

Artificial intelligence and systems biology analysis in stem cell research and therapeutics development

Thayna Silva-Sousa^{1,2,3,4,†}, Júlia Nakanishi Usuda^{1,2,3,4,5,6,†}, Nada Al-Arawe^{1,2,3,4,5,7,8,†}, Irene Hinterseher, MD^{3,4,5,9}, Rusan Catar, PhD⁷, Christian Luecht, MD⁷, Pedro Vallecillo Garcia, PhD^{5,8}, Katarina Riesner, PhD^{5,8}, Alexander Hackel, PhD¹⁰, Lena F. Schimke, MD¹¹, Haroldo Dutra Dias¹², Igor Salerno Filgueiras¹¹, Helder I. Nakaya^{6,13}, Niels Olsen Saraiva Camara, MD⁶, Stefan Fischer¹⁰, Gabriela Riemekasten, MD^{10,14}, Olle Ringdén, MD¹⁴, Olaf Penack, MD^{5,8}, Tobias Winkler, MD^{1,2}, Georg Duda^{1,2}, Dennyson Leandro M. Fonseca^{1,2,5,12,#}, Otávio Cabral-Marques^{6,11,12,13,15,#}, Guido Moll, PhD^{1,2,7,#,*}

¹BIH Center for Regenerative Therapies (BCRT), Charité Universitätsmedizin Berlin, Humboldt-Universität zu Berlin, Berlin Institute of Health (BIH), Berlin 10117, Germany

²Julius Wolff Institute (JWI), Charité Universitätsmedizin Berlin, Humboldt-Universität zu Berlin, Berlin Institute of Health (BIH), Berlin 10117, Germany

³Department of Vascular Surgery, Universitätsklinikum Ruppin-Brandenburg, Medizinische Hochschule Brandenburg Theodor Fontane, Neuruppin 16816, Germany

⁴Fakultät für Gesundheitswissenschaften Brandenburg, Gemeinsame Fakultät der Universität Potsdam, der Medizinischen Hochschule Brandenburg Theodor Fontane, und der Brandenburgischen Technischen Universität Cottbus-Senftenberg, Potsdam 14476, Germany

⁵Berlin Institute of Health (BIH), Charité Universitätsmedizin Berlin, Humboldt-Universität zu Berlin, Berlin Institute of Health (BIH), Berlin 10117, Germany

⁶Department of Clinical and Toxicological Analyses, School of Pharmaceutical Sciences, University of São Paulo (USP), São Paulo, Brazil

⁷Department of Nephrology and Internal Intensive Care Medicine, Charité Universitätsmedizin Berlin, Humboldt-Universität zu Berlin, Berlin Institute of Health (BIH), Berlin 10117, Germany

⁸Department of Hematology, Oncology, and Tumorimmunology, Charité Universitätsmedizin Berlin, Humboldt-Universität zu Berlin, Berlin Institute of Health (BIH), Berlin 10117, Germany

⁹Vascular Surgery, Charité Universitätsmedizin Berlin, Humboldt-Universität zu Berlin, Berlin Institute of Health (BIH), Berlin 10117, Germany

¹⁰Department of Rheumatology and Clinical Immunology, University Medical Center Schleswig Holstein Campus Lübeck, Lübeck, Germany

¹¹Department of Immunology, Institute of Biomedical Sciences, USP, São Paulo, Brazil

¹²Interunit Postgraduate Program on Bioinformatics, Institute of Mathematics and Statistics (IME), USP, São Paulo, Brazil

¹³Department of Medicine, Division of Molecular Medicine, Laboratory of Medical Investigation 29, USP School of Medicine (USPM), São Paulo, Brazil

¹⁴Division of Pediatrics, Department of CLINTEC, Karolinska Institutet, Stockholm, Sweden

¹⁵D'OR Institute Research and Education, São Paulo, Brazil

*Corresponding author: BIH Center for Regenerative Therapies (BCRT), Julius Wolff Institute (JWI), and Department of Nephrology and Internal Intensive Care Medicine, Charité Universitätsmedizin Berlin, Campus Virchow Klinikum (CVK), Augustenburger Platz 1, Berlin D-13353, Germany (guido.moll@charite.de).

†These authors shared first authors.

#These authors are senior authors.

Abstract

Background: Stem cell research has rapidly advanced during the past decades, but the translation into approved clinical products is still lagging behind. Multiple barriers to effective clinical translation exist. We hypothesize that an ineffective use of the existing wealth of data from both product development and clinical trials is a crucial barrier that hampers effective clinical implementation of stem cell therapies.

Methods and Results: Here, we summarize the contribution of systems biology (SysBio) and artificial intelligence (AI) in stem cell research and therapy development, to better understand and overcome these barriers to effective clinical translation. Advancements in cell product profiling technology, clinical trial design, and adjunct clinical monitoring, offer new opportunities for a more integrated understanding of both, product and patient performance. Synergy of SysBioAI analysis is boosting a more rapid, integrated, and informative analysis of large-scale multi-omics data sets of patient and clinical trial outcomes, thus enabling the “Iterative Circle of Refined Clinical Translation.” This SysBioAI-supported concept can assist more effective development and clinical use of stem cell therapeutics through iterative adaptation cycles. This includes product- and patient-centered clinical safety and efficacy/potency evaluation through paired identification of suitable biomarkers of clinical response.

Conclusion: Integrated SysBioAI-use is a powerful tool to optimize the design and outcomes of clinical trials by identifying patient-specific responses, contributing to enhanced treatment safety and efficacy, and to spur new patient-centric and adaptable next-generation deep-medicine approaches.

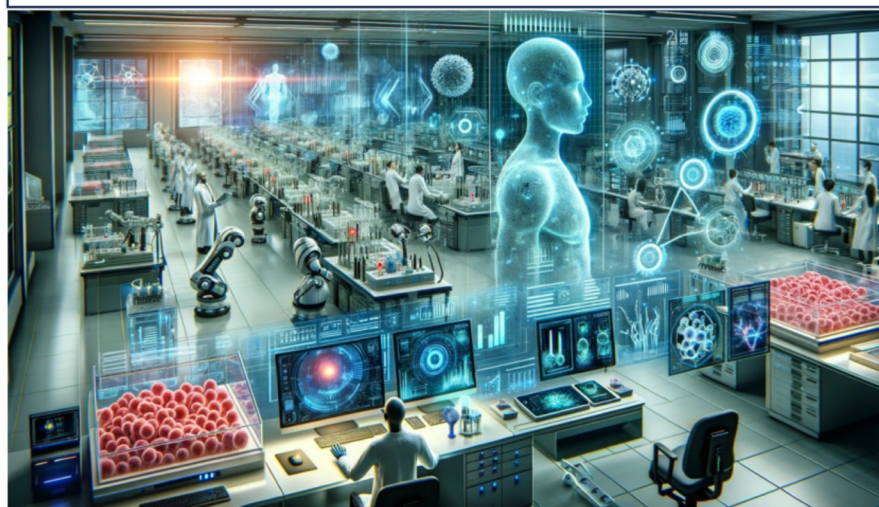
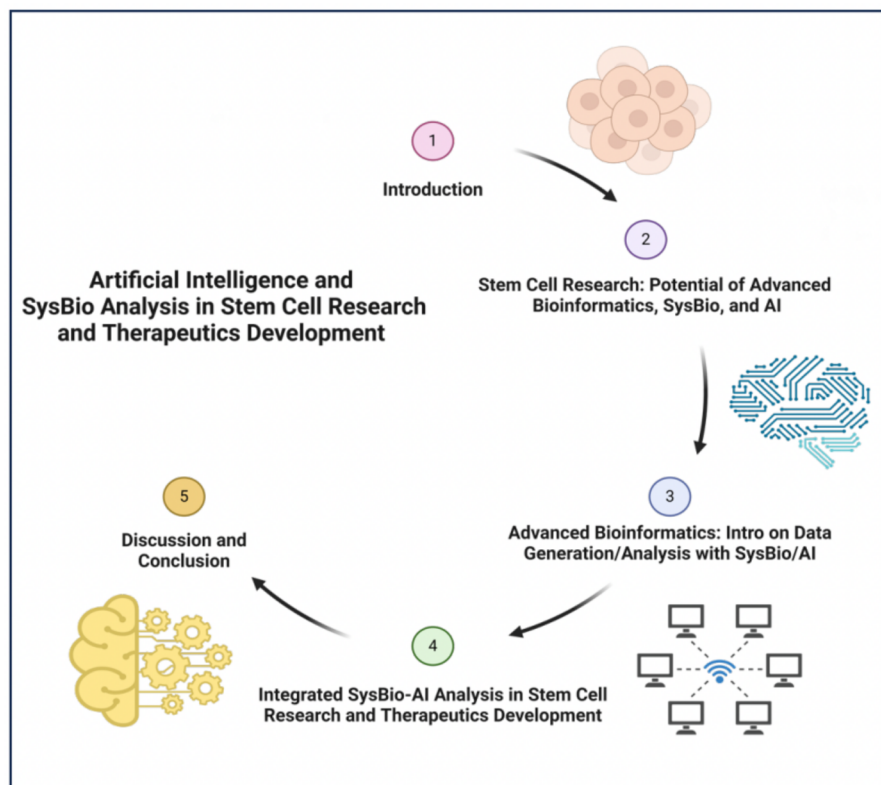
Key words: artificial intelligence; machine learning; systems biology; stem cell research; mesenchymal stromal/stem cells.

Received: 7 May 2025; Accepted: 15 July 2025.

© The Author(s) 2025. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.

Graphical abstract



Significance Statement

The clinical translation of stem cell therapies has been hindered by persistent challenges, including product heterogeneity, incomplete mechanistic understanding, and limited predictive power of current trial designs. This review highlights how the integration of systems biology (SysBio) and artificial intelligence (AI) can transform these limitations into opportunities. By enabling the holistic analysis of multi-omics datasets, patient biomarkers, and clinical outcomes, SysBioAI offers an iterative framework for refining both therapeutic products and trial strategies. This patient-centered, data-driven approach not only promises to accelerate safe and effective deployment of stem cell therapies but also establishes a paradigm for next-generation precision and regenerative medicine.

Introduction

Stem cell research has advanced rapidly over the past decades, setting the basis for novel therapeutic approaches.^{1–5} Translating these early advancements into mature therapies has faced

significant challenges, unraveling numerous barriers to effective clinical translation that must be understood and overcome.^{5–11}

This review article highlights the potential of advanced bioinformatics, artificial intelligence (AI), and systems biology (SysBio) to assist in this important endeavor (Graphical Abstract).

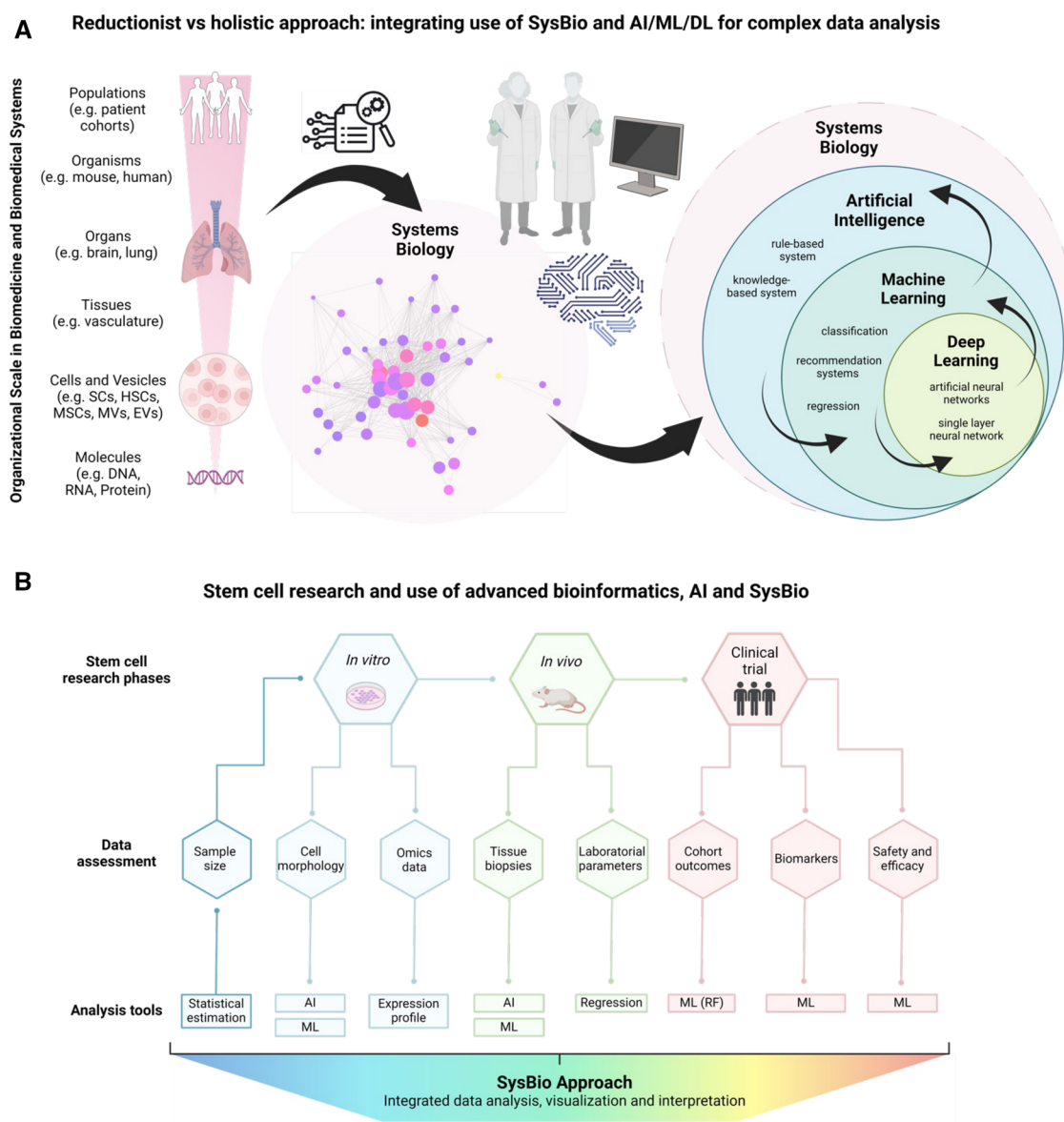


Figure 1. Advanced bioinformatics, systems biology, and artificial intelligence in stem cell research and therapeutics development.

