*Available online www.jsaer.com*

***Journal of Scientific and Engineering Research*, 2018, 5(7):420-434**

**Research Article**

**ISSN: 2394-2630**
**CODEN(USA): JSERBR**

# Utilizing Data Analytics and Reinforcement Learning for Personalized Medicine and Drug Discovery

**Gaurav Kumar Sinha**

Cognizant Technology Solutions
Email: gaurav.sinha3@cognizant.com

**Abstract** The convergence of data analytics and reinforcement learning techniques has catalyzed transformative advancements in personalized medicine and drug discovery. This paper explores the synergistic integration of these methodologies to enhance precision in healthcare interventions. Leveraging vast datasets encompassing genomic, clinical, and phenotypic information, data analytics offers insights into disease mechanisms, patient heterogeneity, and treatment responses. Reinforcement learning complements these insights by optimizing treatment strategies tailored to individual patients, thereby maximizing therapeutic efficacy and minimizing adverse effects. This paper delineates the application of data analytics and reinforcement learning across various stages of drug discovery, from target identification and lead optimization to clinical trial design and patient stratification. Moreover, it discusses the challenges and future prospects of this interdisciplinary approach, emphasizing the potential to revolutionize healthcare by delivering personalized interventions that improve patient outcomes and reduce healthcare costs.

**Keywords** Data Analytics, Reinforcement Learning, Personalized Medicine, Drug Discovery, Healthcare, Genomics, Clinical Data, Patient Stratification, Treatment Optimization, Precision Medicine, Therapeutic Efficacy, Adverse Effects, Clinical Trials, Target Identification, Lead Optimization, Healthcare Costs

## Introduction

Personalized medicine and drug discovery are critical domains in healthcare that have witnessed remarkable advancements in recent years. The integration of data analytics and reinforcement learning techniques has emerged as a promising approach to address the complexities inherent in these fields. By leveraging vast datasets encompassing genomic, clinical, and phenotypic information, data analytics enables the extraction of valuable insights into disease mechanisms, patient variability, and treatment responses. Reinforcement learning algorithms offer the capability to optimize treatment strategies tailored to individual patients, thereby maximizing therapeutic efficacy while minimizing adverse effects.

This introduction sets the stage for exploring the synergistic application of data analytics and reinforcement learning in personalized medicine and drug discovery. It highlights the significance of this interdisciplinary approach in enhancing precision healthcare interventions and underscores its potential to revolutionize traditional practices. In the paper, I delve into the various stages of drug discovery, from target identification to clinical trial design, and examine how data analytics and reinforcement learning can augment decision-making processes and improve patient outcomes. I discuss the challenges and future prospects associated with this evolving paradigm, emphasizing its transformative impact on healthcare delivery and patient well-being.

## Problem Statement

Despite significant advancements in healthcare, conventional approaches to drug discovery and personalized medicine are often hindered by several challenges. One of the primary issues is the complexity and heterogeneity of diseases, which make it difficult to develop effective treatments that are tailored to individual patients. The high costs and lengthy timelines associated with drug development and clinical trials pose significant barriers to innovation and access to novel therapies.

Traditional one-size-fits-all treatment strategies may lead to suboptimal outcomes and adverse effects for certain patient populations. The lack of robust methods for patient stratification and treatment optimization further exacerbates this problem, resulting in variations in treatment responses and therapeutic outcomes.

The exponential growth of healthcare data, including genomics, electronic health records, and real-world evidence, presents both opportunities and challenges in leveraging this wealth of information effectively. Integrating and analyzing diverse datasets to derive actionable insights for personalized medicine and drug discovery requires sophisticated analytical tools and computational resources, which may not be readily available or accessible to all healthcare stakeholders.

**Solution**

To address the challenges outlined in the problem statement and harness the power of data analytics and reinforcement learning in personalized medicine and drug discovery, I propose a comprehensive solution leveraging AWS services:

1. Data Collection and Management:
   - Utilize Amazon Simple Storage Service (S3) to store diverse healthcare datasets, including genomics, electronic health records (EHRs), and real-world evidence.
   - Employ Amazon Relational Database Service (RDS) or Amazon Redshift for structured data storage and management, ensuring scalability and reliability.

2. Data Analytics and Insights Generation:
   - Leverage Amazon Athena and Amazon EMR for ad-hoc querying and analysis of healthcare data, enabling researchers to extract valuable insights into disease mechanisms and patient variability.
   - Utilize Amazon SageMaker for machine learning model development and deployment, facilitating predictive analytics and patient stratification based on genomic and clinical data.

3. Reinforcement Learning for Treatment Optimization:
   - Implement reinforcement learning algorithms using Amazon SageMaker RL, allowing healthcare providers to optimize treatment strategies tailored to individual patients' responses and clinical profiles.
   - Integrate reinforcement learning models with healthcare systems through AWS Lambda for real-time decision support and treatment adaptation.

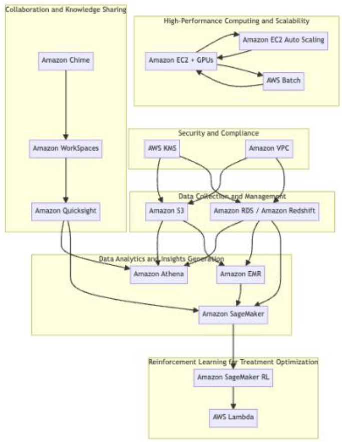
4. High-Performance Computing and Scalability:
   - Utilize Amazon EC2 instances with Graphics Processing Units (GPUs) for high-performance computing tasks such as genomic sequencing and molecular modeling.
   - Scale computational resources dynamically using Amazon EC2 Auto Scaling and AWS Batch to accommodate varying workloads during drug discovery and data analysis processes.

