

# Medical Omics

Jens Allmer<sup>1</sup>

Ralf Hofestädt<sup>2</sup>

1 Hochschule Ruhr West,  
Institute for Measurement  
Engineering and Sensor  
Technology, Mülheim an der  
Ruhr, Germany

2 Bielefeld University, Faculty  
of Technology,  
Bioinformatics/Medical  
Informatics Department,  
Bielefeld, Germany

## Editorial

Diseases are multifaceted, and their manifestation can be very individual. This is reflected in the trend toward precision medicine. For this individualization of diagnosis and treatment, molecular data plays an increasingly important role. These data include, among others, genomic, transcriptomic, and proteomic data. The measurement and evaluation of these data are treated in the corresponding fields of genomics, transcriptomics, and proteomics. The connection to medicine is made through the area of medical omics, which also considers data integration.

In recent years, we have offered workshops on this topic at various local and international events and the annual meetings of the GMDS. The work in this special issue stems from two events, the 1<sup>st</sup> International Applied Bioinformatics Conference and the 67<sup>th</sup> Annual Meeting of the GMDS.

Six contributions have been selected after peer review. They reflect the broad spectrum of medical omics, which relates not only to data and their integration but also to the coherence of the data structure, including the connection to patients and knowledge databases.

Beukers and Allmer [1] deal with the possibility of automated analysis of transcriptomic data. Such data form the basis for higher-value studies in medical omics and reflect the cellular state. The authors compare three workflow management systems (Galaxy, KNIME, and CLC) for the creation and execution of RNA-seq data analysis workflows. They find that each tool has strengths and weaknesses and that the workflows created by each tool lead to different results. The authors also provide recommendations on which tool should be used and share their workflows to promote the development of best practices for RNA-seq data analysis.

The next step after data analysis of raw data is the integration with previously acquired knowledge. Unlu Yazıcı and colleagues [2] write that individual omics datasets are insufficient to fully understand the molecular mechanisms of complex diseases that are influenced by multiple factors. They use machine learning to integrate multiple omics data. They use existing data to create models for prediction and classification. The authors further discuss the importance of interpreting the biological significance of model outputs with corresponding biological mechanisms and the potential clinical applications.

A similar approach is presented by Königs and Dietrich [3]. However, other data types are used here, and the prediction is a pathway. Pathway enrichment is used to analyze gene expression data and identify significant pathways, while the PharMe-BINet database uses correction methods to reduce false-positive results. An analysis was performed on gene expression data from T-cell lymphomas.

Once data has been obtained at the molecular level, and knowledge has been integrated, it should find application in the hospital. For this, Raupach [4] proposes an integration method. A software module integrates molecular data into a hospital information system to improve the safety of drug therapy and performs pharmacogenetic reviews in an existing drug therapy workflow. The results show the importance of regular, pre-emptive genotyping of patients and the prevalence of genetic variants in the population.

While creating and testing new drugs (also for precision medicine) is essential, it is equally important to research why studies fail. Friedrichs [5] presents a web-based platform that categorizes the reasons for the failure of clinical studies. As of September 23, 2022, the database contains 14,232 failed studies, of which 51.5% are already commented on. The goal is to use the data for

information acquisition for decision-making and the reuse of drugs.

All of the presented works go hand in hand with Savoska and colleagues' work [6], which combines various medical data in a personalized way. The authors propose a cloud-based model to integrate health and medical data from multiple sources, including electronic health records and measurement sensors, into a personal health record to improve the individualized prediction of diseases and treatment of patients. Omics data should also be integrated into this model, which complies with the necessary data security and data protection standards.

These contributions show the broad spectrum of the field of medical omics but cannot reflect its breadth. Other aspects, such as teaching, application, and quality assurance, have yet to be considered here. This highlights the importance of establishing the field of medical omics as a working group in the GMDS. We encourage all interested parties to contact and work with us to advance this critical endeavor.

## Notes

### Competing interests

The authors declare that they have no competing interests.

## References

1. Beukers M, Allmer J. Challenges for the development of automated RNA-seq analyses pipelines. *GMS Med Inform Biom Epidemiol.* 2023;19(1):Doc06. DOI: 10.3205/mibe000245
2. Unlu Yazıcı M, Bakir-Gungor B, Yousef M. Integrative analyses in omics data: Machine learning perspective. *GMS Med Inform Biom Epidemiol.* 2023;19(1):Doc05. DOI: 10.3205/mibe000244
3. Königs C, Dietrich T. A web-based pathway enrichment analysis module for the PharMeBInet database. *GMS Med Inform Biom Epidemiol.* 2023;19(1):Doc04. DOI: 10.3205/mibe000243
4. Raupach L. Integration of biological molecular data into an existing drug therapy safety workflow used in hospital information systems. *GMS Med Inform Biom Epidemiol.* 2023;19:Doc03. DOI: 10.3205/mibe000242
5. Friedrichs M. A web-based annotation tool for clinical trial failure reasons. *GMS Med Inform Biom Epidemiol.* 2023;19(1):Doc02. DOI: 10.3205/mibe000241
6. Savoska S, Ristevski B, Blazheska-Tabakovska N, Jolevski I, Bocevska A, Trajkovic V. Integration of heterogeneous medical and biological data with electronic personal health records. 2023;19(1):Doc01. DOI: 10.3205/mibe000240

