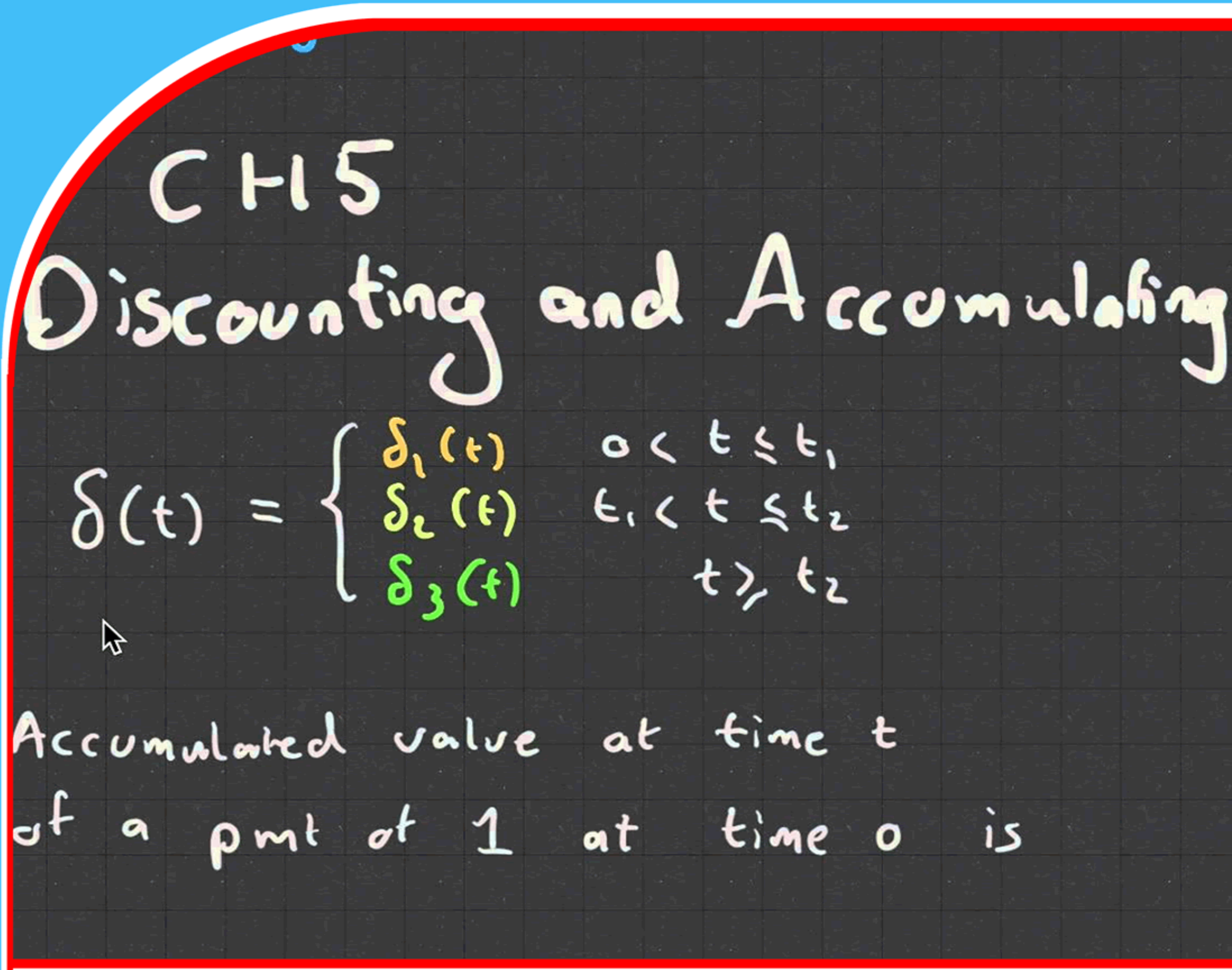
American Journal of
Statistics and Actuarial Science
(AJSAS)



Discovering Disease Biomarkers in Metabolomics via Big Data Analytics

Mia Md Tofayel Gonee Manik, Sadia Islam Nilima, Md Abdullah Al Mahmud, Sadia Sharmin & Rakibul Hasan

AJPO

American Journal of Statistics and Actuarial Science
ISSN 2520-4189 (Online)
Vol.4, Issue 1, pp 35 – 49, 2022
www.ajpojournals.org

Discovering Disease Biomarkers in Metabolomics via Big Data Analytics

Mia Md Tofayel Gonee Manik[1*], Sadia Islam Nilima[3], Md Abdullah Al Mahmud[2], Sadia Sharmin[2] & Rakibul Hasan[3]

[1]Department of Biochemistry and Molecular Biology, University of Dhaka, Dhaka 1000, Bangladesh

[2]Department of Business Administration, International American University, 3440 Wilshire Blvd STE 1000, Los Angeles, CA 90010, United States

[3]Department of Business Administration, National University, Dhaka - Mymensingh Hwy, Gazipur 1704, Bangladesh

Abstract

Purpose: This paper aims to demonstrate the utility of BDA in metabolomics and illness biomarker identification. Additionally, it will describe the metabolomics methodology and procedures, big data for metabolomics, and biomarkers utilizing these strategies. Lastly, it will highlight metabolomics possible future course and advancement as well as its incorporation into the healthcare industry.

Materials and Methods: Metabolomics data is handled via big data, which also helps find illness biomarkers. The research is conducted with simple collective analysis.

Findings: The results have found that metabolomics is a formidable tool, and its implementation has significantly improved our capacity to evaluate disease assumptions.