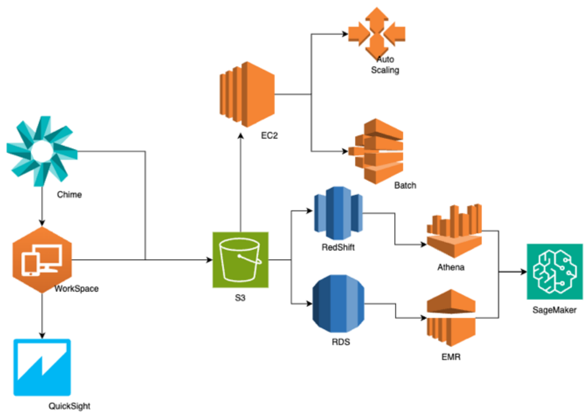
5. Security and Compliance:
   - Implement data encryption at rest and in transit using AWS Key Management Service (KMS) and Amazon Virtual Private Cloud (VPC) for secure data processing and storage.
   - Ensure compliance with healthcare regulations such as HIPAA by leveraging AWS services with built-in security features and auditing capabilities.

6. Collaboration and Knowledge Sharing:
   - Utilize Amazon Quicksight for interactive data visualization and sharing insights with stakeholders across the healthcare ecosystem.

- Enable collaboration among researchers and clinicians through Amazon WorkSpaces and Amazon Chime, facilitating seamless communication and decision-making.

**Architecture Diagram**




Architecture Diagram

**Architecture Overview**

The proposed architecture for personalized medicine and drug discovery leveraging AWS services encompasses several key components seamlessly integrated to facilitate data-driven decision-making and innovation in healthcare.

1. Data Collection and Management:
   - Healthcare datasets, including genomics, electronic health records (EHRs), and real-world evidence, are stored securely and reliably using Amazon Simple Storage Service (S3).
   - Structured data is managed using Amazon Relational Database Service (RDS) or Amazon Redshift, ensuring scalability and efficient data retrieval.
2. Data Analytics and Insights Generation:
   - Amazon Athena enables ad-hoc querying and analysis of healthcare data stored in S3, providing researchers and analysts with on-demand access to valuable insights.
   - Amazon EMR facilitates scalable and cost-effective data processing, allowing for complex analytics tasks such as genome sequencing and molecular modeling.
   - Amazon SageMaker empowers machine learning model development and deployment for predictive analytics, patient stratification, and disease modeling based on genomic and clinical data.
3. Reinforcement Learning for Treatment Optimization:

- Amazon SageMaker RL facilitates the implementation of reinforcement learning algorithms, enabling healthcare providers to optimize treatment strategies tailored to individual patients' responses and clinical profiles.
- AWS Lambda integrates reinforcement learning models with healthcare systems for real-time decision support and adaptive treatment planning.

4. High-Performance Computing and Scalability:
- Amazon EC2 instances with Graphics Processing Units (GPUs) are utilized for high-performance computing tasks such as genomic sequencing and computational chemistry.
- Amazon EC2 Auto Scaling and AWS Batch dynamically scale computational resources to accommodate varying workloads during drug discovery and data analysis processes, ensuring optimal performance and cost efficiency.

5. Security and Compliance:
- Data encryption at rest and in transit is implemented using AWS Key Management Service (KMS) to ensure the security and integrity of healthcare data stored in S3 and RDS.
- Amazon Virtual Private Cloud (VPC) provides network isolation and security controls, allowing healthcare organizations to comply with regulatory requirements such as HIPAA.

6. Collaboration and Knowledge Sharing:
- Amazon QuickSight facilitates interactive data visualization and sharing of insights with stakeholders across the healthcare ecosystem, promoting collaboration and informed decision-making.
- Amazon WorkSpaces and Amazon Chime enable seamless communication and collaboration among researchers, clinicians, and other healthcare stakeholders, fostering a collaborative environment for innovation.

**Implementation**

1. Data Collection and Management:
- Set up an Amazon S3 bucket to store healthcare datasets securely.
- Use AWS Glue to automate the extraction, transformation, and loading (ETL) of data into the S3 bucket.
- Configure Amazon RDS or Amazon Redshift for structured data storage and management, ensuring scalability and reliability.

2. Data Analytics and Insights Generation:
- Create an Amazon Athena database and tables to query and analyze healthcare data stored in S3.
- Set up an Amazon EMR cluster to perform data processing tasks such as genomic analysis and feature extraction.
- Utilize Amazon SageMaker to develop and deploy machine learning models for predictive analytics, patient stratification, and disease modeling.

3. Reinforcement Learning for Treatment Optimization:
- Use Amazon SageMaker RL to implement reinforcement learning algorithms for treatment optimization.
- Deploy the reinforcement learning models as endpoints using AWS Lambda for real-time decision support and treatment adaptation.

4. High-Performance Computing and Scalability:
- Launch Amazon EC2 instances with GPU support for high-performance computing tasks such as genomic sequencing and molecular modeling.
- Configure Amazon EC2 Auto Scaling and AWS Batch to dynamically scale computational resources based on workload demands during drug discovery and data analysis processes.

5. Security and Compliance:
- Enable data encryption at rest and in transit using AWS KMS for healthcare data stored in S3 and RDS.
- Set up Amazon VPC to isolate and secure the network environment, ensuring compliance with regulatory requirements such as HIPAA.