### Corresponding author:

Prof. Dr. Jens Allmer  
Hochschule Ruhr West, University of Applied Sciences,  
Institute for Measurement Engineering and Sensor  
Technology, Duisburger Str. 100, 45479 Mülheim an der  
Ruhr, Germany  
jens@allmer.de

### Please cite as

Allmer J, Hofestädt R. *Medical Omics.* *GMS Med Inform Biom Epidemiol.* 2023;19:Doc07.  
DOI: 10.3205/mibe000246, URN: urn:nbn:de:0183-mibe0002469

### This article is freely available from

<https://doi.org/10.3205/mibe000246>

**Published:** 2023-07-04

### Copyright

©2023 Allmer et al. This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 License. See license information at <http://creativecommons.org/licenses/by/4.0/>.

# Medizinische Omik

Jens Allmer<sup>1</sup>

Ralf Hofestädt<sup>2</sup>

1 Hochschule Ruhr West,  
Institut Mess- und  
Sensortechnik, Mülheim an  
der Ruhr, Deutschland

2 Universität Bielefeld,  
Technische Fakultät,  
Arbeitsgruppe Bioinformatik/  
Medizininformatik, Bielefeld,  
Deutschland

## Editorial

Krankheiten sind vielgestaltig und ihre Ausprägungen können sehr individuell sein. Dies spiegelt sich im Trend zur Präzisionsmedizin wider. Für diese Individualisierung der Diagnose und Behandlung spielen molekulare Daten eine immer wichtigere Rolle. Zu diesen Daten zählen unter anderem genomische, transkriptomische und proteomische Daten. Die Messung und Auswertung dieser Daten werden in den entsprechenden Feldern wie Genomik, Transkriptomik und Proteomik behandelt. Die Verbindung zur Medizin wird durch das Feld medizinische Omik, welches auch die Datenintegration betrachtet, hergestellt. In den letzten Jahren haben wir, im Rahmen verschiedener lokaler und internationaler Veranstaltungen und auch der Jahrestagungen der GMDS, Workshops zu diesem Thema angeboten. Die Arbeiten in diesem Sonderheft entspringen zweier Veranstaltungen, einmal der 1<sup>st</sup> International Applied Bioinformatics Conference und der 67. Jahrestagung der GMDS.

Es wurden nach einem Peer Review-Verfahren sechs Beiträge ausgewählt. Diese bilden das breite Spektrum der medizinischen Omik ab, welches sich natürlich nicht nur auf die Daten und deren Integration, sondern auch auf Kohärenz der Datenstruktur inklusive der Verbindung zu Patienten und Wissensdatenbanken erstreckt.

Beukers and Allmer [1] beschäftigen sich mit der Möglichkeit der automatisierten Analyse von transkriptomischen Daten. Solche Daten bilden die Basis für höherwertige Analysen in der medizinischen Omik und bilden den zellulären Ist-Zustand ab. Die Autoren vergleichen drei Workflow-Management-Systeme (Galaxy, KNIME und CLC) zur Erstellung und Ausführung von RNA-seq-Datenanalyse-Workflows. Sie finden heraus, dass jedes Werkzeug seine eigenen Stärken und Schwächen hat und dass die von jedem Werkzeug erstellten Workflows zu unterschiedlichen Ergebnissen führen. Die Autoren geben auch Empfehlungen darüber ab, welches Werkzeug verwendet

werden sollte und teilen ihre Workflows, um die Entwicklung von Best Practices für die RNA-seq-Datenanalyse zu fördern.

Der nächste Schritt nach der Datenanalyse der Rohdaten ist die Integration mit zuvor akquiriertem Wissen. Ünlü Yazıcı und Kollegen [2] schreiben, dass einzelne Omik-Datensätze nicht ausreichen, um die molekularen Mechanismen von komplexen Krankheiten, die von mehreren Faktoren beeinflusst werden, vollständig zu verstehen. Sie nutzen maschinelles Lernen, um mehrere Omik-Daten zu integrieren. Dabei verwenden sie vorhandene Daten, um Modelle für Vorhersage und Klassifikation zu erstellen. Die Autoren gehen weiter auf die Bedeutung der Interpretation der biologischen Bedeutung von Modellausgaben im Zusammenhang mit entsprechenden biologischen Kenntnissen ein.

Ein ähnlicher Weg wird von Königs und Dietrich [3] vorgestellt. Hier werden jedoch andere Datentypen verwendet und die Vorhersage ist ein Pathway. Pathwayanreicherung wird verwendet, um Geneexpressionsdaten zu analysieren und signifikante Pathways zu identifizieren, während die PharMe-BINet-Datenbank Korrekturmethode verwendet, um falsch-positive Ergebnisse zu reduzieren. Eine Analyse wurde an Genexpressionsdaten von T-Zell-Lymphomen durchgeführt.