Implications to Theory, Practice and Policy: In the future, metabolomics and big data can be combined to help resolve challenging medical problems. By combining metabolomics with advanced blockchain, artificial intelligence, and statistical and computational analysis, metabolomics can improve the identification and finding of biomarkers. Metabolomics is a formidable tool for developing and formulating precision with large data analysis, and its integration with other "omics fields" and the addition of new technology and initiatives will bring ideas of personalized medicine to life. As metabolomics continues to evolve, further discoveries can be made in the future, benefiting the understanding of disorders and the treatment of patients.

Keywords: *Big Data Analytics, Biomarker, Metabolomics, NMR, Chromatography*

1.0 INTRODUCTION

Metabolomics, the global metabolic profiling of cells, tissues, or organisms, is now mature enough to join other omics and close the systems biology loop. Small molecule biochemistry, the foundation of medical research, is reemerging from the shadows, bringing technologies of immense power and insight. The study of endogenous small molecules, or metabolites, in cells, tissues, organs, or organisms is known as metabolomics. These byproducts of a variety of cellular processes offer an indicator of the general state of health of a biological organism. They engage in a vital part in researching various chemical processes in living things and determining their origins of illnesses, how medications affect living things, and how living things influence their surroundings (Alonso et al., 2015). The field of metabolomics is broad and applies to various fields, such as agriculture, environmental sciences, and health. Metabolomics is used in medicine to find possible treatment regimens by applying it to biomarkers, diagnostic, and prognostic clinical trials. external metabolomics, on the other hand, focuses on alterations in the organism brought about by interactions with external elements. By determining plant metabolisms, agricultural metabolomics seeks to increase crop yields and control plant diseases (Nagana Gowda & Raftery, 2021). The rapid development of metabolomics technology platforms and the incorporation of metabolomics into research efforts allow for answering key questions that could not be fully addressed by other omics alone. Examples of this principle are accruing at a rapid pace, made possible by recent developments in analytical chemistry, such as high-field nuclear magnetic resonance (NMR) and mass spectrometry (MS) platforms for small molecule separation, detection, and characterization. User-friendly multivariate data analysis software packages are also available to convolute large data matrices generated in metabolomics experiments (Abdelnabi et al., 2022; Schmidt et al., 2021).

Metabolomics is a field that uses various techniques to analyze and determine the representation of the most abundant metabolites in biological samples. These techniques include Nuclear Magnetic Resonance (NMR) spectroscopy, Mass Spectrometry (MS), Liquid Chromatography (LC-MS), and Gas Chromatography (GC-MS). Targeted metabolomics focuses on quantifying specific values in a specific group of metabolites linked to a certain pathway, while untargeted metabolomics aims to determine as many metabolites present in a given sample to provide an overall view of the metabolism (Schmidt et al., 2021). Big data analytics in metabolomics plays a crucial role in data collection, management, data processing, and integration, providing statistical and computational methods essential for effective metabolomics research. Data collection starts with sample acquisition, which typically involves blood, urine, or body tissue samples. The produced samples are further analyzed with huge datasets offering data on hundreds to thousands of metabolites. Primary database data derived from biological samples and repositories of metabolomics data are compared and analyzed by researchers to perform meta-analysis (Tolani et al., 2021). Processing these large and complicated data sets is a real challenge in metabolomics, as the amount of information can be overwhelming and its variety necessitates keeping track of and storing all types of it most conveniently. Data integrity and avoiding significant differences in data format are essential for analysis. Data quality and standardization establish the compatibility of data collection, processing, and reporting schemes since data derived from different sources can be heterogeneous and not directly comparable (Guasch-Ferré et al., 2018).

Data processing and integration involve basic preprocessing steps to make the data manageable after analysis. Proper normalization helps correct for variations in sample concentration and makes comparisons between the samples feasible. Techniques like logarithmic transformation are utilized to modify the variance in data and make it suitable for statistical analysis (Akid, Shah, et al., 2021; Akid, Wasiew, et al., 2021; Das et al., 2020; Datta et al., 2024; J. Rana et al., 2022; M. J. Rana et al., 2022; Rana et al., 2021). The use of metabolomics in conjunction with other global approaches like genomics and proteomics is currently seen as a possibility. Integrating multiple platforms enhances the power of identification and understanding of complex molecular systems and pathological conditions. For example, combining metabolomics with genomic data would help identify the genetic aspect of the metabolic profile and combine proteomics data involving details of the proteins involved (Van Ravenzwaay et al., 2007).

Statistics and computational methods are the basis of metabolomics data analysis. The current study will use Principal Component Analysis (PCA) and Partial Least Squares Discriminant Analysis (PLS-DA) methods to reduce data dimensions and get feature extraction (Perez De Souza et al., 2020). Machine learning or artificial intelligence can also be integrated in analyzing metabolomics data to search for less obvious relations hidden from conventional analysis techniques. In conclusion, metabolomics data analysis is a critical area for research and development. By utilizing computational methods, identifying biomarkers for disease, and integrating them with other types of data, researchers can develop more effective treatments for various diseases (Steuer et al., 2019; Taherdoost, 2022; Tolani et al., 2021).

The goal of this paper is to demonstrate how BDA can aid in the identification of illness biomarkers and metabolomics. It will also go over the metabolomics methodology and procedures, big data metabolomics, and biomarkers that make use of these technologies. Lastly, it will emphasize the possibility of the future course and advancement of metabolomics, along with its incorporation into the healthcare field area where both the researchers of health and people with related health problems can get benefit from the innovations.