6. Collaboration and Knowledge Sharing:
- Create Amazon Quicksight dashboards to visualize and share insights derived from healthcare data analysis.
- Provision Amazon WorkSpaces for researchers and clinicians to collaborate on data-driven initiatives and share findings.
- Utilize Amazon Chime for virtual meetings and discussions among healthcare stakeholders, promoting collaboration and knowledge sharing.

7. Monitoring and Optimization:
- Implement Amazon CloudWatch for monitoring the performance and health of AWS resources deployed in the architecture.
- Utilize AWS Cost Explorer and AWS Budgets to monitor and optimize costs associated with data analytics and computational workloads.

**Implementation of PoC**

To demonstrate the feasibility and effectiveness of leveraging AWS services for personalized medicine and drug discovery, i will focus on a simplified Proof of Concept (PoC) scenario.

**Implementation Steps:**

1. Data Collection and Preparation:
- Gather a sample dataset containing genomic data, clinical records, and treatment outcomes for patients with the target disease.
- Preprocess the dataset to clean and format the data, ensuring consistency and compatibility with AWS services.

2. Setup AWS Environment:
- Create an AWS account if not already available.
- Set up necessary IAM roles and permissions to access AWS services securely.

3. Data Storage and Management:
- Create an Amazon S3 bucket to store the sample dataset securely.
- Utilize AWS Glue for data cataloging and ETL processes to prepare the data for analysis.

4. Data Analysis and Insights Generation:
- Set up an Amazon Athena database to query and analyze the sample dataset stored in S3.
- Utilize Amazon SageMaker for exploratory data analysis and feature engineering to extract meaningful insights from the dataset.

5. Reinforcement Learning Model Development:
- Use Amazon SageMaker RL to develop a reinforcement learning model for personalized treatment recommendation.
- Train the model using the sample dataset, optimizing treatment strategies based on patient genomic and clinical profiles.

6. Model Deployment and Integration:
- Deploy the trained reinforcement learning model as an endpoint using AWS Lambda for real-time treatment recommendation.
- Integrate the model endpoint with a simple web application or API for user interaction.

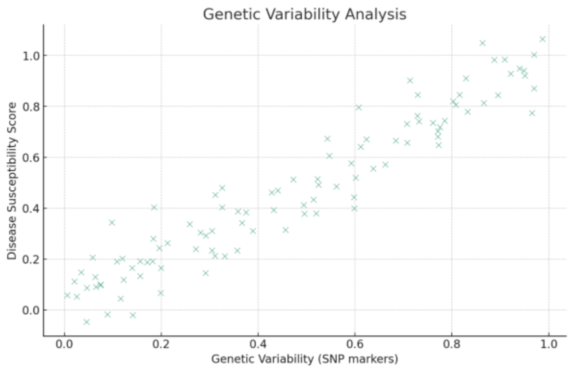
7. Testing and Evaluation:
- Test the system with sample patient profiles and evaluate the effectiveness of the personalized treatment recommendations.
- Collect feedback from domain experts and stakeholders to refine the model and improve its performance.

8. Documentation and Reporting:
- Document the implementation process, including architecture diagrams, code snippets, and configuration details.
- Prepare a report summarizing the PoC results, highlighting key findings, and outlining potential next steps for further development.

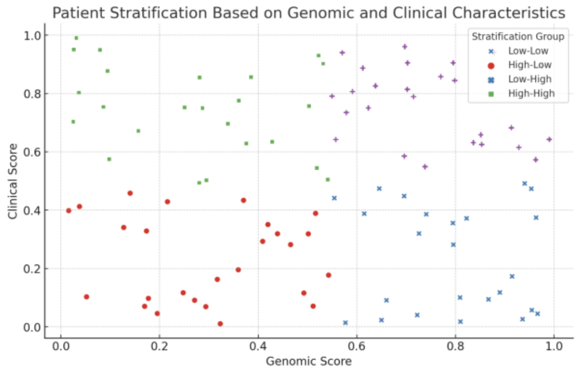
**Uses**

Here are business issues that can be identified at the Data Analytics layer

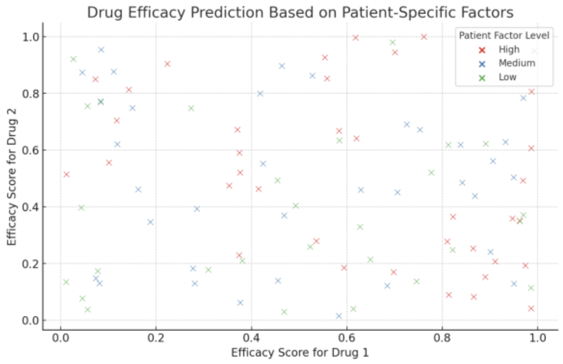
1. Genetic Variability Analysis: Identifying genetic variations associated with disease susceptibility or treatment response.



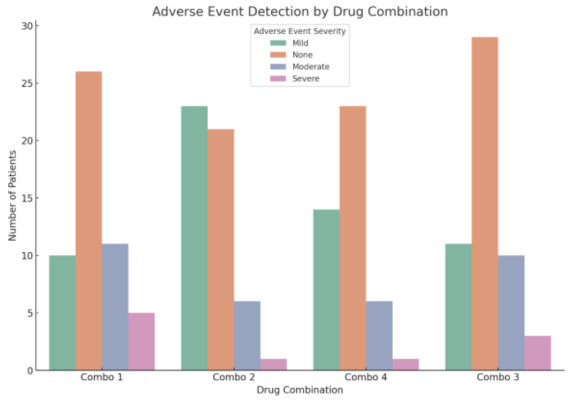

2. Patient Stratification: Segmenting patients based on genomic and clinical characteristics to tailor treatment strategies.