Daten, die auf der molekularen Ebene gewonnen wurden, und Wissen, das integriert wurde, sollten im Krankenhaus Anwendung finden. Dafür schlägt Raupach [4] eine Integrationsmethode vor. Ein Softwaremodul integriert molekulare Daten in ein Krankenhausinformationssystem, um die Sicherheit der Arzneimitteltherapie zu verbessern und führt pharmakogenetische Überprüfungen in einem vorhandenen Arzneimitteltherapie-Workflow durch. Die Ergebnisse zeigen die Wichtigkeit regelmäßiger, präemptiver Genotypisierung von Patienten und die Prävalenz von Genvarianten in der Bevölkerung.

Während es wichtig ist, neue Medikamente (auch für Präzisionsmedizin) zu erzeugen und zu testen, ist es ge-

nauso wichtig, zu erforschen, warum Studien fehlschlagen. Friedrichs [5] stellt eine webbasierte Plattform vor, die Gründe für das Scheitern von klinischen Studien kategorisieren soll. Stand 23. September 2022 enthält die Datenbank 14.232 fehlgeschlagene Studien, von denen 51,5% bereits kommentiert sind. Ziel ist es, die Daten zur Informationsgewinnung für die Entscheidungsfindung und die Wiederverwendung von Medikamenten zu nutzen. Alle vorgestellten Arbeiten gehen Hand in Hand mit der Arbeit von Savoska und Kollegen [6], die verschiedene medizinische Daten personalisiert zusammenführt. Die Autoren schlagen ein cloudbasiertes Modell vor, um Gesundheits- und medizinische Daten aus verschiedenen Quellen, einschließlich elektronischer Gesundheitsakten und Messensoren, in ein persönliches Gesundheitsdossier zu integrieren, um die individualisierte Prognose von Erkrankungen und die Behandlung von Patienten zu verbessern. Omik-Daten sollten auch in dieses Modell integriert werden, das die erforderlichen Datensicherheits- und Datenschutzstandards einhält.

Diese Beiträge zeigen das breite Spektrum des Bereichs medizinische Omik auf, können aber nicht die gesamte Bandbreite widerspiegeln. Weitere Felder wie Lehre, Anwendung und Qualitätssicherung wurden hier nicht betrachtet. Dies zeigt auf, wie wichtig es ist, das Feld medizinische Omik als Arbeitsgruppe in der GMDS zu etablieren. Wir wollen dieses Editorial noch zum Anlass nehmen, alle Interessierten aufzufordern, sich zu melden und mit uns dieses wichtige Vorhaben voranzubringen.

## Anmerkungen

## Interessenkonflikte

Die Autoren erklären, dass sie keine Interessenkonflikte in Zusammenhang mit diesem Artikel haben.

## Literatur

1. Beukers M, Allmer J. Challenges for the development of automated RNA-seq analyses pipelines. *GMS Med Inform Biom Epidemiol.* 2023;19(1):Doc06. DOI: 10.3205/mibe000245
2. Unlu Yazıcı M, Bakir-Gungor B, Yousef M. Integrative analyses in omics data: Machine learning perspective. *GMS Med Inform Biom Epidemiol.* 2023;19(1):Doc05. DOI: 10.3205/mibe000244
3. Königs C, Dietrich T. A web-based pathway enrichment analysis module for the PharMeBInet database. *GMS Med Inform Biom Epidemiol.* 2023;19(1):Doc04. DOI: 10.3205/mibe000243
4. Raupach L. Integration of biological molecular data into an existing drug therapy safety workflow used in hospital information systems. *GMS Med Inform Biom Epidemiol.* 2023;19:Doc03. DOI: 10.3205/mibe000242
5. Friedrichs M. A web-based annotation tool for clinical trial failure reasons. *GMS Med Inform Biom Epidemiol.* 2023;19(1):Doc02. DOI: 10.3205/mibe000241
6. Savoska S, Ristevski B, Blazheska-Tabakovska N, Jolevski I, Bocevska A, Trajkovik V. Integration of heterogeneous medical and biological data with electronic personal health records. 2023;19(1):Doc01. DOI: 10.3205/mibe000240

### Korrespondenzadresse:

Prof. Dr. Jens Allmer  
Hochschule Ruhr West, Institut Mess- und Sensortechnik,  
Mülheim an der Ruhr, Deutschland  
jens@allmer.de

### Bitte zitieren als

Allmer J, Hofestädt R. Medical Omics. *GMS Med Inform Biom Epidemiol.* 2023;19:Doc07.  
DOI: 10.3205/mibe000246, URN: urn:nbn:de:0183-mibe0002469

### Artikel online frei zugänglich unter

<https://doi.org/10.3205/mibe000246>

Veröffentlicht: 04.07.2023

### Copyright

©2023 Allmer et al. Dieser Artikel ist ein Open-Access-Artikel und steht unter den Lizenzbedingungen der Creative Commons Attribution 4.0 License (Namensnennung). Lizenz-Angaben siehe <http://creativecommons.org/licenses/by/4.0/>.