2.0 LITERATURE REVIEW

A review explores the potential of metabolomics in studying insect virus infections, a field that has been largely overlooked. Metabolomics has enabled researchers to detect and analyze dynamic changes in host metabolism triggered by viral infections, enhancing our understanding of the complex dynamics of virus-host interactions in insects. The transformative power of metabolomics in this realm is highlighted, and prospects include the integration of multi-omics approaches, the creation of insect-specific metabolite databases, and the emerging field of insect immunometabolism. The review emphasizes the importance of metabolomics in uncovering the subtle yet significant effects of virus infections in insects, calling for greater exploration and recognition in this essential research domain (Taherdoost, 2022). Diabetes, a global health issue, is a growing concern, with complications like cardiovascular and kidney disease. The identification of biomarkers for diabetes screening, diagnosis, and prognosis can help in personalized prevention and treatment. Metabolomics, a technique that can identify and quantify multiple biomarkers simultaneously, can provide information on underlying metabolic pathways and identify diabetes mechanisms. Metabolomics has been applied in epidemiological studies to identify novel biomarkers for type 2 diabetes and its complications, such as branched-chain amino acids, phenylalanine metabolites, energy metabolism metabolites, and lipid metabolism. By integrating metabolomics with other omics data, such as genetics, transcriptomics, proteomics, and clinical data, a comprehensive metabolic network can be presented, facilitating causal inference.

Metabolomics can deepen molecular understanding, identify potential therapeutic targets, and improve the prevention and management of T2D and its complications (Jin & Ma, 2021).

Another paper explores the potential of Big Data Analytics in healthcare, focusing on its application in patient treatment and health management. The research, conducted on 217 Polish medical facilities, reveals that these facilities are increasingly adopting data-based healthcare, using both structured and unstructured data for analytics in administrative, business, and clinical areas. The study confirms that medical facilities are working on both structured and unstructured data sources, including databases, transaction data, unstructured content from emails and documents, and data from devices and sensors. However, the use of social media data is less common, indicating that decisions in medical facilities are highly data-driven. The study confirms the trend towards data-based healthcare and its benefits (Batko & Ślęzak, 2022). Again, big data is revolutionizing biomedical research by integrating data from various sources, including laboratory experiments, clinical investigations, healthcare records, and the Internet of Things. The increasing use of omics technologies, such as genomics, epigenomics, transcriptomics, proteomics, metabolomics, and pharmacogenomics, is enabling future advances in personalized medicine and improved patient care. This data is expected to aid in disease diagnosis and patient care management. The integration of big data from multi-omics technologies, such as meta and co-expression analyses, is crucial for a better understanding of human diseases. Future directions for big data in biomedical research and healthcare include policies, funding, and interdisciplinary collaborations (Cremin et al., 2022; Guasch-Ferré et al., 2018; Nagana Gowda & Raftery, 2021).

Moreover, the blockchain combination and artificial intelligence (AI) are rapidly gaining popularity in the market (Akid, Shah, et al., 2021; Akid, Wasiew, et al., 2021; Das et al., 2020; Datta et al., 2024; Hasan et al., 2022; Hilal et al., 2022; Jakhrani et al., 2012; Nicholls et al., 2021; Rana et al., 2021; Sobuz et al., 2022; Taherdoost, 2022). Blockchains automate instalments, allowing for a decentralized exchange of personal data and information. AI provides robots with the ability to reason and make decisions, while blockchains provide a secure, trustworthy system for Bitcoin payments. The combination of AI and blockchains has led to a critical review of the current situation and rationale of the integration, focusing on practical application and addressing gaps and problems in the literature. Both paradigms offer new opportunities for the digital money industry (Taherdoost, 2022). Medical research has significantly evolved since the publication of the full human genome in 2003, enabling the understanding of the biological functions of various macro and micro molecules. This revolution has been further accelerated by the advent of high-throughput technologies, including genomics, transcriptomics, proteomics, and metabolomics, which collectively form multi-omics technologies. These technologies are used in the biomedical field to diagnose, manage, monitor, treat, and discover novel disease biomarkers. Genotyping arrays and other transcriptomic studies have helped to elucidate gene expression patterns in different biological states, while proteomics and metabolomics predict the role of different biological molecules in an organism. These technologies offer an integrative approach to understanding the information that underlies diseases, such as cancer, cardiovascular diseases, neurodegenerative diseases, and diabetes. By analyzing multi-omics data, researchers can identify the interrelationships of involved biomolecules and their functions, providing valuable insights into disease progression and treatment (Alonso et al., 2015; Bekri, 2016; Cremin et al., 2022).

Metabolomics plays a crucial role in identifying disease biomarkers, which can significantly impact disease diagnosis and treatment. For instance, in cancer research, metabolomics has been used to discover reliable biomarkers for detecting prostate cancer and type 2 diabetes. These biomarkers can be used to design personalized treatment strategies, enhancing patient survival and reducing the impact of diabetes on public health. However, validation is necessary to integrate metabolomics results into clinical practice, ensuring the accuracy of biomarkers for other groups and conditions. Challenges in incorporating biomarkers into daily medicine include regulation approval, cost consideration, and the need for specialized equipment and a skilled workforce. This paper suggests that big data analysis can help in metabolomics and identifying disease biomarkers, highlighting the potential future direction and progression of metabolomics and its integration into the healthcare domain.