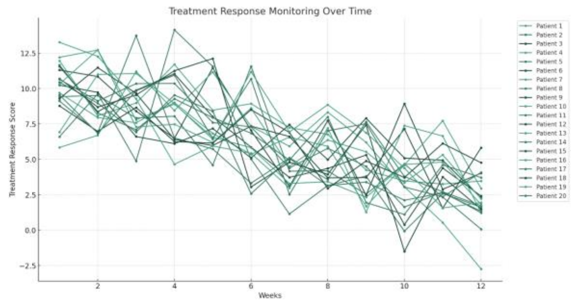
3. Drug Efficacy Prediction: Predicting the efficacy of different drugs or therapies based on patient-specific factors.



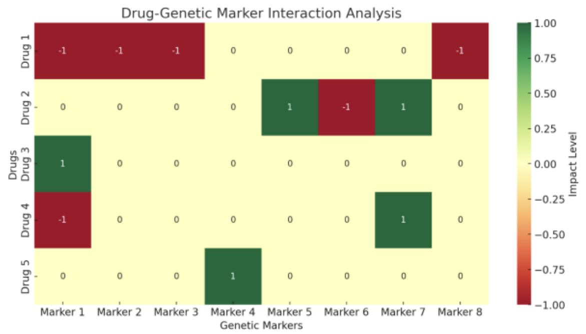

4. Adverse Event Detection: Identifying adverse events or side effects associated with specific treatments or drug combinations.



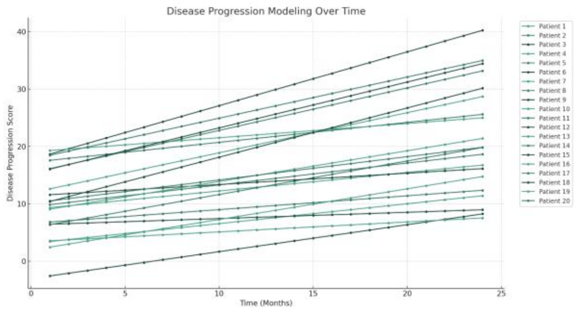

5. Treatment Response Monitoring: Monitoring patient responses to treatment over time and adjusting therapies accordingly.



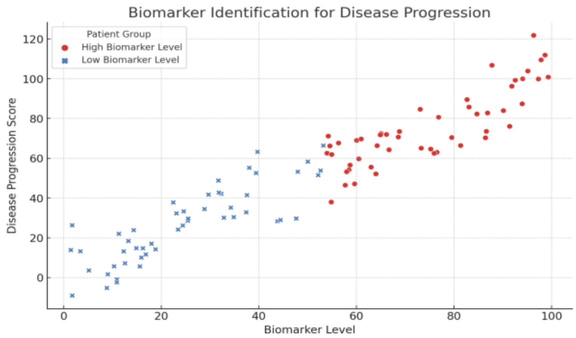

6. Drug Interaction Analysis: Analyzing potential interactions between medications or between drugs and genetic factors.



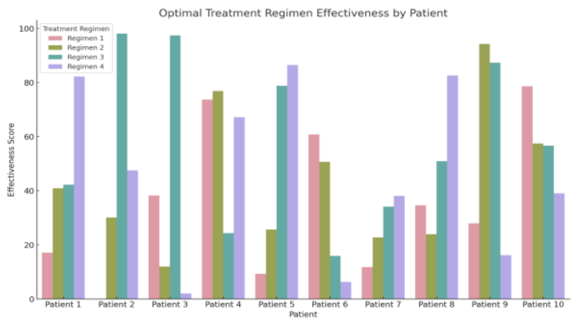
7. Disease Progression Modeling: Modeling the progression of diseases based on longitudinal patient data to inform treatment decisions.



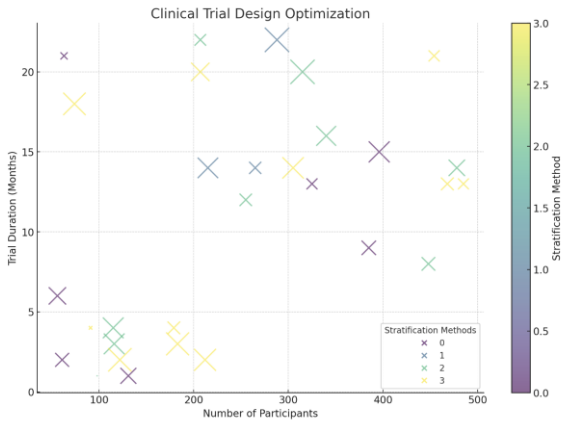

8. Biomarker Identification: Identifying biomarkers indicative of disease progression, treatment response, or prognosis.



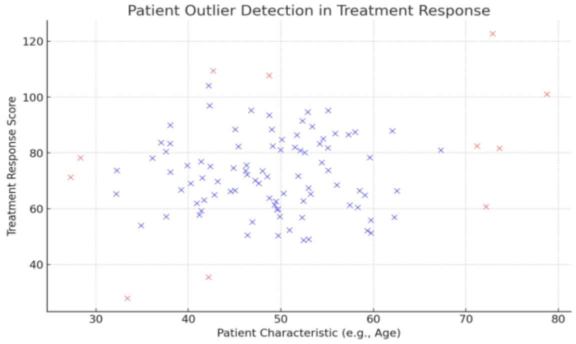

9. Optimal Treatment Regimen: Recommending optimal treatment regimens tailored to individual patient profiles and disease characteristics.