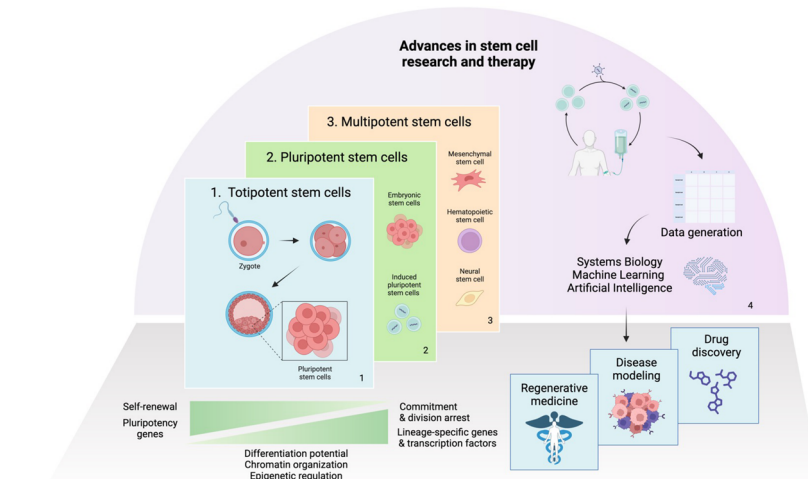
(A) Reductionist vs holistic approach: Integrating SysBio and AI/ML/DL for complex data analysis. Demonstration of the scale progression, from molecular studies to population-based research (cohorts), with enhancement of biomedical research through the integration of advanced bioinformatic data analysis, systems biology (SysBio), artificial intelligence (AI), machine learning (ML), artificial neural networks (ANNs), and deep learning (DL). (B) Use of advanced bioinformatics, AI and SysBio in preclinical and clinical stages of stem research and development. Datasets from all three phases in stem cell research (*in vitro*, *in vivo*, and pre-clinical studies) benefit from the integration of SysBio and AI analysis tools. These studies involve the use of AI or ML, employing their individual strengths in different settings. Typically, pre-clinical analysis includes omics data, cell morphology, tissue biopsies, laboratory parameters, but also analysis of clinical cohort outcomes and biomarkers with subsequent safety and efficacy analysis. In conclusion, the SysBio approach enables integrated data analysis, visualization, and interpretation of complex biological systems.

This current Concise Review Article provides an overview on the concepts underlying the stem cell SysBioAI analysis done in our previous Scoping Review.¹² The Scoping Review reported a concise timeline of publications on stem cell SysBioAI topic registered at PubMed between 2000 and 2024 and numerous examples of the reported daily experience and practical use cases. In short, the prior Scoping Review found an 8 to 10-fold global increase in SysBioAI-related research output between 2000 and 2021, with a 10-fold increase in AI-related publications since 2010, and an increasing shift from preclinical basic research to clinically oriented translational stem cell research. As summarized in this prior Scoping Review,¹² SysBioAI-supported data analysis is now common at all stages,

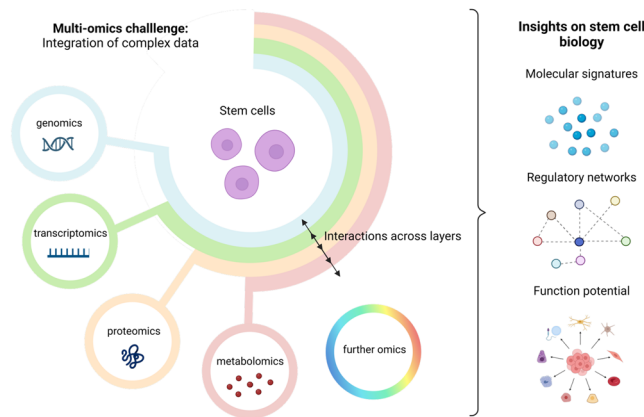
from molecular level to population level analysis, and from preclinical *in vitro* and *in vivo* research to clinical research (Figure 1A and B).

Based on the results and examples from this prior scoping review,¹² it becomes clear that SysBio, in particular in combination with AI technology, has already proven invaluable in boosting outcomes in many data-intensive research fields, particularly in enhancing the analysis of large-scale multi-omics and 3D-spatial-temporal data.^{12–17} Recent success in clinical development of chimeric antigen receptor (CAR)-based cell and gene therapies and advanced therapy medicinal products (CGTs/ATMPs) underscores the potential of SysBioAI analysis for more cost-effective CGT/ATMP development in the

A Opportunities for use of advanced bioinformatics in stem cell research and therapy



B Multi-omics challenge in integration of complex data



C Technical challenges in advanced data generation and analysis

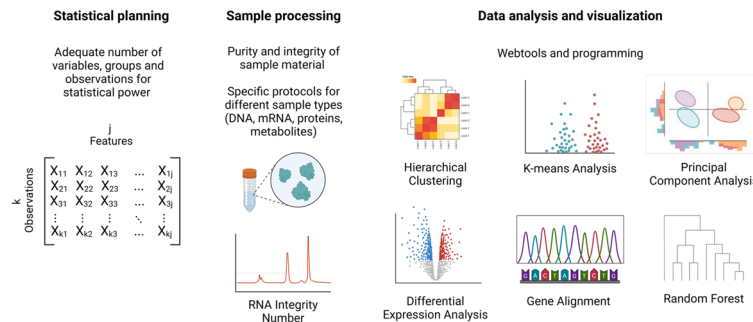


Figure 2. Opportunities and challenges in use of advanced bioinformatics stem cell research and therapeutics development.

(A) Opportunities in use of advanced bioinformatics in stem cell research and therapy. Popular stem cell types studied and employed in stem cell research and therapeutic settings grouped according to self-renewal capacity, lineage commitment and differentiation capacity (left panel). Multiple advances in SC research through integration of SysBio and AI analysis revolutionize regenerative medicine and cell therapy (right panel): This approach bears the potential to offer new treatment approaches, personalizing therapies based on SC with more effective and tailored treatments to meet individual patient needs. Promising tools such as SysBio, AI/ML drive the continuous development of advanced and personalized therapies. (1) Totipotent SCs: Originating from fertilization, play a unique and essential role in early embryonic development, limited to the initial phase. They can differentiate into any cell type, establishing the foundations for formation of all tissues and organs, (2) Pluripotent SCs: Located in the inner cell mass of the blastocyst are formed days after fertilization during embryogenesis, then, specialize into the three germ layers, eventually giving rise to a variety of tissues and organs in the developing organism. Induced pluripotent stem cells (iPSCs) are created through the reprogramming of adult cells, acquiring characteristics similar to embryonic pluripotent cells, and (3) Multipotent SCs: Are crucial for tissue regeneration and maintenance in multicellular organisms. Mesenchymal Stromal/Stem Cells (MSCs), originate from vascularized tissue sites like bone marrow and adipose tissue, and demonstrate potential differentiation into bone, fat, and cartilage cells, along with immunomodulatory properties. Hematopoietic Stem/Progenitor Cells (HSPCs) are found in the bone marrow, and are crucial for hematopoiesis, playing a vital role in transplants to treat blood diseases. Neural Stem Cells (NSCs) are found in the nervous system and have an essential role in neurogenesis, with promising therapeutic implications for neurological disorders. (B) Multiomics challenge with integration of complex data. Different types of “omics” technologies (left panel), including genomics, transcriptomics, proteomics, metabolomics and other types of high-throughput assessments of different biological components. A major challenge is the integration of multiple data types and sources across different omics layers in SysBio and AI based analysis. Omics studies provide crucial insights on several aspects of stem cell biology, including molecular signatures, regulatory networks, and functional properties, such as the differentiation potential of stem cells (right panel). (C) Technical challenges in advanced data generation and analysis. An adequate statistical planning of the type and number of features (j) and number of observations (k) in the dataset matrix, paired with optimized sample material and processing methods result in robust data. This can be analyzed and visualized with popular bioinformatics tools, available independently on various websites (webtools), or through computer programming in languages such as R and Python. Some examples include the hierarchical clustering, k-means analysis, principal component analysis, differential expression analysis, gene alignment and random forest.

future.^{18–27} Some recent practical examples include the use of AI in automated CAR-T cell manufacturing, safety and efficacy monitoring, and AI-driven predictive approaches for improving CAR-T cell signaling/prevent CAR-T cell exhaustion, to improve their “fitness” and clinical efficacy in patients.^{19,20,26} Overall, this approach may yield safer, more effective, and cost-efficient results.

Similarly to CAR-T-development, crucial factors that still hamper stem cell therapy development include: (1) High costs for manufacturing and eventual therapy, (2) Large product heterogeneity and partly insufficient definition, and (3) Incomplete understanding of the mechanism of action (MoA) and optimal clinical use.^{5–10,28–32} Another setback to clinical success may be the delayed adoption of advanced bioinformatics, SysBio, and AI tools, that may be needed to effectively harness the wealth of data generated during stem cell therapy development and adjunct clinical trial testing.^{13–16,33,34} Although thousands of clinical trials with diverse stem cell products have been performed in the past,^{4,8,12,35} advanced SysBioAI data analysis tools have only arrived rather late in the stem cell clinical trials arena compared to other fields.¹²

Opposed to traditional reductionist scientific methods that are based on “taking things apart,” SysBio is a holistic approach “to better understand the larger picture.” SysBio employs advanced computational and mathematical analysis to better model and understand complex biological systems in an integrative and holistic fashion.³⁶ This includes advanced data analysis and computational tools, for example artificial intelligence, machine learning, and deep learning (AI/ML/DL, respectively; Figure 1).^{12–16,36–40} Here, we provide a comprehensive review of the interesting interplay between SysBio and AI/ML/DL in stem cell biology and therapeutics development, exploring both advancements and challenges in this novel field to overcome potential bottlenecks and accelerate its effective implementation for the betterment of human health.

Potential of advanced bioinformatics, SysBio, and AI in stem cell research

Systems biology and artificial intelligence (SysBioAI) analysis approaches and tools have already proven its great value in other data-intense disciplines, such as immunology and infectious biology, that strive for a holistic understanding of disease process on different biological levels from molecular to population level analysis (Figure 1A, left panel).^{31,33,41–51} This includes: (1) Analysis on molecular level (eg, multi-omics and single-cell RNAseq),^{13,14,34,45,52,53} (2) Cellular, tissue, and organ analysis (eg, 3D-spatial and 4D-spatio-temporal analysis of cellular features in their 3D context, appreciating both space and time),^{54–57} and (3) Studies and modeling of complex “living” systems [eg, organ-on-chip (OoC) models, animal models, and human studies/clinical trials from early to advanced stage].^{58–60}

Novel SysBio data analysis tools are increasingly useful for the holistic analysis of datasets in medicine and stem cell research.^{13–16,36–38,61–66} Here, SysBio approaches go beyond traditional analysis by enabling the operator to better contextualize and understand the complexity of data using a “human-like artificial intelligence.” They combine the strength of conventional computational analysis/computing power with “self-learning” computational algorithms and neural networks that increasingly mimic complex non-linear learning processes of the human brain (Figure 1A, right panel).