3.0 MATERIALS AND METHODS

The research was conducted using a simple collective method where several papers were selected primarily. Then the procedure followed the flow chart of Figure 1:

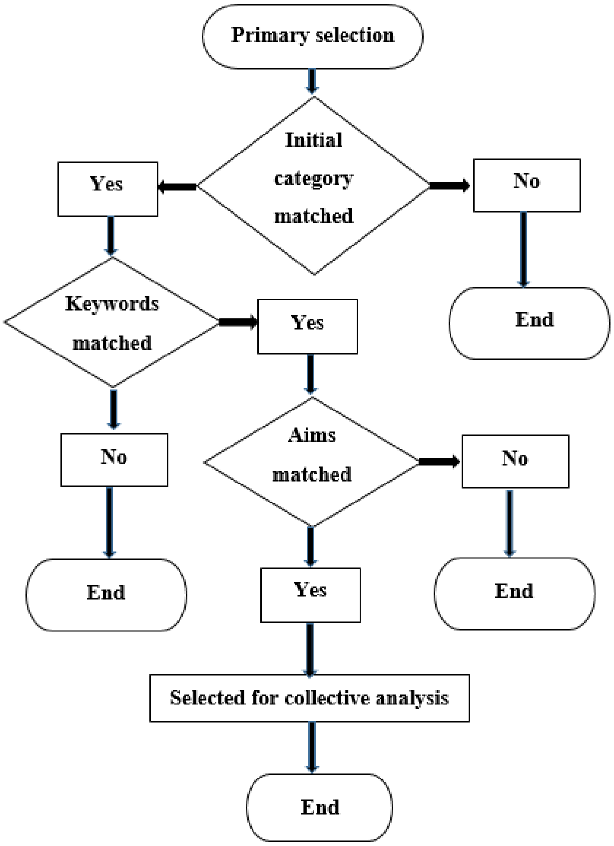


Figure 1: Flowchart for Data Selection

After following the procedure, several articles were rejected due to not matching the criteria. Finally detected papers were analyzed and the results of the analysis are categorized and explained in the result section.

4.0 FINDINGS

After following the procedures of simple collective method, a total of 20 papers were found to match the aim of this paper. From the selected papers analysis, the results found are as follows:

Role of Big Data Analytics in Metabolomics

In metabolomics, big data analytics is essential for data management and collection, data processing and integration, and the provision of statistical and computational techniques (Bekri, 2016; Liebal et al., 2020; Perez De Souza et al., 2020; Tolani et al., 2021). For metabolomics research to be effective, these are necessary.

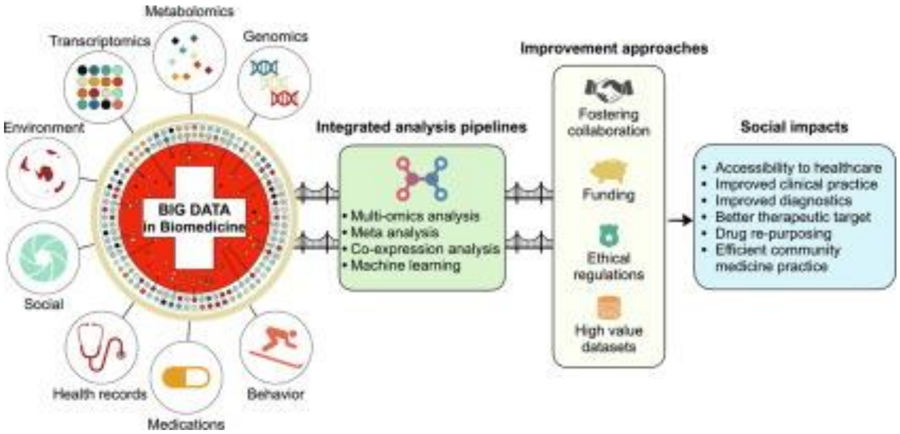


Figure 2: Advanced Usage of Big-Time Data Analysis for Metabolomics (Cremin et al., 2022)

Data Collection and Management

Metabolomics involves data collection, a process that begins with sample acquisition, typically involving blood, urine, or body tissue samples. These samples are then analyzed using large datasets, providing insights on hundreds to thousands of metabolites. Researchers compare primary database data from biological samples and repositories of metabolomics data for meta-analysis. However, processing these complex data sets can be challenging due to the overwhelming amount of information and the need for efficient storage. Data integrity and standardization are crucial for avoiding biases and ensuring the compatibility of data collection, processing, and reporting schemes. This standardization is essential for the replicability of metabolomics research and for making sound judgments based on Big Data perspectives.

Data Processing and Integration

Metabolomics data is complex and requires preprocessing steps to make it manageable for statistical analysis. Proper normalization corrects variations in sample concentration, enabling comparisons between samples. Techniques like logarithmic transformation modify data variance and make it suitable for statistical analysis. Combining metabolomics with genomics and proteomics can enhance the identification and understanding of complex molecular systems and pathological conditions. Combining genomic data with proteomics data can identify the genetic aspect of metabolic profiles. Despite computational complexity, there is potential for further development of biomarkers and therapeutic targets through integrating multi-omics in the future.

Statistical and Computational Technique

Metabolomics data analysis relies on statistical and computational methods, such as Principal Component Analysis (PCA) and Partial Least Squares Discriminant Analysis (PLS-DA). These methods help in reducing data dimensions and extracting features, enabling classification between different samples, such as diseased and healthy, based on metabolite content. The integration of machine learning and artificial intelligence in metabolomics data allows for the search for hidden relations in a wide range of data. For example, a dataset for training a model could be used with metabolomic data to predict disease results or classify samples based on metabolic profiles. Successful applications of statistical and computational methods in metabolomics include identifying new biomarkers for diagnosis and prognosis in cancer research and integrating metabolomics data with other types of genomic data to enhance understanding of specific diseases' molecular profiles.

Diagnostic Biomarkers

A set of procedures must be followed to identify a disease's biomarkers. Finding the appropriate biomarkers for a condition can be accomplished by following these procedures, which will help establish the best potential therapy options.