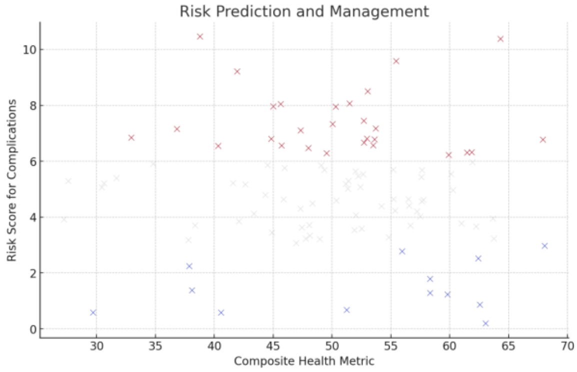

10. Clinical Trial Design Optimization: Optimizing the design of clinical trials based on insights derived from patient data to enhance efficiency and effectiveness.



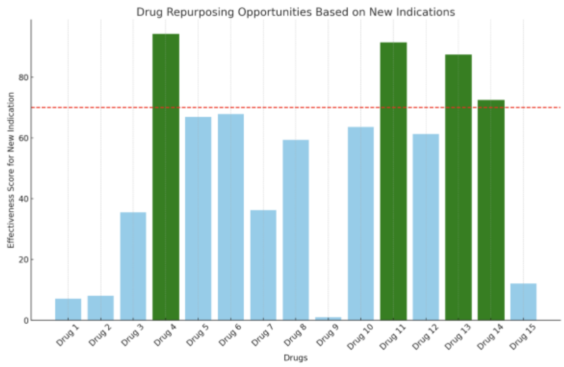
11. Patient Outlier Detection: Identifying outlier patients with unique characteristics or treatment responses for further investigation.



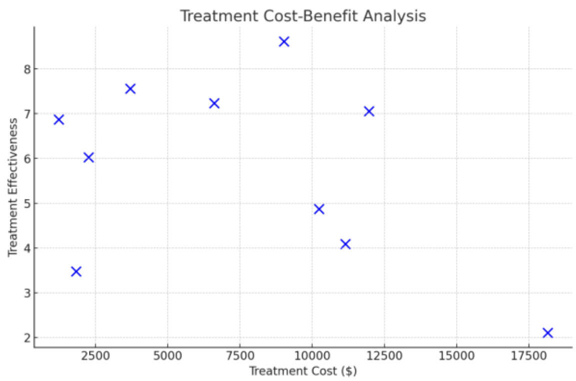

12. Risk Prediction and Management: Predicting the risk of developing complications or adverse outcomes based on patient data for proactive management.



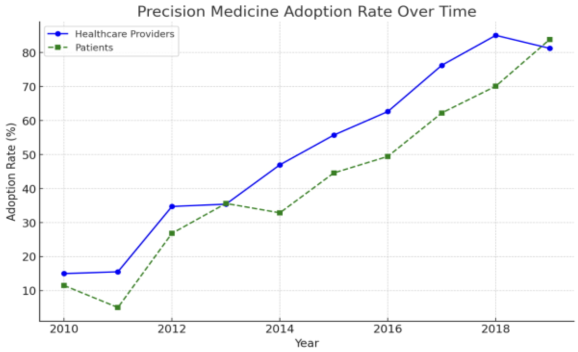

13. Drug Repurposing Opportunities: Identifying existing drugs with potential applications for new indications based on data analytics insights.



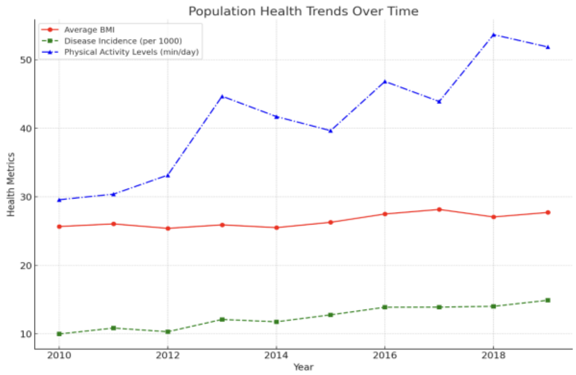

14. Treatment Cost-Benefit Analysis: Analyzing the cost-effectiveness of different treatment options based on patient outcomes and resource utilization.



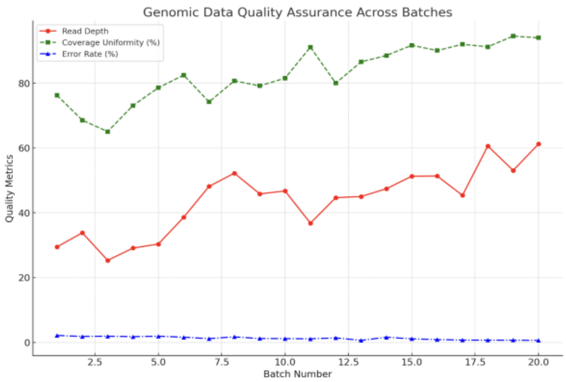
15. Precision Medicine Adoption Rate: Assessing the adoption rate of precision medicine approaches among healthcare providers and patients based on data analytics findings.