These developments are fueled by the rapid progress in commonly available and affordable computer hardware and software, which can easily handle large datasets in conjunction with external servers.^{67,68} Developing novel bioinformatics software and data-processing tools has been another decisive step, integrating SysBio with AI/ML/DL tools in all phases of research (Figure 1B).^{12,42,43,47,50} Together, these innovations provide prospects for a rapid acceleration in medical breakthroughs and betterment of human health in the near future, including the stem cell field.

Stem cell research has emerged as one of the most dynamic fields in biomedical research, potentially transforming regenerative medicine, disease modeling, and drug discovery (Figures 1 and 2).^{2,5,65,69,70} Their unique cell-intrinsic properties make stem cells a valuable tool for studying and understanding complex biological processes, such as embryonic development, tissue regeneration, and cellular homeostasis, and thus, also for developing concomitant innovative therapeutic interventions.^{71–74} In this complex and ever-evolving research landscape, the convergence of two cutting-edge bioinformatic disciplines, SysBio and AI/ML/DL, is now paving the way for new transformative breakthroughs in stem cell science,^{13–16,39} including advancements in treating complex diseases like Parkinson’s, cardiovascular diseases, and diabetes.⁷³

A defining feature of stem cells is their ability to self-renew and differentiate into various specialized cell and tissue types depending on their state of potency (Figure 2A, left panel), thus, distinguishing between totipotent (eg, the fertilized egg), pluripotent (eg, ESCs and iPSCs), and more lineage-committed multipotent adult and perinatal stem cells (eg, HSCs, MSCs, and DSCs).^{5,71–77} Totipotent stem cells can differentiate into all cell types within the organism, while pluripotent stem cells (PSCs), for example embryonic and induced pluripotent stem cells (ESCs and iPSCs, respectively), can differentiate into a wide range of more specialized cell types. Multipotent stem cells, also known as adult stem and progenitor cells (ASCs/ASPCs), have a more restricted differentiation capacity according to their developmental origin,^{5,71–77} including hematopoietic stem and progenitor cells (HSCs/HSPCs), diverse adult and perinatal-tissue derived mesenchymal stem/stromal cells (eg, MSCs and DSCs),^{5,10,76–78} and other types of more lineage committed tissue-specific stem cells, for example skin stem cells, entailing subsets of basal and epidermal stem cells.^{3,74}

The complexity and dynamic nature of stem cells and the multitude of factors influencing their behavior pose a significant challenge and an opportunity for scientific exploration and clinical applications.⁶⁵ Stem cell niche-dependent control of self-renewal and asymmetric division, lineage commitment, division arrest, and differentiation potential is governed by the tight control of pluripotency and lineage-specific genes/transcription factors in conjunction with chromatin organization and epigenetic regulation of accessibility and transcription (Figure 2A, bottom left).^{74,75}

Traditional reductionist approaches often fall short in providing comprehensive insights into the intricate regulatory networks that govern stem cell fate and function.³³ Here, SysBio with its strong holistic perspective can play a vital role in understanding biological systems as a whole, encompassing diverse molecular interactions between genes, proteins, and cellular pathways (Figure 2A, right).^{31,33,41–51} This approach allows scientists to unravel the properties of complex biological systems

and provides a foundation for predicting cellular behavior with unprecedented accuracy.

SysBio- and AI-driven computational models can analyze/integrate data from diverse sources by offering the analytical power and pipelines required to handle the massive datasets needed to decipher intricate patterns. This can predict cell behavior and identify novel therapeutic targets.^{61,79–81} The synergy of SysBio and AI enhances our comprehension of the intricate nature of stem cells and propels the field toward personalized medicine and precision therapies (Figure 2A, right).⁸²

In turn, clinical trials benefit by becoming more targeted and patient-centric.⁸³ With AI-driven predictive models, researchers can better identify patient-specific responses to ensure more effective and safer treatments.^{10,79,84} By better understanding the genetic/epigenetic factors that govern stem cell responses, we are entering an era where regenerative medicine may be tailor-made for each patient.

Advanced bioinformatics: data generation and analysis with SysBio and AI

Advanced bioinformatics plays a pivotal role in better understanding complex biological systems although multiple challenges exist (Figure 2B and C).⁸⁵ The rise of high-throughput omics technologies, in particular the transition from bulk to single-cell RNA sequencing (scRNAseq) and other multi-omics approaches,⁸⁶ has transformed data generation to yield extensive datasets that require sophisticated analytical approaches across different layers of molecular analysis, such as genomics, transcriptomics, proteomics, metabolomics, and other methods (Figure 2B).⁸⁷ Proper integration of these omics data with the patients' clinical background can reveal new insights into the complex interplay between typical molecular signatures, characteristic modulation of decisive regulatory networks, and respective functional outcomes.^{14,62} Different SysBio applications are characterized by three major approaches: (1) Top-down, (2) Bottom-up, and (3) Middle-out intermediate approaches, as described previously.^{88–92}

Top-down approaches

Begin with high-throughput technologies to measure certain classes of macromolecules (eg, RNA with transcriptomics), which are then analyzed with robust algorithms.¹³ Crucial steps in this analysis include (Figure 2C): (1) statistical planning of sample collection and analysis, ideally before the initiation of the experimental comparison/clinical trial, to adjust for the sample size accordingly (eg, power analysis of a number of samples required to obtain a significant result; Figure 2C, left),^{93–95} (2) biological sample processing, requiring suitable starting material and fixation for later processing (eg, purity and integrity; Figure 2C, center), and (3) data analysis, visualization, and interpretation (eg, employing analytical tools; Figure 2C, right). Here, SysBio can be helpful in the integration of results from various analytical software tools and AI/ML/DL techniques to facilitate the classification and interpretation of results from complex data visualization approaches.

The last point, data analysis/visualization/interpretation, is often considered the most challenging step, since it employs multiple analytical tools that require an advanced understanding of bioinformatics software (Figure 2C, right),¹³ including: (1) hierarchical clustering (HCL) based on sample similarity,

(2) K-means analysis based on dividing variables into a user-selected number of groups, (3) principle component analysis (PCA) and t-distributed stochastic neighbor embedding (t-SNE) of high-dimensional data to reduce the number of dimensions for improved visualization and comprehension, (4) differential expression and peak comparison analysis to determine if a set of genes is statistically different between two conditions, (5) gene alignment and gene ontology (GO) terms enrichment analysis to align identified genetic signatures to certain biological process stored in data bases, and eventually, (6) random forest (RF) and support vector machine (SVM) analysis as novel machine learning approaches to conduct ranking and regression analysis, as outlined in the following sections.

Bottom-up and middle-out approaches

In contrast to going into more detail by breaking the whole system apart, bottom-up approaches form detailed models from subunits of data to simulate whole systems for modeling of physiological conditions (eg, genome-scale or flow simulations).^{89–92} This approach has been used for tissue-specific simulations, for example systems pharmacology, to model blood flow/tissue responses with simulation of drug distribution, safety, and efficacy.^{91,92,96–98} It is not always clear whether these models accurately reflect clinical conditions. Thus, comparative studies of both top-down and bottom-up approaches need to be carried out to compare their ability to predict the clinical safety and potency of certain drugs.^{89,91} To reduce the uncertainty in drug models, “intermediate or middle-out” approaches can be used, where top-down and bottom-up data analysis is combined.⁹² All these approaches have great potential to reduce the need for animal models and the length of clinical trials.

Strong need for statistical planning before starting your study

Before delving into the meticulous details of big data analytics, it is essential to evaluate key aspects that will greatly impact the final results, including the size of the dataset and an appropriate number of groups (variables and rows; Figure 2C).^{93–95,99} This decision significantly affects the subsequent selection of statistical tests, thereby shaping the methodological framework of the computational analysis.⁹³

For inferential statistics or AI/ML/DL approaches, it is important to remember that confidence intervals (CI) and/or *P*-value results are directly influenced by the sample size and data distribution of each variable. Here, various analysis methods exist to evaluate the suitability of a dataset for conducting robust statistical tests, including multiple linear regression, logistic regression, and the Wilcoxon-Signed-Rank test.^{100–102} Tools such as the G-Power and its contained testing methods serve as valuable tools for ensuring the dataset's adequacy and reliability for meaningful statistical analyses, thereby contributing to the overall rigor and validity of the research endeavor.^{50,103}

Summarizing, the dataset's structure, including the number of columns and rows, plays a vital role in the success of integrative big data and multi-omics analysis.⁴⁶ The multitude of columns—representing variables or features in the dataset—influences the choice of statistical methods (eg, regression analysis),^{50,104} which is particularly pertinent when exploring the relationships between different variables across different omics layers.⁴⁶ Simultaneously, the number of rows reflecting individual observations or samples becomes essential in

determining the statistical power and precision of the study. The meticulous alignment of the dataset composition with suitable statistical tests during the planning phase ensures a robust foundation for accurate analyses and meaningful interpretations, paving the way for a comprehensive understanding of stem cell characteristics through integrative multi-omics analysis.¹⁰⁵

Intricacies of biological sample preparation and data integration strategies

Sample collection and processing demand careful consideration of the specific requirements tailored to the different readout methods (Figure 2C, center).¹⁰⁶ Whether employing genomics, transcriptomics, proteomics, or metabolomics, each distinct method comes with its own set of intricate sample processing requisites.^{87,107–109} Sample processing kits for a combined analyte extraction/readout with different omics approaches in one workflow are available and can reduce sampling requirements.^{87,108} The use of reference samples should be considered in multicenter studies.¹⁰⁹ When pursuing single omics approaches, the intricate requirements of each platform must be considered early on in experimental and clinical studies to prevent sample collection- and processing-related artifacts/degradation that may falsify the correct representation of the sample intrinsic information. The precautions below may increase the chance of an optimal and reproducible readout of the original sample encoded information.

DNA is generally considered a fairly stable molecule at ambient conditions, thus, also allowing for DNA extraction from biospecimens resting for a substantial time under ambient conditions—preferentially assisted by sequence alignment to already processed complete datasets/genomes found in specialized online databases (eg, EMBL-EBI, BLAST, and NCBI nucleotide search and sequence analysis databases).¹¹⁰ Nonetheless, any genomic analysis methods—such as different generations of DNA sequencing techniques¹¹¹—will aim to extract nonfragmented high-quality and high-purity DNA considering the anticipated biological source material of interest. These principle steps are paramount to eventually yield an accurate representation of the desired genome upon sequence alignment in well-established online databases, thus, requiring the necessary attention of investigators to optimal isolation and purification techniques to prevent DNA degradation/fragmentation/contamination.¹¹²

In transcriptomics and gene expression analysis, stringent protocols are required to maintain the RNA integrity and transcript stability to prevent potential bias, for example RNase inhibitors and cold storage.¹¹³ On optimal extraction yield and molecular integrity, for example a lower cutoff of a RNA integrity number (RIN) value >7 is often used in bulk RNAseq approaches,¹¹⁴ of the different types of RNA species, such as messenger RNA (mRNA), micro RNA (miRNA), and long noncoding RNA (lncRNAs), to just name a few, is paramount—due to the high sensitivity of RNA to degradation by different types of ribonucleases (RNases, eg, endonucleases and exonucleases) and different intrinsic stability of different RNA species depending to their susceptibility to cell intrinsic/extrinsic DNA degradation systems/mechanisms.¹¹⁵

Proteomic and metabolomic methods often involve some form of cold-storage of the samples/biospecimens with variable temperature requirements depending on the sample type (eg,

storage of blood, plasma, serum, urine, or tissue samples at 4 °C fridges to –20 °C freezers for short term storage if no other means are available, or –80 to –150 degrees freezers for prolonged storage of samples with minimal degradation, or –196 degrees liquid or vapor-phase liquid nitrogen of cells to maintain their viability upon thawing).^{30,116} The initial cold storage of biospecimens is often followed either by direct analysis of sample aliquots with suitable assays (eg, Affymetrix, ELISA, OLINK, or HPLC/GC platforms),^{117–119} or by specialized analyte extraction procedures to capture the full spectrum of analytes in a sample.^{120,121}

Typical analytes include increasingly larger panels of cytokines, chemokines, growth factors, or small-sized metabolites indicative of biological processes, such as regenerative, immunomodulatory, or stem cell niche-related functions.^{122,123} For some sensitive analytes, analysis of fresh samples may be mandated, especially in the setting of routine analysis, as specified in the respective ISO guidelines.¹²⁴