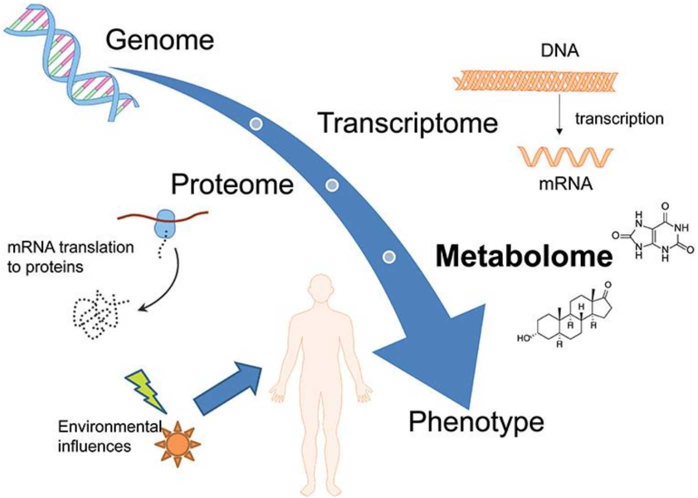


Figure 3: Metabolomics Strategies in Biomarkers(Steuer Et Al., 2019)

Biomarker Discovery Process: Discovering biomarkers in metabolomics involves using a sample from a diseased patient and a sample from a non-disease patient. These samples are then processed for various assays, such as Mass spectrometry, to detect the metabolites that characterize one group against the other. The aim is to identify metabolites that could be biomarkers for the disease. However, once the potential biomarkers have been discovered, they have to be further investigated for their reliability and the specificity of the results. This validation process involves testing on more subjects and often on samples with higher variance to ensure that the biomarkers correctly identify diseased and healthy individuals. Some approaches used to analyze the performance of potential biomarkers include the receiver operating characteristic (ROC) analysis. Biomarker selection criteria include the ability of the biomarker to fully represent the disease for which it is measured, temporal stability, and pathophysiologic relevance. Biomarkers that fulfill these parameters can be advanced into diagnostic tests, prognostic indicators, or therapeutic interventions.

Many examples have shown that metabolomics can successfully discover disease biomarkers. For instance, metabolomics has been employed in cancer research to perform biomarker discovery for diagnosis and prognosis. Another study found a mix of metabolites that are reliable biomarkers for detecting PC, which led to the creation of a noninvasive test for detecting the disease. In another study, metabolomics was applied to identify novel biomarkers for type 2 diabetes. Scientists found that this particular metabolic profile could point to the occurrence of diabetes many years before the manifestation of the first signs of the disease. This can have a positive impact on early detection and early management of the condition and, therefore, reduce the impact of diabetes on the health of the populace. These biomarkers have a large effect on disease diagnosis and treatment. Identifying reliable biomarkers that precede clinical symptoms allows for timely diagnosis and management, enhancing patient survival. Furthermore, biomarkers can be used to design personalized treatment strategies, for they give information on the pattern of the disease's metabolism in the system.

The identification of biomarkers requires validation to integrate metabolomics results into clinical practice. This step assists in confirming that the biomarkers that have been deemed possible for diagnostic use are accurate to other groups and conditions as well. These validation studies are usually conducted on large-scale clinical trials to confirm the efficacy of the biomarkers. Later, biomarkers can be utilized in clinical practice through diagnostic tests or therapeutic interventions. For example, a biomarker that was recently qualified through validation, which is a biomarker for cancer, can be used for a blood test for early detection, thereby increasing the chances of managing the disease. Similarly, biomarkers of metabolic diseases may be employed to determine the best course of action for the metabolism of the individual. However, using biomarkers in clinical practice is not without some challenges. Several barriers contribute to the lack of biomarkers incorporated into the daily practice of medicine: There is a section on challenges associated with biomarkers, with notable difficulties that include regulation approval for the biomarkers, cost consideration of the biomarkers, and other factors such as the need for specialized equipment and skilled workforce. Resolving these concerns requires collaboration between researchers, clinicians, and database authorities to ensure that the biomarkers are not only hypothetical and end phenotypes but also practical within different clinical settings.

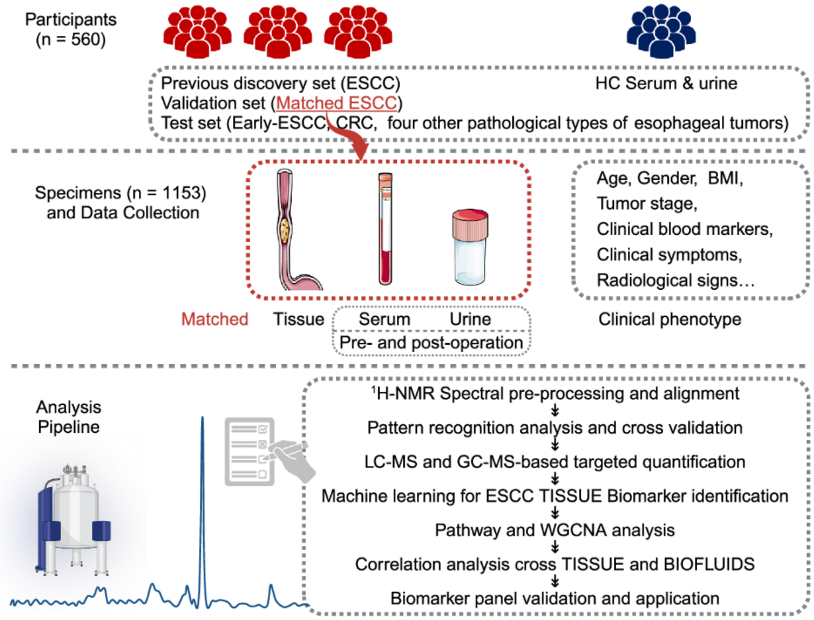


Figure 4: Characteristic Big Data Analysis for Metabolome Atlas (Steuer Et Al., 2019)

Challenges in Big Data Analytics for Metabolomics

To guarantee a successful data analysis process, it's crucial to take into account a few potential obstacles while working with big data analytics for metabolomics (Guo et al., 2022; Odenkirk et al., 2021). A number of challenges are mentioned in Figure 5.