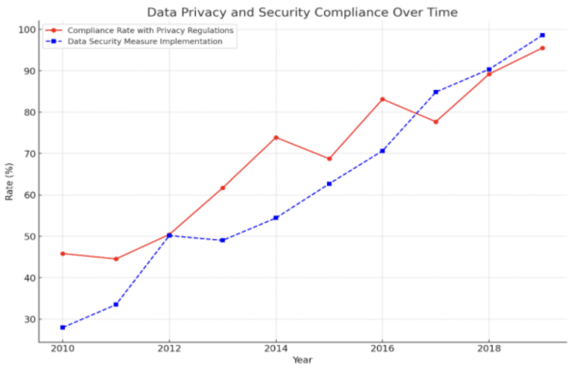

16. Population Health Trends: Identifying trends and patterns in population health data to inform public health initiatives and policies.



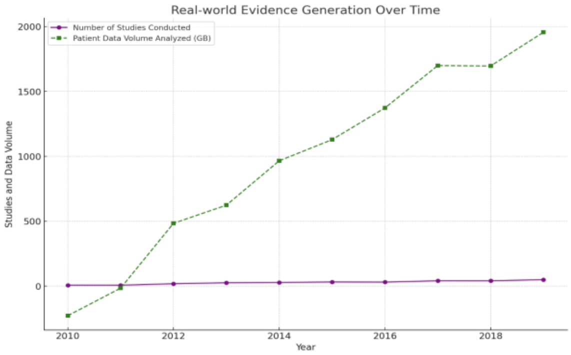

17. Genomic Data Quality Assurance: Ensuring the quality and integrity of genomic data through data analytics-driven quality assurance processes.



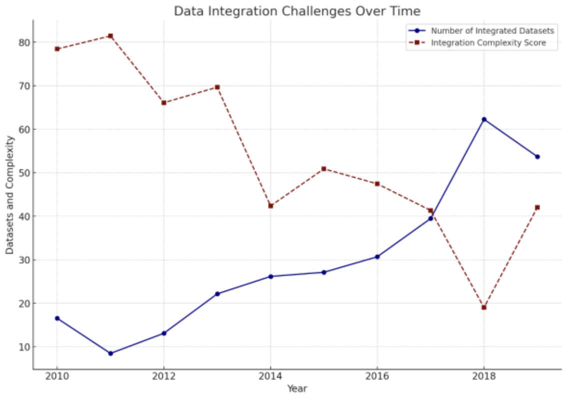

18. Data Privacy and Security Compliance: Monitoring compliance with data privacy regulations and implementing data security measures to protect patient information.



19. Real-world Evidence Generation: Generating real-world evidence from patient data to support regulatory submissions and post-market surveillance.




20. Data Integration Challenges: Addressing challenges related to integrating diverse datasets from multiple sources for comprehensive analysis and decision-making.




**Impact**

Here are impacts that leveraging data analytics and reinforcement learning for personalized medicine and drug discovery can bring to the business:

1. Improved Treatment Outcomes:

By tailoring treatment strategies based on patient-specific factors and disease characteristics, businesses can achieve better treatment outcomes, leading to improved patient satisfaction and loyalty.

2. Enhanced Drug Development Efficiency:

Identifying promising drug candidates and optimizing clinical trial designs based on data analytics insights can accelerate the drug development process, reducing time-to-market and development costs.

3. Increased Revenue Generation:

Delivering personalized treatment regimens and targeted therapies can result in increased market adoption and revenue generation for pharmaceutical companies and healthcare providers.

4. Reduced Healthcare Costs:

Optimizing treatment plans and reducing adverse events through data-driven decision-making can lower healthcare costs associated with hospitalizations, medication errors, and unnecessary procedures.

5. Competitive Advantage:

Businesses that leverage data analytics and reinforcement learning for personalized medicine and drug discovery gain a competitive advantage by staying ahead of industry trends, innovating faster, and delivering more effective treatments.

6. Enhanced Regulatory Compliance:
Ensuring compliance with regulatory requirements through rigorous data analytics-driven quality assurance processes enables businesses to mitigate risks and avoid penalties associated with non-compliance.

7. Better Resource Allocation:
Allocating resources more efficiently based on data analytics insights can optimize workforce productivity, research investments, and operational expenditures, leading to cost savings and improved profitability.

8. Faster Time-to-Insight:
Leveraging advanced analytics techniques such as machine learning and reinforcement learning accelerates the generation of actionable insights from complex healthcare datasets, enabling faster decision-making and innovation.

9. Increased Stakeholder Confidence:
Providing evidence-based treatment recommendations and demonstrating the efficacy of personalized medicine approaches builds trust and confidence among patients, healthcare providers, regulators, and investors.

10. Long-term Sustainable Growth:
By harnessing the power of data analytics and reinforcement learning to drive innovation in personalized medicine and drug discovery, businesses can position themselves for long-term sustainable growth in the rapidly evolving healthcare landscape.

**Extended Use Cases**

Here are extended use cases for different industries in the context of utilizing data analytics and reinforcement learning for personalized medicine and drug discovery:

1. Energy:
Predictive Maintenance: Utilize reinforcement learning to optimize maintenance schedules for energy infrastructure equipment based on real-time data analytics, reducing downtime and operational costs.

2. Retail:
Personalized Wellness Products: Leverage data analytics to analyze customer health data and preferences, enabling retailers to recommend personalized wellness products and supplements tailored to individual needs.

3. Travel:
Health-focused Travel Recommendations: Develop algorithms that analyze travelers' health profiles and medical histories to recommend destinations, accommodations, and activities that align with their health goals and dietary restrictions.