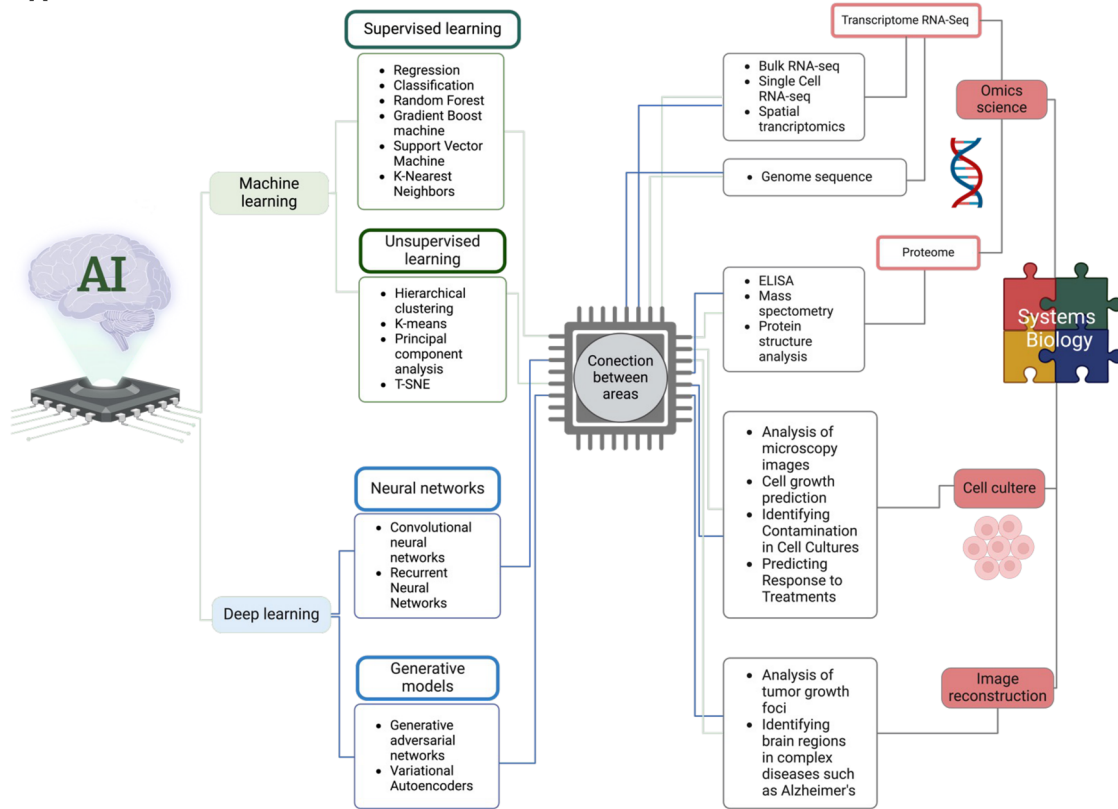
Apart from the detailed transcriptomic analysis of immune cells and other cellular subsets with scRNAseq profiling, the high-parametric analysis of pre-defined clusters of differentiation (CD) marker panels with both flow and mass cytometry (FACS and CyTOF, respectively), still forms the strong analytic foundation of cellular analysis in most research labs today.^{114,125–133} These profiling platforms allow for a solid readout of predefined surface and intracellular CD markers on different types of stem cells, immune cells, and therapeutic cell products according to their cell type intrinsic developmental differentiation pathways.¹²⁵ If the necessary requirements for biosample sourcing and subsequent analysis are met (eg, correct setting and speed of analysis, with adequate ambient temperature control and preservatives during shipping, and validated readout equipment with suitable standardized antibody panels, such as DuraClone panels), these analysis platforms may actually allow for a reproducible, standardized, and accredited/certified readout in specialized facilities/labs.^{114,125–129,133}

The aforementioned details of proper statistical planning, sample collection, and processing are tightly linked to the broader challenges faced in bioinformatics, particularly concerning correct data integration and representation strategies.^{134–137} The diverse nature of omics data requires innovative approaches to integrate information from the different analytic platforms cohesively, thus, requiring suitable strategies to harmonize the data from varied methodologies, ensuring a comprehensive and accurate representation of biological systems.⁸⁷ A correct match and representation of training and target datasets is essential for AI-based analysis.^{134–137} Thus, developing and adopting advanced computational techniques and standardized protocols for data generation and integration is imperative in overcoming the complexities inherent in the diverse array of omics datasets, ultimately enhancing the reliability and utility of biological insights derived from integrated analyses.^{109,138}

Bioinformatics tools for visualization and interpretation of SC data

Both stem cell research and its adjunct clinical trials generate vast datasets on different technology platforms, requiring sophisticated analytical tools for effective interpretation with integrated use of SysBio and AI (Figure 3A).^{12,16} Thus, nowadays, the effective use of advanced bioinformatics tools and

A Intersection of AI and SysBio in stem cell research and therapeutics development



B Typical machine learning algorithms employed in AI and SysBio analysis

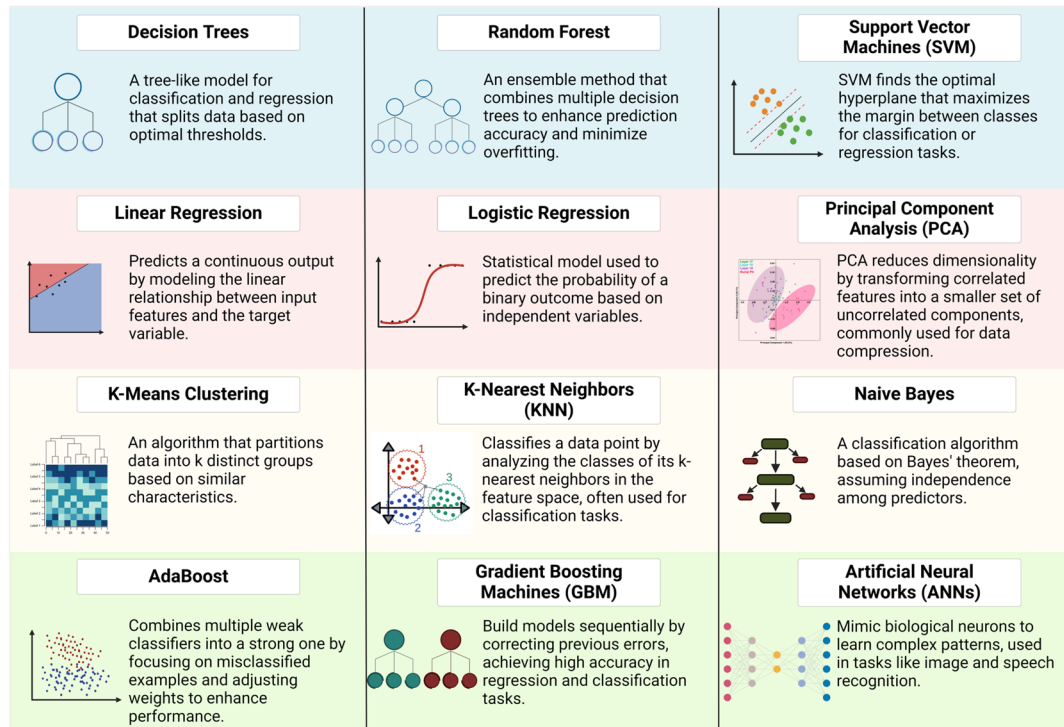


Figure 3. Increasing integrated use of AI and SysBio approaches.

(A) Intersection of AI and SysBio analysis in stem cell research and therapeutics development. The blue and green colors represent the AI/DL and AI/ML approaches, respectively. The blue and green lines represent connections that AI approaches can make to SysBio. The central square represents the connection between the approaches. (B) Typical machine learning algorithms and artificial neural networks analysis employed in SysBio and AI analysis. More detailed elaboration and practical examples can be found [Supplementary Tables S1-S5](#).

software plays a crucial role in visualizing these highly complex datasets, thereby offering researchers the means to explore and better understand the molecular signatures, differentiation potentials, and regulatory networks that govern stem cell behavior and to communicate their results.¹⁶

Some freely available powerful web tools for analyzing omics data include: (1) Morpheus bioinformatics for the simulation and integration of cell-based models with ordinary differential equations and reaction-diffusion systems,¹³⁹ (2) Network analyst to provide a systems-level data understanding with the integration of gene expression data with protein-protein interaction (PPI) networks,¹⁴⁰ and (3) The Galaxy Platform for accessible, reproducible and collaborative biomedical analysis,¹⁴¹ amongst others. These tools enable scientists to gain meaningful insights from the wealth of data through graphical representations, statistical analysis, and interactive platforms.

Furthermore, SysBio researchers often use diverse power tools to integrate two or more datasets from the commonly

used RNA sequencing approach.^{46,104,142,143} There are typically two different approaches to follow: (1) Consensus analysis,^{88,142} or (2) ComBat analysis for statistical correction of batch effects,¹⁴⁴ for example by employing the Network analysis web tool.⁴⁶ These web tools are easily accessible and widely used in SysBio analysis. Another option is to employ programming languages like R,¹⁴⁵ Python,¹⁴⁶ and Shell/Linux—GNU.¹⁴⁷ For both R and Python, open-source software packages are available at websites such as Bioconductor,¹⁴⁸ that can facilitate complex statistical analysis. In general, the motto—learning by doing—applies.

Within the realm of programming languages, sophisticated analyses with ML methods employs statistical algorithms that are based on mathematical optimization and prediction models (Figure 3B and Table 1). These learn from existing datasets to generalize and predict outcomes for previously unseen data. In turn, data mining focuses on discovering unknown properties in the data (knowledge discovery) through employing ML

Table 1. Use of machine learning algorithm and artificial neural networks.

Author and references	Algorithm	Type	Description	Biological application
Liu et al. ¹⁴⁹	Linear regression	Supervised	Predicts a continuous output (numerical value) based on the linear relationship between the input variables (features) and the target output.	Gene expression analysis, dose-response relationships
Ban et al. ¹⁵¹	Logistic regression	Supervised	Models binary or categorical outcomes by estimating probabilities using a logistic function, often applied to classification tasks.	Disease diagnosis and risk prediction, gene mutation and drug response prediction
Wu et al. ¹⁵³	Decision trees	Supervised	Constructs a tree-like model of decisions by splitting the dataset into subsets based on feature values, useful for both classification and regression.	Disease diagnosis and classification, gene function prediction
Imai et al. ¹⁵⁵	Random forest	Supervised	An ensemble learning method combining multiple decision trees to improve prediction accuracy and reduce overfitting.	Biomarker discovery, genome-wide association studies (GWAS), ecology and evolutionary biology
Lien et al. ¹⁵⁷	Support vector machines (SVM)	Supervised	Finds the optimal hyperplane that maximizes the margin between different classes for classification or regression tasks.	Protein structure and function prediction, transmembrane-protein topology prediction
Mota et al. ¹⁵⁹	K-nearest neighbors (KNN)	Supervised	Classifies a data point by analyzing the classes of its k-nearest neighbors in the feature space, often used for classification tasks.	Protein function prediction, medical image analysis
Wood et al. ¹⁶¹	Naive Bayes	Supervised	A probabilistic classifier that assumes independence between features, efficient for large datasets and suitable for classification problems.	Disease classification (eg, classifying cancer subtypes), microbial classification, text mining in bioinformatics, gene classification
Buggenthin et al. ¹⁶³	Artificial neural networks (ANNs)	Supervised	Mimics the structure of biological neurons to learn complex patterns in data, used for tasks like image recognition, speech recognition, and more.	Drug discovery, personalized medicine, medical data mining, metabolic network modeling
Katarzyna et al. ¹⁶⁴	Dense neural networks	Supervised	A fully connected neural network where each node in one layer is connected to every node in the subsequent layer, used for a wide range of tasks.	Gene expression analysis, predicting drug-target interactions in drug discovery using molecular data.
Khouj et al. ¹⁶⁶	K-means clustering	Unsupervised	Groups data into K clusters by minimizing the distance between points and their assigned cluster centers based on feature similarity.	Grouping similar gene expression profiles, image segmentation, clustering cells in single-cell RNA sequencing
Zimmermann et al. ¹⁶⁸	Principal component analysis (PCA)	Unsupervised	Reduces dimensionality by transforming correlated features into a smaller set of uncorrelated components, often used for data compression.	Visualizing gene expression data, reducing high-dimensional omics data
Taherkhani et al. ¹⁶⁹	AdaBoost (adaptive boosting)	Ensemble	Combines multiple weak classifiers into a strong one by focusing on misclassified examples and adjusting weights to improve model performance.	Improving cancer diagnosis from imaging, integrating diverse biological datasets
Rapakoulia et al. ¹⁷⁰	Gradient boosting machines (GBM)	Ensemble	Sequentially builds models by correcting the errors of previous models, commonly used for regression and classification with high accuracy.	Predicting disease risk factors and discovering censored survival outcomes, identifying key genetic drivers of diseases

methods, but with different goals. Thus, ML algorithms are described as “self-learning-algorithms” that perform tasks without explicit instructions.¹⁷²

Popular ML methods include (Figure 3B), for example (1) Regression Analysis (eg, multiple linear regression),¹⁰⁰ (2) Principal Component Analysis (PCA; to simplify the complexity of high-dimensional data while retaining trends and patterns),¹⁷³ (3) Random Forests (RF; a combination of tree predictors, where each tree depends on values of a random vector sampled independently and with the same distribution for all trees in the forest),¹⁷⁴ (4) Support Vector Machine (SVM learning—a powerful analysis tool for classification and subtyping of data),^{175,176} and (5) Bootstrap (eg, tests or resampling methods that work through assigning measures of accuracy to sample estimates, including measures such as bias, variance, confidence intervals).¹⁷⁷

Regression analysis provides insights into the relationships between variables and enables forecasting of trends by modeling how changes in one or more independent variables impact a dependent variable. Linear regression is used when the dependent variable is continuous and the relationship with independent variables is assumed to be linear, allowing for the prediction of future values based on this linear relationship. Binomial logistic regression is applied when the dependent variable is categorical with two possible outcomes. This method estimates the probability of a particular outcome by modeling the relationship between the independent variables and the odds of the outcome, using a logistic function to ensure predictions fall between 0 and 1.