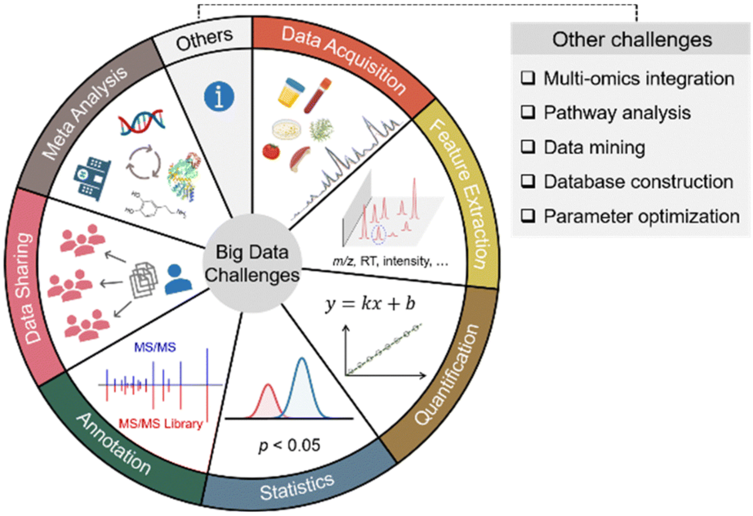


Figure 5: Various Challenges for Big-Time Data Big Data Analytics for Metabolomics (Odenkirk et al., 2021)

Data Complexity and Volume

Metabolomics is a field that involves the analysis of large quantities of data, often involving thousands of distinct metabolites. This data can be challenging to manage and analyze due to its complexity. To handle this, researchers use high-performance computing resources and the cloud to process the data efficiently. Data thinning, a process that involves data condensation, is used to make merging datasets more manageable and related. Another challenge is the compatibility of metabolomics with other platforms like genomics and proteomics, which is essential for providing a comprehensive view of biological systems but also complicates data analysis. High-end computational methods are required to analyze and extract features from diverse, large data sets.

Technical and Analytical Challenges

The most challenging avenues for further growth in the area of metabolomics are the technical and analytical points connected to data gathering, analysis, and interpretation. These challenges are among the many that are now being recognized as crucial and requiring development. Metabolomics data can be utilized with accuracy and reliability, contingent upon the chosen techniques of analysis. Despite their popularity, popular techniques like NMR spectroscopy and mass spectrometry are neither sensitive, precise, nor repeatable. One more technological issue with metabolomics is data handling, which entails managing substantial amounts of data produced throughout the procedure. Thousands of metabolites, each with varying quantities and intricate relationships, typically comprise the noisy data. Finding any relevant information is almost impossible. Combining data from metabolomics with other molecular technologies like proteomics and genomics results in more dimensional data and the complexity of the data and the analysis that follows, which is handled by applying higher-level algorithms for computation.

The type of analysis performed on the data produced is another often-checked technique in the context of analysis. However, because metabolomics data is high-dimensional, it might be difficult to identify clear patterns and relationships between different things. Furthermore, introducing unpredictability and noise into the data may mask powerful biological signals and lead to misunderstandings. However, these are only a few of the difficulties the researchers encounter, which is why they constantly create and alter different statistical and bioinformatics techniques to examine the metabolomics data and offer a deeper comprehension.

Ethical and Privacy Concerns

Ethical and privacy considerations are critical, particularly in the context of managing large amounts of data in metabolomics. Because personal data is involved in the collection, storing, and analysis of Metabolomics data, Metabolomics information may qualify as Sensitive Personal Data. The hazard that metabolomics poses to participant re-identification is the main ethical problem in this investigation. Since the data is frequently anonymized, sharing the data may raise privacy concerns for some people, however, these are not specific to metabolomics; combining the metabolomics data with the other omics data may put the privacy of the concerned parties at risk. Measures like data encryption and access restrictions should be implemented to lessen this danger.

Metabolomics data may be used for something other than diagnosis or research. This brings up the subject of consent about the intended use of an individual's data as well as the individual's general rights to the data being gathered. Regarding these worries, participants must be informed about and provide consent for the specific data being gathered and used through metabolomics. This is a required ethical practice for researchers and the institutions with which they are affiliated. Last but not least, the fact that metabolomics is an interdisciplinary field of study with participants from throughout the globe raises several additional ethical conduct issues. These are the following: How would the researchers note the variations in national legislation about data and moral behavior?

To foster cooperation, information must be made available for research and development without impairing individuals' liberties and rights.

Possible Developments and Innovations

Numerous cutting-edge technologies and fresh approaches that could improve the activity are made possible by the ongoing growth of metabolomics. These advancements may promote precision medicine, the identification of biomarkers, and the incorporation of metabolomics data into other omics fields.

Emerging Technologies

Metabolomics data is being transformed by AI and blockchain technology, enhancing biomarker discovery and reducing the need for centralized analysis. AI's machine learning capabilities enable the identification of biomarkers in diseases, allowing early detection and diagnosis. Blockchain offers a secure, efficient storage solution for metabolomics data among multiple research entities, ensuring originality and protection from piracy. This technology can also support alliance formation without sharing personal data. While not commonly used in metabolomics, these technologies may help counter negative tendencies such as data volume, ownership, and compatibility issues. As a result, these technologies will become integral to metabolomics analysis, advancing biomarker discovery and precision medicine effectiveness.