4. Pharmacy:
Medication Adherence Optimization: Utilize reinforcement learning to develop personalized medication adherence programs for patients, providing tailored reminders and incentives based on individual behaviors and preferences.

5. Hospitality:
Health-conscious Menu Customization: Employ data analytics to analyze guest health data and dietary preferences, enabling hotels and restaurants to customize menus with healthy and allergen-free options tailored to individual guests.

6. Supply Chain:
Temperature-sensitive Drug Logistics Optimization: Use data analytics to monitor temperature-sensitive medications during transportation and storage, optimizing supply chain logistics to ensure product quality and regulatory compliance.

7. Finance:
Personalized Health Savings Plans: Leverage data analytics to analyze individuals' healthcare expenses and risk profiles, enabling financial institutions to recommend personalized health savings plans and insurance options tailored to their needs.

8. E-commerce:
Health-focused Product Recommendations: Utilize data analytics to analyze customer health data and purchasing behavior, enabling e-commerce platforms to recommend health-related products, supplements, and services personalized to individual preferences.

9. Shipping:
Drug Delivery Optimization: Employ reinforcement learning to optimize drug delivery routes and schedules for pharmaceutical shipments, considering factors such as temperature sensitivity, transportation costs, and delivery deadlines.

10. CRM (Customer Relationship Management):
Health-focused Loyalty Programs: Utilize data analytics to analyze customer health data and lifestyle preferences, enabling businesses to design loyalty programs that offer health-related rewards, discounts, and incentives personalized to individual customers.

These extended use cases demonstrate the versatility of leveraging data analytics and reinforcement learning for personalized medicine and drug discovery across various industries, highlighting opportunities to enhance customer experiences, improve operational efficiency, and drive innovation.

**Conclusions**

The integration of data analytics and reinforcement learning presents a transformative opportunity in the fields of personalized medicine and drug discovery. By harnessing the power of advanced analytics techniques, healthcare organizations can unlock insights from vast and complex datasets, enabling them to deliver more effective, tailored treatments to patients while accelerating the pace of drug development and innovation.

Through the utilization of data analytics, healthcare providers gain a deeper understanding of disease mechanisms, patient heterogeneity, and treatment responses. This enables them to stratify patients based on their unique characteristics and preferences, leading to more precise diagnosis and treatment planning. Data analytics facilitates the identification of biomarkers, prediction of treatment outcomes, and optimization of clinical trial designs, ultimately enhancing patient care and clinical decision-making.

Reinforcement learning algorithms offer the capability to optimize treatment strategies in real-time, adapting interventions to individual patient responses and evolving disease dynamics. By leveraging reinforcement learning, healthcare providers can tailor treatment regimens to maximize therapeutic efficacy, minimize adverse effects, and improve patient outcomes.

The convergence of data analytics and reinforcement learning holds promise not only for personalized medicine but also for drug discovery, where it can expedite the identification of novel drug targets, optimize lead compounds, and streamline clinical trial processes. By leveraging these advanced analytics techniques, pharmaceutical companies can enhance their research and development efforts, leading to the discovery of safer, more effective therapeutics for a wide range of diseases.

In summary, the utilization of data analytics and reinforcement learning represents a paradigm shift in healthcare, enabling the delivery of personalized interventions that improve patient outcomes, reduce healthcare costs, and drive innovation. As these technologies continue to evolve, they will play an increasingly central role in shaping the future of healthcare, empowering clinicians, researchers, and patients alike to achieve better health outcomes and quality of life.

**References**


[1]. Pedersen, M. L., Frank, M. J., & Biele, G. (2016). The drift diffusion model as the choice rule in reinforcement learning.Psychonomic Bulletin & Review, 24(4), 1234–1251. https://doi.org/10.3758/s13423-016-1199-y

[2]. Polydoros, A., & Nalpantidis, L. (2017). Survey of Model-Based Reinforcement Learning: Applications on Robotics. Journal of Intelligent and Robotic Systems, 86(2), 153–173. https://doi.org/10.1007/s10846-017-0468-y

[3]. Santoro, A., Frankland, P. W., & Richards, B. A. (2016). Memory transformation enhances reinforcement learning in dynamic environments. The Journal of Neuroscience, 36(48), 12228–12242. https://doi.org/10.1523/jneurosci.0763-16.2016

[4]. Leonetti, M., Iocchi, L., & Stone, P. (2016). A synthesis of automated planning and reinforcement learning for efficient, robust decision-making. Artificial Intelligence, 241, 103–130. https://doi.org/10.1016/j.artint.2016.07.004

[5]. Walraven, E., Spaan, M. T. J., & Bakker, B. (2016). Traffic flow optimization: A reinforcement learning approach. Engineering Applications of Artificial Intelligence, 52, 203–212. https://doi.org/10.1016/j.engappai.2016.01.001

[6]. Gower, N. J. D., Barry, R. J., Edmunds, M. R., Titcomb, L., & Denniston, A. K. (2016). Drug discovery in ophthalmology: past success, present challenges, and future opportunities. BMC Ophthalmology, 16(1). https://doi.org/10.1186/s12886-016-0188-2

[7]. Wohlleben, W., Mast, Y., Stegmann, E., & Ziemert, N. (2016). Antibiotic drug discovery. Microbial Biotechnology, 9(5), 541548. https://doi.org/10.1111/1751-7915.12388