Random Forest and Gradient Boosting Machines (GBM) are advanced ensemble methods that build on decision trees to enhance predictive accuracy. Random Forest constructs multiple decision trees using random subsets of the data and features, combining their predictions to improve robustness and reduce overfitting. Gradient Boosting Machines create a series of decision trees sequentially, where each new tree focuses on correcting the errors of the previous ones, thereby refining the model's performance through iterative improvements. Both methods leverage decision trees but differ in their approach to aggregating predictions and addressing model errors.

Support Vector Machines (SVMs) are classification and regression techniques that identify optimal hyperplanes that best separate different classes by maximizing the margin between data points of different classes, thus, ensuring a clear separation. For non-linear relationships, SVMs use kernel functions to transform the data into higher dimensions where a linear separation is possible. This approach is particularly effective in high-dimensional spaces and for complex classification problems.

K-means clustering partitions the data into a predefined number of clusters by iteratively assigning data points to the nearest cluster center and updating the centers to minimize the within-cluster variance. This method is effective for finding natural groupings in data but requires specifying the number of clusters in advance. Clustering encompasses various techniques for grouping data into clusters without predefined labels, including hierarchical clustering. These methods aim to uncover inherent structures in the data and can adapt to different types of cluster shapes and densities.

The ML algorithms enable the identification of complex patterns and non-linear relationships within datasets, thus,

providing valuable insights for more in-depth and predictive analyses.¹⁷⁸ Artificial neural networks and deep learning approaches have recently surpassed many previous ML approaches in performance.

Integrated SysBio-AI analysis in SC research and therapeutics development

Integrated AI and SysBio analysis has become a field of increasing interest in both stem cell research and concomitant clinical trial analysis (Figure 3).³⁶ Importantly, the integrative SysBio approaches offer the opportunity for connecting the molecular components to physiological functions and organismal phenotypes within both a single biological scale (eg, on type of tissue) but also among different scales (eg, cells, tissues, or organ systems), through quantitative reasoning, computational models and high-throughput experimental technologies such as scRNAseq analysis (Figure 4).³⁶ This allows SysBio researchers to better decode the flow of information from genes, proteins, and subcellular components, intracellular/extracellular pathways, up to the regulatory control of cell, tissue, organ, and organismal level functions, which are all highly relevant in stem cell research and respective clinical trial efforts.

Although the stem cell field has been one of the pioneers in multi-omics analysis, this has often been limited to detailed lab-research-centered preclinical, experimental, and modeling approaches in the past to better understand intrinsic stem cell behavior (eg, potency and fate decisions).^{14,34,53,61,179} Long et al. recently gave an elaborate update on scRNAseq method advancements in MSC research, demonstrating the rapid evolution of methodology in this field.³⁴ In exemplary fashion, Miura et al. employed scRNAseq to identify an LRRC75A-expressing subpopulation of MSCs involved in the secretion of vascular endothelial growth factor (VEGF) under ischemic conditions in their quest to explore phenotypic and functional heterogeneity in MSC products in relation to their angiogenic potency.⁵³ Furthermore, scRNAseq of “out-of-thaw MSCs” can identify differences related to both tissue source and inter-donor variations,¹⁷⁹ thus, widening the scope of earlier studies on the potential impact of cryopreservation and freeze-thawing for clinical MSC product performance.^{9,30,116,180–189}

By integrating different omics datasets with SysBio and advanced AI/ML/DL, it is possible to obtain a systemic view of the biological networks and stem cell (niche) interactions with either the local or the systemic environment, that exert impact on the potency, lineage differentiation, and the resulting progeny of stem cells.^{13,14,39,61,62,87} Tools such as Clusterprofiler (An R package for comparing biological themes among gene clusters),¹⁹⁰ EnrichR,¹⁹¹ and GSEA¹⁹² (Gene set enrichment analysis for interpreting genome-wide expression profiles) identify biological functions and cellular processes enriched in a set of genes associated with stem cell biology (Figure 4A).¹⁹³

There are various categories of stem cells with varying degrees of plasticity, but despite their distinct phenotypes and functions, all these varieties of stem cells share common fundamental characteristics in their biology, such as the activation and repression of biological processes linked to a set of stem cell related genes.^{65,75,193,194} Currently, scRNAseq stands out for its ability to provide more refined details about the characteristic expression profile in various types of stem cells.^{194,195}

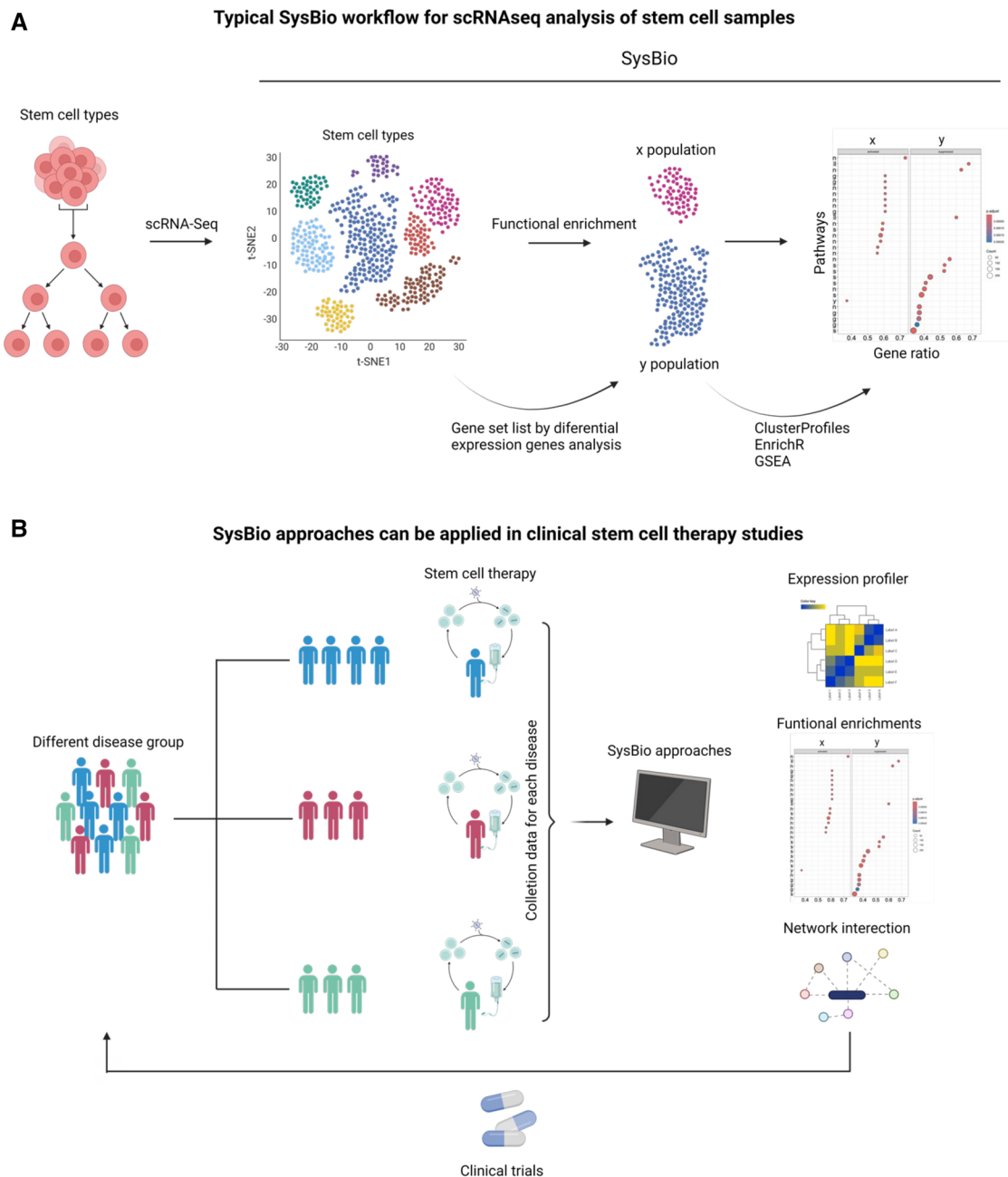


Figure 4. SysBio workflow applied to data from stem cell research and therapy development.

(A) Typical SysBio workflow for scRNAseq analysis of stem cell samples. Bioinformatic workflow in SysBio assisted scRNAseq analysis of stem cell samples: scRNAseq can reveal the gene expression profile in different types of stem cells and through SysBio it is possible to identify different sets of genes being differentially expressed in each type of stem cell. Furthermore, other SysBio methods such as functional enrichment can be applied to better understand biological processes. (B) SysBio approaches can be applied in clinical stem cell therapy studies. The application of integrative SysBio methods to data generated in clinical stem cell studies can generate a better understanding for different complex diseases and enable more substantiated guidance in clinical trials.

Recently, Barata et al. identified core functional modules of stem cells and potential novel “stemness genes” by integrating known stemness gene signatures.¹⁹⁴ The authors compared 21 distinct stemness signatures for humans and mice, identifying a significant overlap between both species, thus, allowing for the identification of integrated stemness signatures (ISS) comprised of genes frequently occurring among stemness signatures.

In a rat model of Parkinson’s disease (PD),¹⁹⁵ Tiklová et al. employed scRNAseq in combination with histological analysis

to characterize cellular diversity in intracerebral cellular grafts of transplanted human ESCs and fetal tissue. They demonstrated that neurons and astrocytes are major components of both graft types and identified a perivascular-like cell type in the stem cell-derived grafts.

Advanced tools like Harmony, LIGER, and Seurat (R toolkits for alignment of high-throughput scRNAseq datasets) offer powerful statistical approaches,^{196–198} for example including batch-effect correction methods, moderated estimation of fold change and dispersion for RNAseq data with DESeq2, thus,

simplifying the analysis and facilitating the visualization of complex data. These tools enable a more quantitative analysis focused on the individual strength of certain expression signatures rather than the mere presence of differential expression.¹⁹⁸ Here, batch-correction normalization and removal of batch effects is decisive for appropriate quantitative representation.

Furthermore, this approach allows for functional enrichment analysis, identifying distinct molecular signatures for each stem cell population and potential correlations between populations based on the expression profile. The synergy among various

disciplines in interpreting results from complex biological models can lead to more sophisticated future conclusions regarding stem cell-related outcomes,¹⁹⁴ thereby establishing a solid foundation for the identification of specific markers and potential therapeutic targets and paving the way for promising avenues in future clinical trials and advancements in biomedical research.¹⁹⁹

The ability to map critical signaling pathways allows for a better biological understanding of stem cell-related processes in patients and respective therapeutic interventions, and thus,