Personalized Medicine

Metabolomics has made a significant contribution to personalized medicine by identifying individual metabolic profiles. By comparing these profiles, scientists can identify disease biomarkers and predict the likelihood of developing a specific illness or reaction to a specific therapy. This information is useful in designing effective treatments with fewer side effects. For example, in oncology, metabolomics can observe metabolic shifts in cancer cells, enabling treatments targeting the patient's metabolic profile. In cardiovascular diseases, metabolomics can predict a patient's response to specific drugs, controlling dosage and reducing harm. Utilizing metabolomics data for individualizing treatments is a revolution in healthcare, enhancing care quality and reducing treatment costs. With further research and the discovery of more biomarkers, metabolomics will become even more relevant in personalized medicine, enhancing the delivery of individualized health solutions.

Multi-Omics Integration

Metabolomics is a popular field in omics research, combined with genomics, proteomics, and transcriptomics. This integrated multi-omics approach provides a systemic view of biological processes, enabling simultaneous analysis of multiple molecules and pathways. It allows for the definition of relationships between metabolite alterations, genetic differences, protein levels, and other molecular phenomena, identifying novel biomarkers and potential therapeutic targets. Merging metabolomics with genomics has shown that gene-metabolite connections are crucial to disease onset and progression. However, integrating multi-omics data presents challenges in data management and analysis, such as high-performance computational methods and models, and harmonization of different data types. Despite these challenges, integrating data from different omics layers is essential for the identification of biomarkers more accurately and for disease treatment. Disease treatment technology and methodologies integrating multi-omics data are expected to become more common in metabolomics studies.

5.0 CONCLUSION AND RECOMMENDATIONS

Big data analytics is radically changing the field of metabolomics by offering methods for obtaining biomarkers essential for metabolite profiling-based disease diagnosis, prognosis, and treatment. When metabolomics is combined with the most advanced blockchain, artificial intelligence, and statistical and computational analysis it improves the identification and finding of biomarkers. The field of biomarker research has found that metabolomics is a formidable tool, and its implementation has significantly improved our capacity to evaluate disease assumptions. Biomarker big-time data discovery is made possible by metabolomics. Metabolomics using Big Data Analytics. It has become possible to use metabolomics to investigate how medications and other treatments affect the metabolism of the host. Measuring the metabolites generated by the microbiota can provide insight into the activities of the gut flora. This important technique for developing and formulating precision with large data analysis is metabolomics. There are many challenges concerning data concerns, technological limitations, and ethical dilemmas. Nevertheless, additional study and development on these issues is presently underway, which will result in metabolomics analyses that are more effective and efficient. The trend of more research being done on multi-omics shows that comprehensive methods must be used to understand intricate systems, which are critical to the advancement of customized healthcare. The best practices for data administration and analysis should be implemented along with ongoing technological advancements to address various duties in a field. The integration of metabolomics with other "omics fields" and the addition of new technology and initiatives to bring ideas of personalized medicine to life. Using big data analysis and since metabolomics is always changing, further discoveries can be made in the future. Advantageous for the understanding of disorders and the treatment of patients.

REFERENCES

Abdelnabi, S., Hasan, R., & Fritz, M. (2022). Open-domain, content-based, multi-modal fact-checking of out-of-context images via online resources. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition,

Akid, A. S. M., Shah, S. M. A., Sobuz, M. D. H. R., Tam, V. W. Y., & Anik, S. H. (2021). Combined influence of waste steel fibre and fly ash on rheological and mechanical performance of fibre-reinforced concrete. Australian Journal of Civil Engineering, 19(2), 208-224. https://doi.org/10.1080/14488353.2020.1857927

Akid, A. S. M., Wasiew, Q. A., Sobuz, M. H. R., Rahman, T., & Tam, V. W. (2021). Flexural behavior of corroded reinforced concrete beam strengthened with jute fiber reinforced polymer. Advances in Structural Engineering, 24(7), 1269-1282. https://doi.org/10.1177/1369433220974783

Alonso, A., Marsal, S., & Julià, A. (2015). Analytical methods in untargeted metabolomics: state of the art in 2015. Frontiers in bioengineering and biotechnology, 3, 23. https://doi.org/https://doi.org/10.3389/fbioe.2015.00023

Batko, K., & Ślęzak, A. (2022). The use of Big Data Analytics in healthcare. Journal of big Data, 9(1), 3. https://doi.org/https://doi.org/10.1186/s40537-021-00553-4

Bekri, S. (2016). The role of metabolomics in precision medicine. Expert Review of Precision Medicine and Drug Development, 1(6), 517-532. https://doi.org/https://doi.org/10.1080/23808993.2016.1273067

Cremin, C. J., Dash, S., & Huang, X. (2022). Big data: historic advances and emerging trends in biomedical research. Current Research in Biotechnology, 4, 138-151. https://doi.org/https://doi.org/10.1016/j.crbiot.2022.02.004

Das, S., Habibur Rahman Sobuz, M., Tam, V. W. Y., Akid, A. S. M., Sutan, N. M., & Rahman, F. M. M. (2020). Effects of incorporating hybrid fibres on rheological and mechanical properties of fibre reinforced concrete. Construction and Building Materials, 262, 120561. https://doi.org/https://doi.org/10.1016/j.conbuildmat.2020.120561

Datta, S. D., Islam, M., Rahman Sobuz, M. H., Ahmed, S., & Kar, M. (2024). Artificial intelligence and machine learning applications in the project lifecycle of the construction industry: A comprehensive review. Heliyon, 10(5), e26888. https://doi.org/https://doi.org/10.1016/j.heliyon.2024.e26888

Guasch-Ferré, M., Bhupathiraju, S. N., & Hu, F. B. (2018). Use of metabolomics in improving assessment of dietary intake. Clinical chemistry, 64(1), 82-98. https://doi.org/https://doi.org/10.1373/clinchem.2017.272344

