care Plan is developed for each patient based on the risk assessment. Patient behavior is monitored on a monthly basis through Remote Monitoring Sensor Networks. A Rapid Response Team is also sent to the patient’s household to investigate emergency cases. This system is unique in that it incorporates needs assessments and feedback loops, and combines computationally intensive simulations and real-time sensor networks. Over time, policy-relevant data is gathered, stored in data silos, and used in aggregate form to optimize the overall performance of the Disease Management Center. However, datasets are often locked away in walled gardens, and are statistically toxic to one another. Even if a hospital system digitizes its patient medical records, it does bar anyone from accessing the records to analyze them statistically. Data silos are broken down to collect and integrate a diverse range of public and non-public healthcare data sources. Patient medical records, public health records and datasets concerning non-medical determinants of health are integrated. Descriptive analytics is used to mine healthcare data to generate testable evidence regarding healthcare outcomes, health delivery systems, and health policies.

3.6.2. Predictive Analytics

In an era of data-driven processes, the health-care industry generates a vast amount of data across various sectors. This phenomenon is particularly relevant in a clinical environment where the primary focus is patient care. As methods and tools for the analysis of clinical and operational big data are readily maturing, the gap between the potential for medical data insights and the usage in everyday clinical practice by health-care professionals still persists. The Paper aims to provide an overview of beneficial strategies for the development of a data-driven health-care ecosystem, which helps bridge the gap between clinical data insights and patient-tailored, evidence-based clinical decision-making. These strategies aggregate aspects from interfacing with electronic health records, data collection and analysis, current state-of-the-art medical data analysis tools and methods, knowledge generation from medical big data, up to the development of predictive, data-driven models supporting personalized medicine environments.

Given the complexity and interdisciplinary nature of the health-care domain, the paper targets a wide audience, including medical staff, medical information scientists, bioinformaticians, data scientists, decision-makers, health-care IT professionals and regulation bodies. A bi-directional, mutual learning paradigm between clinical-end users and data scientists will facilitate the development of justified, validated and accepted clinical decision-support tools. This is highlighted in the clinical end-to-end use-case scenario developed in the paper. Broadly covering the steps from a patient's admittance to hospital, data collection and analysis, knowledge generation and patient similarity models will assist in the improvement of patient-specific interpretation of medical data up to the generation of evidence-based, patient-specific treatment plans.

3.7. Conclusion

The rapid digitization of health data offers a unique opportunity to build a robust data-driven healthcare ecosystem. Clinical decision-making, traditionally guided by expert knowledge and the accumulation of personal experience, may now also be informed by accumulated knowledge emerging from the digitization of individual medical data and the creation of population health databases that surround the patient. This ecosystem consists of an integrated data platform that collects personal medical information and medical devices, databases that collect the accumulated knowledge and the professional experience from the individual medical treatment of each patient and the accumulation of medical data of the patient's population, and a set of data-driven applications that aim to provide evidence that supports clinical decision-making. Regulatory and scientific policy framework problems should be addressed before this vision is realized. On its own, increasing the quantity of data that can be collected about an individual patient, by collecting personal medical data from medical devices and electronic health records, as well as combining it with big data analytics, can monitor the resource consumption and health outcome of a patient in an unprecedented resolution and dimension. While medical doctors are treating a single patient, their decisions are informed by considerations that emerge from treating many other individual patients, and the accumulation of problems encountered in the treatment of many patients.

3.7.1. Future Trends

This section has several facets of future trends that will change various areas of electronic health and medical records. With increasing prevalence of digital health data, the use of blockchain for secure, transparent and encrypted data transfer will become the expectation. Eventually, Electronic Health Records (EHRs) will acquire the ability to not only store records of patient visits, but also provide useful diagnostic advice. As this transformation takes place, EHRs will be helmed as diagnostic aids by physicians for writing prescriptions and recommendations, as seamless tools that alert doctors to abnormal health markers and nudges doctors toward further diagnostic tests, as well as a resource for exploring prognoses generally provided by the assistant in hundreds of types to help clue providers into aspects of patient care that might improve outcomes. Large multi-national EHR distributors already building predictive tools into their records include Roche–Flatiron. Rapid adoption and transparency in how the various models establish their predictions will determine whether music assistants end up ruling the healthcare system, or serving as unethical crutches for care providers to manipulate patients against best practices. To avoid discrimination, any AI-based clinical decision-making support will need to be auditable from physician-time and provider-time creation of algorithmically-derived recommendations. As the resolution of this conflict unfolds, patients or relevant power of attorney will increasingly demand permanent access to models fabricated by Machine Learning (ML) to guard against malpractice, essentially bounds on when then models were developed and the data they were developed from. The timeline from blanket privacy guarantees for models, to patients insisting upon free transfer of secondary decision support devices rigged with heightened early-warning performance based on models previously acting on their visit activity will be much faster. Furthermore, designing the forthcoming generation of electronic health or medical records as both intuitive software utilizable by physicians during patient visits and also data logging mechanisms for insurance and litigation will be an around-the-clock task requiring extensive input from multiple entities.

References

Google Health. (2024). AI identifies early signs of lung cancer in CT scans. Nature Medicine.

IBM Watson Health. (2024). AI matches cancer patients with targeted therapies. Journal of Clinical Oncology.

PathAI. (2024). AI-powered digital pathology for rapid cancer detection. The Lancet Digital Health.

Johns Hopkins University. (2024). Digital twin technology simulates heart treatments. Journal of the American College of Cardiology.