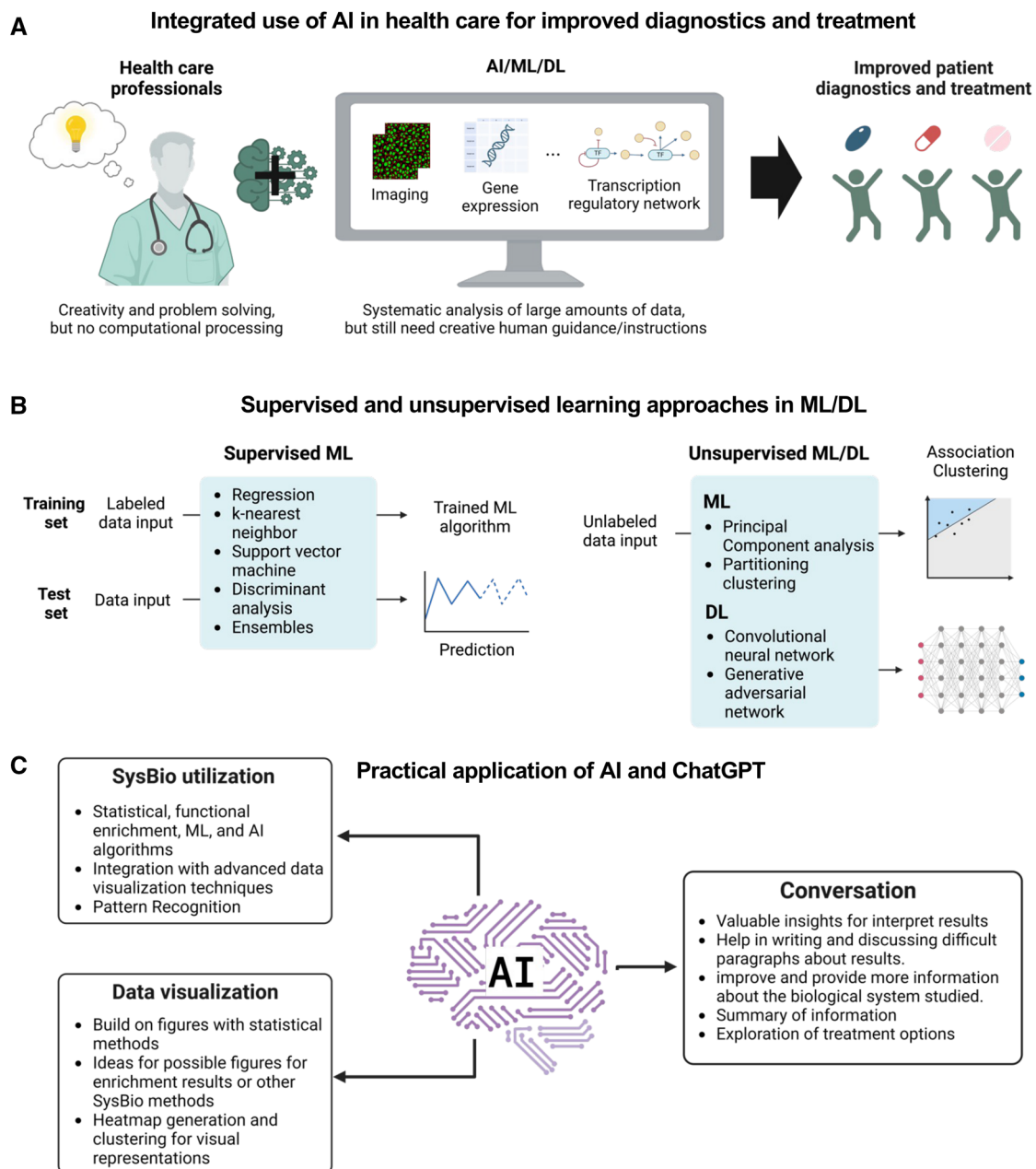


Figure 5. Enhancing interpretation of biomedical results through use of SysBio-AI synergy.

(A) Integrated use of AI in health care for improved diagnostics and treatment. The illustration depicts a health care scenario where human creativity, problem solving skills, and adjunct decision-making processes of health care professions are increasingly supported by the integrated use of SysBio and AI/ML/DL analysis tools, to yield improved outcomes for patients. (B) Supervised and unsupervised learning approaches in ML/DL. Simplified overview of the types of supervised and unsupervised AI/ML and AI/DL. Supervised and unsupervised artificial learning approaches utilize distinct statistical methods, including regression, clustering, and dimensionality reduction, to enhance the interpretation of large datasets. (C) Practical application of AI and ChatGPT. The increasing impact of AI use in biomedicine and health care is exemplified by the ChatGPT chatbot example, which can be employed to optimize statistical methods, result interpretation, and fostering constructive discussions to enhance treatments.

a more refined customization of clinical protocols.²⁰⁰ This not only bears the potential to enhance the treatment efficacy but it can also contribute to patient safety by minimizing the risks of adverse events. Moreover, a comprehensive understanding of the molecular characteristics of stem cell biology through SysBio analysis in the preclinical setting may facilitate the selection of patients for clinical trials, thus, maximizing the chances of success and expediting the translation of stem cell-based therapies into clinical practice (Figure 4B).⁸⁴

SysBio analysis in stem cell research and therapeutics development

Recent developments aim to employ SysBio for a better understanding of more complex physiological processes related to stem cell biology, for example organismal aging, regeneration, and therapeutic responses in a systemic context.^{62,66,69} McNamara et al. recognized the value of SysBio for studying complex biological systems and hypothesis generation in the context of regenerative medicine, seeking to replace or repair tissues with compromised function to improve their functionality,⁶² while Raj et al. reviewed the value of SysBio for drug development.⁶⁹ Both studies exemplify the adaptability of SysBio to fields related to stem cells, regenerative medicine, and pharmacology.

Researchers in data-intense biomedical fields (eg, immunology and infectious biology) are the early adopters of SysBio analysis, a common tool for analyzing complex data in these fields.^{31,44,46,48,49} This early adoption of SysBio analysis was driven by the need to better integrate human clinical trial results and patient data in proper relation to their adjunct mechanistic analysis in the laboratory to best ensure the practical relevance of their experimental research.⁷⁰ Stem cell research and adjunct clinical trials rapidly follow suit to better integrate more holistic SysBio analysis in their development pipelines to yield significant results.¹⁴

In the past decades, stem cell product-based clinical trials have matured from small concept studies on few patients (eg, phase I studies with typically no more than 9 patients in the 3 × 3 design for early-stage dose-escalation trials), to an increasing use of advanced phase IIb and III studies, to prove the safety and efficacy of a rather diverse panel of distinct stem cell products in multiple clinical indications for regional and international marketing approval.^{9,18,201–208} Such advanced studies provide the ideal training ground, complexity, and volume of data to explore the strengths of both SysBio and AI for advanced integrated data analysis (Figure 4B).

Examples of clinical stem cell trials with adjunct mechanistic studies include the approval of the AT-MSC-based cell product Cx601 for the treatment of complex perianal fistulas in Crohn's disease,^{201,209} systemic intravascular MSC therapy for the treatment of HSCT-related complications, for example acute Graft-versus-Host Disease (GvHD) and HSC engraftment failure,^{5,9,202–205} systemic MSC therapy for sepsis and pulmonary complications, for example acute lung injury (ALI), acute respiratory distress syndrome (ARDS) and coronavirus-induced disease 2019 (COVID-19),^{5,29,31,210,211} but also local intramuscular administration of MSC-like cells (PLX-PAD) for treatment of muscle injury following hip arthroplasty and critical limb ischemia (CLI).^{206–208}

These advanced studies are characterized by more complex clinical trial design, larger patient cohorts—often including hundreds of patients, and an abundance of product-, patient-, and biomarker-related data. These are collected during the trials to better understand the trial outcome and gain more value from these precious studies. This approach aims to overcome the “black box” effect of completing trials with a too narrow or singular focus on the primary endpoint, but to gain more knowledge from a diverse panel of secondary readouts and adjunct biomarker panels.^{125,128,129,212}

Thus, the need to better integrate preclinical (mechanistic) data with complex patient data and clinical trial outcomes has arrived in stem cell research and clinical trials. Here, SysBio approaches (in particular with the use of AI/ML/DL) offer a much deeper integration and understanding when integrating and analyzing the different obtained datasets (eg, cell product properties vs patient outcomes) to better understand the reasons for clinical trial success or failure, and to be able to further develop the established technology pipelines to successful therapeutics that eventually meet the high requirements of marketing approval.^{18,213–215}

From SysBio to multimodal AI-/ML-/DL-based analysis in stem cell research

To overcome limitations in high-throughput data analysis, SysBio analysis can benefit from the integration of AI/ML/DL approaches and, recently, multimodal learning as one of its most advanced disciplines (Figure 5A).^{39–43,45,47,216}

A particular challenge for AI considering the analysis of data is to use different modalities or sources of data at the same time in an integrated fashion, thus, requiring sophisticated computational tools and advanced processing power for multimodal data fusion and analysis (eg, see developments at the Fraunhofer-MEOS-Center).²¹⁶ Classically, this would apply to simultaneous/integrated analysis of different sensory data sources, such as text, picture, or audio, to represent/understand the necessary sensory environment to fulfill a dedicated task (eg, robotic-assisted surgery).²¹⁷

Multimodal learning can be of great value for integrated analysis of data from different sources in SysBio, for example in biomedicine, often different types of omics data, microscopy, phenotypic profiling, patient information. Although still in an early stage, in the future, multimodal AI may yield more contextualized analysis with a deeper understanding of the data in their larger biological context instead of evaluating parameters individually, for example integrated multiparametric analysis of (stem) cell phenotype, viability, metabolic activity, and transcriptomic profile.^{39,114}

This may be useful for: (1) Multiparametric imaging of cells/tissues in vitro, but also for automated imaging analysis of complex tissues, organs, and radiology data, (2) Combined analysis of gene expression, genome annotation, and protein binding, and (3) Identification of specific markers, transcription factors, and transcriptional regulatory networks etc.^{39,40,81} This may improve the speed, efficacy, and accuracy of diagnosis and treatment (Figure 5A).⁸¹ Nonetheless, most early ML/DL advances were/are still achieved with data originating from one source.

Kufel et al. recently gave an overview of distinct AI/ML/DL approaches with examples of their medical use (as summarized below and in Figure 5B).⁸¹ Kupfel et al. provide brief explanations of: (1) Commonly used ML models and methods (eg, k-Nearest Neighbor Algorithms, Linear Regression Algorithms, Logistic Regression Algorithms, Naïve Bayes Classifier Algorithms, Support Vector Machines, AdaBoost, and XGBoost) and (2) DL and neural networks that resemble the human brain [eg, Artificial Neural Networks (ANNs) with single layer ML, Deep Neural Networks (DNNs) with multiple hidden layers between input and output layer, and Convolutional Neural Networks (CNNs) consisting of input, spline, auxiliary, and output layers].⁸¹

Many of the ML/AI algorithms are designed for data regression or classification using available datasets, where an ML algorithm, supported by a mathematical model, generates predictions or specific decisions. The main types of ML are supervised and unsupervised learning (Figure 5B). Supervised ML assumes that the model has been trained on a similar dataset as the problem, consisting of input and output data, while unsupervised ML differs in its use of unannotated data, which were not previously labelled by humans nor algorithms. In supervised ML, once the model learns the relationship between the input and output, it can classify new unknown datasets and make predictions or decisions using classification (eg, binary two-class classification yes/no response) or regression (eg, prediction of continuous values based on input data to forecast future costs). In unsupervised ML, the model learns from input data without expected values. The algorithm focuses on grouping data based on their characteristics. The goal is to teach the machine to detect patterns and group the data without a single correct answer, using either clustering (eg, grouping the data based on similarities and differences) or association (eg, using specific relationships between data). In turn, DL and neural networks employ a much more complex layered architecture.⁸¹

The amount of data needed for ML training/testing depends on several factors, such as the type of ML problem (eg, supervised models need more data than unsupervised ones), model complexity (eg, the more complexity, the more data needed), data quality and accuracy (eg, amount of noise).²¹⁸ To get a first estimate, the “rule-of-thumb” may be used for smaller datasets, where you need at least 10 times as many data points as there are features in your dataset, to ensure a minimum quality input and avoid sample bias and underfitting.²¹⁹ Sample size and number of features are two different attributes that vary in different data types, for example images and omics data have different feature dimensions, but the number of samples depends on available data.

For AI-based analysis of larger datasets, more advanced statistical methods to adequately estimate the necessary sample size must be employed to guarantee optimal performance (see section strong need for statistical planning before starting your study),⁹⁹ including checks of dataset size, number of groups (features/variables and rows), data quality and robustness of analysis,^{50,93–95,104} using, for example multiple linear regression, logistic regression, Wilcoxon-Signed-Rank test with tools such as G-Power software.^{50,100–103} For ML, it is advised to first train with 70%-80% of the sample data and then test with 20%-30% of the volume, and in particular, a correct match and representation of both training and target datasets is essential.^{134–137}

Several strategies can be employed to reduce the amount of data needed for an ML model to optimize computational calculation/processing requirements, for example using PCA with recursive feature elimination (RFE) to identify and remove redundant features from a dataset.²¹⁸ Dimensionality reduction techniques, such as singular value decomposition (SVD) and t-distributed stochastic neighbor embedding (t-SNE), can lower the number of dimensions in a dataset while preserving important information. Synthetic data generation techniques like generative adversarial networks (GANs) can generate more training examples from existing datasets.