Guo, J., Yu, H., Xing, S., & Huan, T. (2022). Addressing big data challenges in mass spectrometry-based metabolomics. Chemical Communications, 58(72), 9979-9990. https://doi.org/https://doi.org/10.1039/D2CC03598G

Hasan, N. M. S., Sobuz, M. H. R., Khan, M. M. H., Mim, N. J., Meraz, M. M., Datta, S. D., Rana, M. J., Saha, A., Akid, A. S. M., Mehedi, M. T., Houda, M., & Sutan, N. M. (2022). Integration of Rice Husk Ash as Supplementary Cementitious Material in the Production of Sustainable High-Strength Concrete. Materials, 15(22), 8171. https://www.mdpi.com/1996-1944/15/22/8171

Hilal, W., Gadsden, S. A., & Yawney, J. (2022). Financial fraud: a review of anomaly detection techniques and recent advances. Expert systems With applications, 193, 116429.

Jakhrani, A., Samo, S., Sobuz, H. R., Uddin, M. A., Ahsan, M., & Hasan, N. M. S. (2012). Assessment of dissolved salts concentration of seawater in the vicinity of Karachi. International Journal of Structural and Civil Engineering, 1(2), 61-69.

Jin, Q., & Ma, R. C. W. (2021). Metabolomics in diabetes and diabetic complications: insights from epidemiological studies. Cells, 10(11), 2832. https://doi.org/https://doi.org/10.3390/cells10112832

Liebal, U. W., Phan, A. N., Sudhakar, M., Raman, K., & Blank, L. M. (2020). Machine learning applications for mass spectrometry-based metabolomics. Metabolites, 10(6), 243. https://doi.org/https://doi.org/10.3390/metabo10060243

Nagana Gowda, G., & Raftery, D. (2021). NMR-based metabolomics. Cancer Metabolomics: Methods and Applications, 19-37. https://doi.org/https://doi.org/10.1007/978-3-030-51652-9_2

Nicholls, J., Kuppa, A., & Le-Khac, N.-A. (2021). Financial cybercrime: A comprehensive survey of deep learning approaches to tackle the evolving financial crime landscape. Ieee Access, 9, 163965-163986.

Odenkirk, M. T., Reif, D. M., & Baker, E. S. (2021). Multiomic big data analysis challenges: increasing confidence in the interpretation of artificial intelligence assessments. Analytical chemistry, 93(22), 7763-7773. https://doi.org/https://doi.org/10.1021/acs.analchem.0c04850

Perez De Souza, L., Alseekh, S., Brotman, Y., & Fernie, A. R. (2020). Network-based strategies in metabolomics data analysis and interpretation: from molecular networking to biological interpretation. Expert Review of Proteomics, 17(4), 243-255. https://doi.org/https://doi.org/10.1080/14789450.2020.1766975

Rana, J., Hasan, R., Sobuz, H. R., & Tam, V. W. Y. (2022). Impact assessment of window to wall ratio on energy consumption of an office building of subtropical monsoon climatic country Bangladesh. International Journal of Construction Management, 22(13), 2528-2553. https://doi.org/10.1080/15623599.2020.1808561

Rana, M. J., Hasan, M. R., & Sobuz, M. H. R. (2022). An investigation on the impact of shading devices on energy consumption of commercial buildings in the contexts of subtropical climate. Smart and Sustainable Built Environment, 11(3), 661-691. https://doi.org/10.1108/SASBE-09-2020-0131

Rana, M. J., Hasan, M. R., Sobuz, M. H. R., & Sutan, N. M. (2021). Evaluation of passive design strategies to achieve NZEB in the corporate facilities: The context of Bangladeshi subtropical monsoon climate. International Journal of Building Pathology and Adaptation, 39(4), 619-654.

Schmidt, D. R., Patel, R., Kirsch, D. G., Lewis, C. A., Vander Heiden, M. G., & Locasale, J. W. (2021). Metabolomics in cancer research and emerging applications in clinical oncology. CA: a cancer journal for clinicians, 71(4), 333-358. https://doi.org/https://doi.org/10.3322/caac.21670

Sobuz, M. H. R., Datta, S. D., & Rahman, M. (2022). Evaluating the Properties of Demolished Aggregate Concrete with Non-destructive Assessment. In S. Arthur, M. Saitoh, & S. K. Pal (Eds.), Advances in Civil Engineering, Lecture Notes in Civil Engineering (pp. 223-233). Springer Singapore. https://doi.org/10.1007/978-981-16-5547-0_22

Steuer, A. E., Brockbals, L., & Kraemer, T. (2019). Metabolomic strategies in biomarker research–new approach for indirect identification of drug consumption and sample manipulation in clinical and forensic toxicology? Frontiers in Chemistry, 7, 319. https://doi.org/https://doi.org/10.3389/fchem.2019.00319

Taherdoost, H. (2022). Blockchain technology and artificial intelligence together: a critical review on applications. Applied Sciences, 12(24), 12948. https://doi.org/https://doi.org/10.3390/app122412948

Tolani, P., Gupta, S., Yadav, K., Aggarwal, S., & Yadav, A. K. (2021). Big data, integrative omics and network biology. Advances in protein chemistry and structural biology, 127, 127-160. https://doi.org/https://doi.org/10.1016/bs.apcsb.2021.03.006

Van Ravenzwaay, B., Cunha, G. C.-P., Leibold, E., Looser, R., Mellert, W., Prokoudine, A., Walk, T., & Wiemer, J. (2007). The use of metabolomics for the discovery of new biomarkers of effect. Toxicology letters, 172(1-2), 21-28. https://doi.org/https://doi.org/10.1016/j.toxlet.2007.05.021