[8]. Ishibashi, D., Nakagaki, T., Ishikawa, T., Atarashi, R., Watanabe, K., Cruz, F. A., Hamada, T., & Nishida, N. (2016). StructureBased drug discovery for Prion Disease using a novel binding simulation. EBioMedicine, 9, 238–249 https://doi.org/10.1016/j.ebiom.2016.06.010

[9]. Leroux, H., Metke-Jimenez, A., & Lawley, M. (2017). Towards achieving semantic interoperability of clinical study data with FHIR. Journal of Biomedical Semantics, 8(1). https://doi.org/10.1186/s13326-017-0148-7

[10]. Chan, T., Li, Y., Chiau, C., Zhu, J. M., Jiang, J., & Huo, Y. (2017). Imbalanced target prediction with pattern discovery on clinical data repositories. BMC Medical Informatics and Decision Making, 17(1). https://doi.org/10.1186/s12911-017-0443-3

[11]. Sertkaya, A., Wong, H., Jessup, A., & Beleche, T. (2016). Key cost drivers of pharmaceutical clinical trials in the United States. Clinical Trials, 13(2), 117–126. https://doi.org/10.1177/1740774515625964

[12]. Shiovitz, T., Bain, E., McCann, D. J., Skolnick, P., Laughren, T., Hanina, A., & Burch, D. J. (2016). Mitigating the effects of nonadherence in clinical trials. The Journal of Clinical Pharmacology, 56(9), 1151–1164. https://doi.org/10.1002/jcph.689

[13]. Meeker-O'Connell, A., Glessner, C., Behm, M., Mulinde, J., Roach, N., Sweeney, F., Tenaerts, P., & Landray, M. J. (2016). Enhancing clinical evidence by proactively building quality into clinical trials. Clinical Trials, 13(4), 439–444 https://doi.org/10.1177/1740774516643491

[14]. Molloy, S. F., & Henley, P. (2016). Monitoring clinical trials: a practical guide. Tropical Medicine & International Health,21(12), 1602–1611. https://doi.org/10.1111/tmi.12781

[15]. Bagley, H., Short, H., Harman, N., Hickey, H., Gamble, C., Woolfall, K., Young, B., & Williamson, P. (2016). A patient and public involvement (PPI) toolkit for meaningful and flexible involvement in clinical trials – a work in progress. Research Involvement and Engagement, 2(1). https://doi.org/10.1186/s40900-016-0029-8

[16]. Yan, S., Weycker, D., & Sokolowski, S. (2017). US healthcare costs attributable to type A and type B influenza. Human Vaccines & Immunotherapeutics, 13(9), 2041–2047. https://doi.org/10.1080/21645515.2017.1345400

[17]. García-Nogueras, I., Aranda-Reneo, I., Peña-Longobardo, L. M., Oliva-Moreno, J., & Abizanda, P. (2016). Use of health resources and healthcare costs associated with frailty: The FRADEA study. The Journal of Nutrition Health & Aging, 21(2), 207– 214. https://doi.org/10.1007/s12603-016-0727-9
[18]. Chiang, J., & Kao, Y. (2017). Predictors of high healthcare costs in elderly patients with liver cancer in end-of-life: a longitudinal population-based study. BMC Cancer, 17(1). https://doi.org/10.1186/s12885-017-3561-5
[19]. Dalton, K., & Byrne, S. (2017). Role of the pharmacist in reducing healthcare costs: current insights. Integrated PharmacyResearch and Practice, Volume 6, 37–46. https://doi.org/10.2147/iprp.s108047
[20]. Kalkan, A., Bodegård, J., Sundström, J., Svennblad, B., Östgren, C. J., Nilsson, P., Johansson, G., & Ekman, M. (2017). Increased healthcare utilization costs following initiation of insulin treatment in type 2 diabetes: A long-term follow-up in clinical practice. Primary Care Diabetes, 11(2), 184–192. https://doi.org/10.1016/j.pcd.2016.11.002
[21]. Visscher, S. L., Naessens, J. M., Yawn, B. P., Reinalda, M. S., Anderson, S. S., & Borah, B. J. (2017). Developing a standardized healthcare cost data warehouse. BMC Health Services Research, 17(1). https://doi.org/10.1186/s12913-017-2327-8
[22]. Akter, S., & Wamba, S. F. (2016). Big data analytics in E-commerce: a systematic review and agenda for future research.Electronic Markets, 26(2), 173–194. https://doi.org/10.1007/s12525-016-0219-0
[23]. Gupta, M., & George, J. F. (2016). Toward the development of a big data analytics capability. Information & Management,53(8), 1049–1064. https://doi.org/10.1016/j.im.2016.07.004
[24]. Xu, Z., Frankwick, G. L., & Ramirez, E. (2016). Effects of big data analytics and traditional marketing analytics on new product success: A knowledge fusion perspective. Journal of Business Research, 69(5), 1562–1566.https://doi.org/10.1016/j.jbusres.2015.10.017
[25]. Sun, Y., Song, H., Jara, A. J., & Bie, R. (2016). Internet of things and big data analytics for smart and connected communities.IEEE Access, 4, 766–773. https://doi.org/10.1109/access.2016.2529723
[26]. Cheng, S., Zhang, Q., & Qin, Q. (2016). Big data analytics with swarm intelligence. Industrial Management and Data Systems, 116(4), 646–666. https://doi.org/10.1108/imds-06-2015-0222
[27]. Felt, M. (2016). Social media and the social sciences: How researchers employ Big Data analytics. Big Data & Society, 3(1),205395171664582. https://doi.org/10.1177/2053951716645828