Examples of AI/ML/DL-based analysis in the stem cell and RegMed field

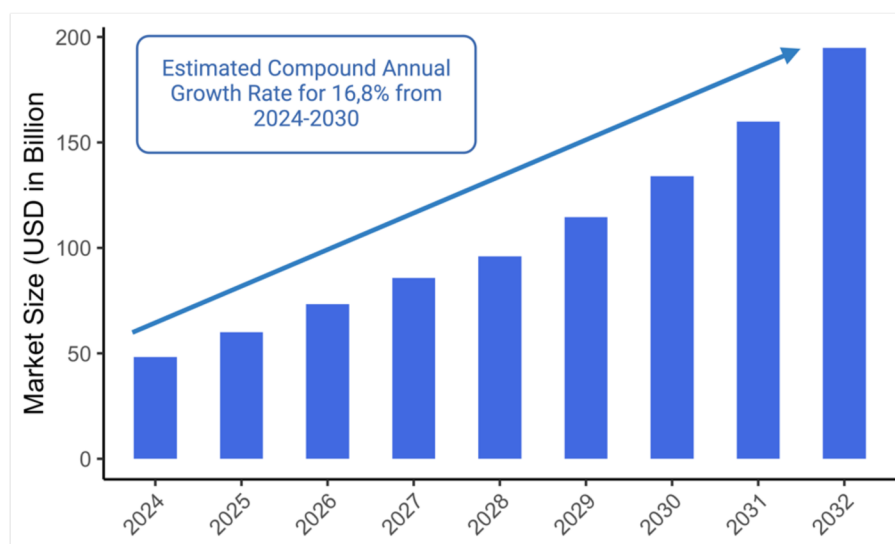
Next, we also wish to briefly introduce practical example(s) for using AI in stem cell research, therapeutics development, and regenerative medicine (RegMed). Others have already given more in-depth overviews of practical applications.^{81,82,220,221} As outlined in our recent Scoping Review,¹² the stem cell SysBioAI field has been expanding at near-exponential rates in recent years. A comprehensive summary of the past and current developments is provided in this adjunct paper.

Similar to using the SysBio approach, one of the most typical scenarios where AI/ML/DL can be of substantial use is the analysis of large datasets from stem cell experiments and clinical trials that quantify numerous molecular and environmental variables related to clinical safety and efficacy. Due to the large and intricate network of potential molecular interactions, such analysis presents considerable challenges for any analyst, clinician, scientist, or regular entity alike. This includes, among others, the causal analysis of stem cell tumorigenicity or cancerous progression *in vitro* (eg, as a safety measure before clinical use) and *in vivo* (eg, to detect host intrinsic or therapy-induced cancerous transformation), for example by using stem cell phenotypic and genotypic datasets, and analysis of characteristic or subtype-specific cancer markers.^{7,28,222–228}

More specialized questions include, for example hemocompatibility and hematotoxicity analysis of CAR-T and IV stem cell therapeutics.^{7,8,10,28,222,227,228} Other topics of increasing interest are neuropsychology and neuroimmunology in the interplay of general body health, immunological status, and thus, regenerative capacity with typical manifestations of poor brain health or mental illness.^{44,229,230} In the past, decisions were often based on individual biomarkers, while AI-based analysis also enables better identification of more complex molecular (autoantibody) network signatures (eg, multiple autoantibody-GPCR network interaction signatures instead of one).^{41–43}

In stem cell-based tissue engineering, recent innovations in advanced scaffold design, 3D-printing, and 3D-spatiotemporal cellular analysis have provided a new boost, but the stem cell-based replacement of damaged tissues and organs is still substantially held back by the enormous engineering, logistic, and analytical challenges required to grow functional replacements.^{231–233} Tissue and organ replacement are still constrained by donor shortages, but potential alternatives such as organ generation and replacement with tissue-specific stem/progenitor cells and iPSCs are increasingly explored.²³⁴ The iPSCs offer personalized treatment without ethical or rejection concerns,²³⁴ but considerable practical challenges considering the reprogramming and differentiation efficiency still exist, alongside outcome variability. Here, AI/ML/DL, plays a crucial role in

A Regenerative medicine market and adjunct medical need for RegMed therapies



B Bottleneck for new therapies: low likelihood of approval (LoA) from phase 1

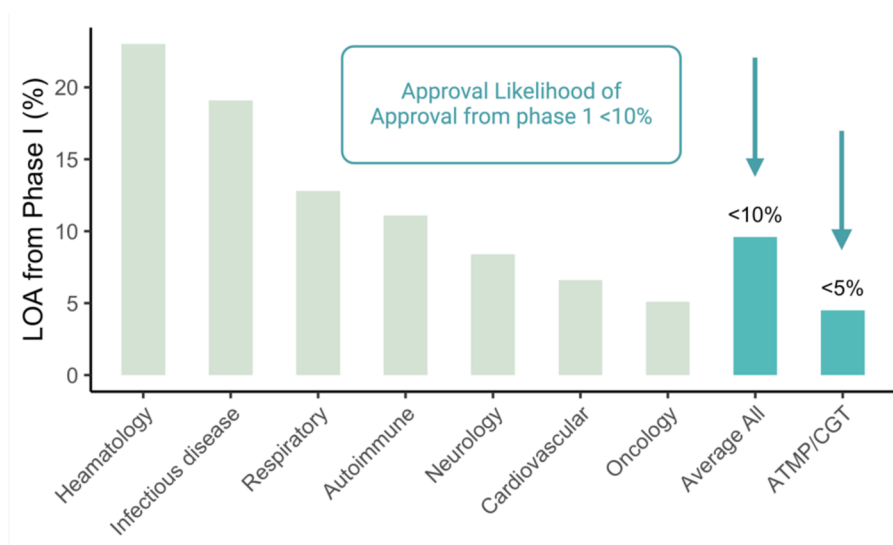


Figure 6. Medical need and bottleneck for new therapy development in RegMed.

(A) Regenerative medicine market and adjunct medical need for RegMed therapies. The global RegMed market size was valued at around USD 30 billion in 2023 and is expected to grow at a compound annual growth rate (CAGR) of around 16.79% from 2024 to 2030. RegMed entails a collection of techniques and technologies, including stem cell technologies, that addresses the high medical need to restore lost organ and tissue function to restore patient health, mobility, physical and mental well-being, and overall quality of life (QoL). The descriptive data were extracted from the website below, Grand-View-Research-Report: <https://www.grandviewresearch.com/industry-analysis/regenerative-medicine-market#>. (B) Bottleneck for new therapies: the low likelihood of approval from phase 1 studies. Individual likelihood of approval (LoA) from phase 1 (%) for the development of conventional therapeutics is depicted according to different classes of clinical indications, including the average LoA of <10% reported by Mullard et al. in 2016 (the average LoA of <10% is depicted here as reported by Mullard, while only a representative selection of indications is shown, which may have a particular need for regenerative therapeutics). An estimated lower average LoA of <5% for the development of more complex biological advanced therapy medicinal products (ATMPs) and cell and gene therapies (CGTs) is shown, based on the lower approval rates reported for ATMPs/CGTs. The highest LoAs for therapeutics are reported in infectious disease and hematology studies with approximately 18%-25% (data extracted from Mullard et al. 2016). Similarly, higher success rates are also typically achieved for ATMPs/CGTs in the hematology and oncology field, for example: for the development of various CAR-T therapeutics.

the quality control assessment of iPSCs, predicting induction and differentiation, and characterizing cellular functions (Figure 5B).²³⁵

Liu et al. recently published an example on using AI/ML/DL in the meta-analysis of published MSC research on cartilage

repair.¹⁴⁹ The authors aimed to better use the existing wealth of data to predict crucial impactors on MSC efficacy in cartilage repair to overcome inconsistencies in their therapeutic use. They first identified 36 articles and respective datasets in PUBMED to compile a small data base, including data from

15 clinical trials and 29 animal models. They employed artificial neural network (ANN) analysis (Figure 5B), to compare the relative strength of potential impactors, such as cartilage defect area, defect depth, type of damage, body weight, tissue source of therapeutic MSCs, or implanted cell numbers, to identify the factors that are most important for therapeutic outcome. Eventually, the authors identified defect area, defect depth, and implanted MSC number are the three most critical impactors on cartilage repair in their model, with a dose of 17-25 million MSC/defect providing optimal cartilage repair. Interestingly, the authors conclude that this approach/model can also be adapted for use in other clinical indications/treatments. The authors also emphasize that handling incomplete or missing data was a crucial aspect in the development of this model, thus, emphasizing the need for sufficiently sized datasets and, in this case, actual databases in the development of AI models.

AI-based advanced chatbots, such as the Generative Pre-Trained Transformer chatbot (ChatGPT), can prove useful in stem cell research and SysBio data visualization and interpretation (Figure 5C and Tables S1-S4).²³⁶ ChatGPT is of use in many ways, for example: (1) to improve data cleanup and coding readability/documentation, (2) to interpret data more effectively, generate visual summaries, and improve your writing, and (3) to prompt engineering/designs, stay on top of new developments and increase novelty in your stem cell research field.¹⁷ ChatGPT is a software engineering solution built on Transformer Architecture a DL neural network known for its ability to model complex dependencies in data sequences, such as text. It employs pre-trained transformers—GPTs, which are the foundation models (Tables S1-S4). ChatGPT can “understand” and generate text based on broader and more intricate contexts. It is important to note that, like any other ML/AI model, ChatGPT may occasionally produce minor errors and requires a subject matter expert to interact with it effectively and rectify any mistakes. As more and more AI/ML models stand out for their ability to process increasingly large volumes of generated data and efficiently identify complex patterns, their capacity to learn from past experience and dynamically adjust to new information provides the crucial key for more effective data handling and interpretation in the future. The combination of human creativity and intellect with the strength and efficacy of AI-based analysis can provide a strong dynamic partnership that combines the strengths of both AI/ML and human intellect to advance knowledge and further accelerate and drive innovation in the fields of stem cell research and therapeutics development.

Overcoming barriers to stem cell therapy development through the use of advanced SysBioAI bioinformatics and the framework for its implementation

Bodies wear out, tissue thins and tears, organs stop functioning, and cells lose their biological way; as a result we become ill and disabled.²³⁷ This has been our reality, until we entered the era of regenerative medicine (RegMed) and its advanced therapies, determined to restore diminished physiology and quality of life (QoL).²³⁷ Stem cell therapies are a crucial component of the CGTs/ATMPs currently under development to address this important unmet medical need.¹⁸ Integrating advanced bioinformatics, SysBio, and AI, in analyzing complex biological data

and advanced clinical trials is spurring promising next-generation patient-centric deep-medicine approaches.²²⁰ To develop these novel stem cell-based advanced therapies,¹⁸ rigorous clinical trials must be conducted to establish their safety and efficacy profile and to obtain regulatory approval for their clinical use, with systematic side studies to verify the underlying mechanism of action of the proposed therapeutic effect.^{5,7,9,18,28}

The growing medical need for regenerative therapies and adjunct growth potential of the regenerative medicine market (Figure 6A)²³⁸ is hampered by the low approval rates for new therapies and ATMPs/CGTs (Figure 6B). The likelihood of approval (LoA) rates of advanced clinical trials in relation to prior phase 1 trials are estimated to be less than 10% and 5%, respectively.^{239–241} The integration of advanced bioinformatic tools, such as the use of advanced SysBio and AI, bears the potential to positively counteract the high failure rates (typically around 90% failure rate from early stage to regulatory approval) and long timelines (typically 12–15 years, estimated reduction 3–4 years through AI-enabled developments) that are common in the conventional development of drugs, stem cell therapeutics, and for the development of the complex ATMPs/CGTs in particular.^{220,239–241}

Thus, to pave the way for faster and more efficient access to novel therapies, we here illustrate our “Iterative Circle of Refined Clinical Translation” (Figure 7A–C), including a framework for its implementation supported by SysBioAI analysis tools.¹² This iterative concept identifies at least three major needs to achieve better success rates in ATMPs/CGTs development: (1) Better use of open data and advanced data analysis tools (eg, integrated use of SysBio and AI), (2) Improved clinical trial design with option for iterative adjustments (eg, inclusion of exploratory biomarkers and informative secondary endpoints in addition to the primary endpoint), and (3) Iterative advancements/adjustment in biomedical engineering of product properties, clinical delivery and patient stratification, in case the study fails to meet its primary endpoints, but shows safety and promising efficacy in secondary endpoints.

Typical limitations to the implementation of such iterative learning approaches in clinical translation include: (1) Limitations in funding for complex studies, requiring suitable funding tools, such as larger EU-funded grants; (2) Lack of experience, training, planning and execution of advanced SysBioAI analysis in clinical trials; and (3) Resulting from points 1 and 2—General shortcomings in availability and sufficient quality of the primary data (as outlined in section advanced bioinformatics: data generation and analysis with SysBio and AI), data incompleteness and bias, need appropriate training datasets and batch correction, but also challenges resulting from model interpretation. However, the limitations in point 3 are mainly technical problems that can be overcome by improvements in points 1 and 2. Accordingly, the integrated use of open data and advanced data analysis tools, such as SysBioAI, is a key element often still lacking in many clinical trials and developmental approaches (Point 1 in Figure 7A). Its increased use could promote a better understanding and harnessing of the wealth of data (Figure 7B), which may be crucial to eventually overcoming existing barriers to effective clinical translation.^{13–16}

Many cell therapy trials still yield contradictory efficacy results, at least in part due to the immune safety and efficacy challenges in their preclinical assessment and progression to mature cell therapies, such as CAR-T and MSCs.^{8,10,227,242,243} Intriguingly, CAR-T cell therapy trials typically have a

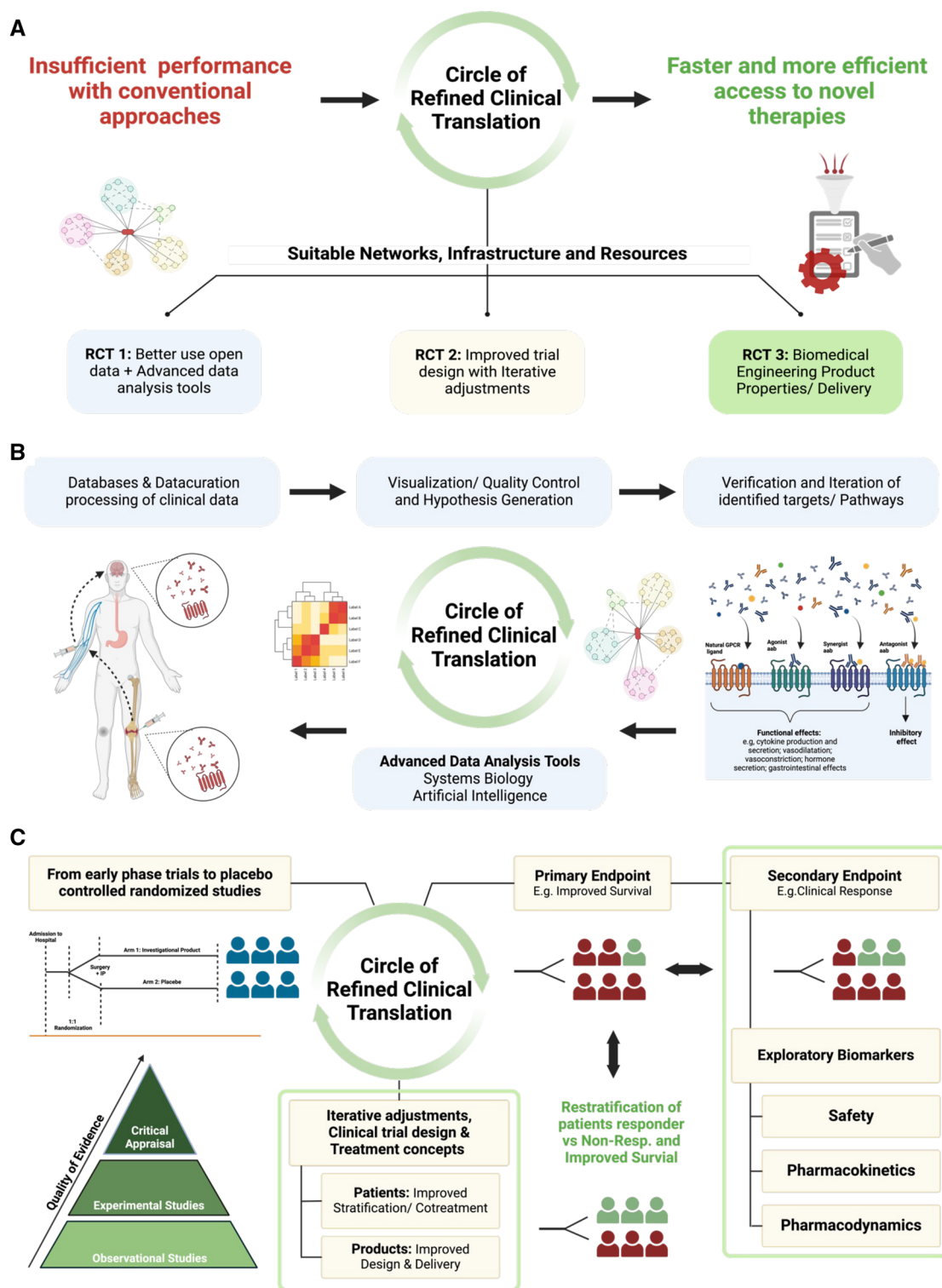


Figure 7. The iterative circle of refined clinical translation.

(A) Insufficient performance with conventional approaches mandates adjustments to yield faster and more efficient access to novel therapies, including: (1) better use of open data and advanced bioinformatics, (2) improved trials design with iterative adjustments, and (3) advancements in biomedical engineering to improve product properties and clinical delivery. (B) Measures to support better use of open data and advanced bioinformatics, such as SysBio and AI approaches. (C) Measures for improved clinical trial design, including stronger focus on secondary endpoints and exploratory biomarkers to support iterative adjustments in clinical trials design and treatment concepts for qualified restratification of patients according to therapy responder (green) and non-responder (red) status and improved survival.

well-defined goal of targeting and eradicating tumor cells in the oncological setting, or the autoreactive B-cells/plasma cells in the autoimmune setting,^{18,20,21,23–26} which facilitates the development of adequate potency assays and its analysis.^{20,26} In contrast, MSCs are utilized in a multitude of degenerative and inflammatory disorders, where the exact targets of therapy response and their preclinical/clinical potency assessment are often yet to be defined.^{204,243–248} Thus, in the case of MSCs, the medical and regulatory need to find an “assay matrix,” “surrogate markers,” or “patient biomarkers” that are indicative of therapy response and its MoA in patients is paramount.^{243,248–253} Novel profiling technologies combined with SysBioAI-analysis may allow rapid and adequate patient monitoring (Section advanced bioinformatics: data generation and analysis with SysBio and AI).

Examples include *in vitro* assays for down-modulation of characteristic pro-inflammatory mediators or donor cell subsets in mixed lymphocyte reactions (MLRs) and other similar assays, or modulation of cellular properties, such as apoptosis, that correlate with functional response in patients.^{248,254} Such *in vitro* markers should correlate with adjunct *in vivo* modulation of similar mediators and cell subsets in patients,^{248,254} for example skewing of proinflammatory activated T-cells and M1 myeloid cells into anti-inflammatory Tregs and M2 macrophages, with adjunct downmodulation of characteristic soluble mediators, such as CRP and IL-6. Given their clinical relevance, such biomarker and potency assays should be done by certified laboratories, to warrant the timeliness, robustness, and validity of such measurements.^{114,125–129} Importantly, SysBioAI supported analysis may be of advantage for the frequently required rapid and cost-effective decision-making processes in such a translational setting.^{227,242}

The MSCs have a multitude of properties and effector pathways that modulate responder immune cell populations in the host to collectively mitigate pathology.^{249–253} In addition, MSCs are often tested in advanced or refractory conditions (eg, steroid-resistant GvHD),²⁰⁴ where the host of pathobiology is versatile.²⁵⁵ Accordingly, one of the solutions that needs to be identified is the preemptive stratification of the patients who are likely to respond to MSC therapy.^{30,180,204,255–257} This may include similar approaches as done in therapy for Her2-positive breast cancer patients,^{258–260} which are selected for (immune)-therapy targeting Her2, while Her2 negative patients are treated otherwise. Thus, prior definition and subsequent assessment/screening of suitable patient inclusion/exclusion criteria with SysBioAI approaches could be a valuable asset to yield improved clinical trial outcomes.¹² The assessment of predefined patient parameters with subsequent analysis employing SysBioAI tools could provide the necessary accuracy and speed need for rapid clinical decision making, as already done in the assessment of histopathology, radiology, or in the kidney transplant setting to monitor potential rejection of allografts with an online multimodal diagnostic algorithm the “Banff Automation System.”^{12,261–271}

Considering MSC therapy, a practical example for inclusion/exclusion criteria may entail patient stratification according to their immune-aging status, with cutoffs for patient immune-age-related biomarkers.²⁷² MSC therapy for aGvHD seems to be more effective in children than adults,²⁷³ while severe COVID-19 patients of advanced age appear to be refractory to MSC treatment (patients below the age of 60 years appeared to respond better than patients above 60 years of age).³¹ This disparate outcome may be related to the low quantity or entire

lack or of suitable “host responder cells” that need to be present in the patient to respond to the application of MSC therapeutics, for example by being reprogrammed into beneficial anti-inflammatory repair cells.^{30,116,248,254}

In line with this argument, host biomarkers of potential therapy response to MSC applications could include the identification of certain biomarkers or circulating immune cell subsets in the patient indicative of “immunoaging” before and after therapy (eg, soluble aging markers such as GDF5 or circulating TEMRA T cells), or the induction of distinct immune cell subsets, that are indicative of therapy response (eg, Tregs and M2 myeloid cells macrophages).^{128,129} Either way, any SysBioAI-supported monitoring and evaluation of patient suitability for MSC treatment may require highly automated manufacturing, with suitable automated online-assessment of the required parameters during the manufacturing and release process, in conjunction with the assessment of *in vivo* parameters before and after therapy, to evaluate treatment safety and efficacy. This will require suitable technology development, standardization and regulation, to allow its comparable use in multi-centric/national clinical trials, to eventually demonstrate MSC product safety and efficacy.^{114,125–129}

Importantly, negative trial results are less likely to be published than positive outcomes, which confounds meta-analysis based only on published outcomes.^{274–276} This leads to a substantial loss of information considering non-published failed studies, and thus, loss in translational efficacy (Figure 7). The consequences are fewer learning outcomes from failed studies, waste of precious resources, and overall delayed access to novel therapies, with the financial loss carried by society as a whole. Thus, better use of advanced bioinformatics tools is important for an improved integrated understanding of large volumes of data, as is assessment, publishing, and access to information *per se*. This points out the role of “open data” policies and publishing policies for publicly and privately funded clinical studies alike to make required data more accessible. Eventually, one of the most crucial points remains to acquire suitable information through improved iterative clinical trial design (Figure 7C).

Conclusions, limitations, and outlook

Improvements in analytical capacity are crucial to overcoming existing barriers to effective clinical translation in stem cell therapy development and deployment as mature, approved clinical products. However, integrating SysBioAI-based approaches in analyzing biological data and clinical trials goes well beyond expanding analytical capabilities. The increasing synergy of SysBioAI actually supports a much more integrated understanding of complex multilayered information on multiple parameters from different data sources, typical for stem cell research, therapeutics development, and adjunct clinical trials. This synergy also enables a more patient-centric view even in the large, randomized trials necessary for therapy approval.

Deep-profiling of cell therapies’ phenotypic and functional properties with adjunct assessment of multiple secondary endpoints and biomarkers in patients is crucial to provide the necessary database for integrated SysBioAI-assisted analysis. In turn, this effort assists to better understand the clinical safety profile, pharmacokinetics, and pharmacodynamics of these novel therapeutics. This mandates careful planning with an integrated readout and analysis concept. Typical limitations include shortcomings in availability and quality of primary data, data

incompleteness and bias, with a need for batch correction and appropriate training datasets, but also challenges resulting from model interpretation. If data assessment is planned well in advance to avoid such limitations, iterative adjustments in clinical trial design and treatments concepts, based on the identification of favorable product and patient parameters with restratification of patients according to therapy responder and non-responder status (or more stringent parameters, such as survival), may be a suitable way to support evidence-based iterative improvements to therapy performance, and thus, successful outcomes and product development in the long-run.

Here, we provide the necessary background, explanations, and structure for non-experts in advanced bioinformatics and SysBioAI-based analysis to familiarize themselves with these important new developments and to plan the next steps to implement such technology in their manufacturing and clinical trial workflow.¹² Future discussion of interest may include more detailed discussion on related topics, such as: (1) Emerging technologies (including organoids and 3D-bioprinting single cell and spatial transcriptomics for capturing cellular heterogeneity, and integration of wearable devices or real-time monitoring technologies with AI for personalized therapy); (2) Advanced Considerations on Data Challenges (including handling of missing, incomplete, or noisy data, which are common in omics and clinical datasets, standardization issues in data collection from multicenter trials, ethical and logistical concerns regarding data sharing and open data policies, computational bottlenecks in multimodel AI analysis, risk of algorithmic bias and overfitting in small, diverse patient cohorts, and challenges in the interpretation of high-dimensional data from poorly annotated datasets); and (3) Considerations on Different Stem Cell Types (including how SysBioAI-analysis can be tailored for specific stem cell types, such as MSCs, iPSCs, and ESCs, which each have their unique challenges).

Author contributions

Guido Moll was invited to contribute the review article to the journal, first conceived, designed, and supervised the writing of the review article and the final version of the manuscript, and is the corresponding author. All listed authors listed have made a substantial and intellectual contribution to the work, contributed to the manuscript writing, revisions, read and approved the submitted version.

Supplementary material

Supplementary material is available at *Stem Cells Translational Medicine* online.

Funding

G.M.'s contributions were made possible by funding from the German Federal Ministry for Education and Research (BMBF) and German Research Foundation (DFG; projects Nephroprotection #394046635, subproject A03, as part of CRC 1365, and EXPAND-PD; CA2816/1-1 and IMMME) and through the BIH Center for Regenerative Therapies (BCRT) and Berlin-Brandenburg School for Regenerative Therapies (BSRT, GSC203), respectively, and in part by the European Union's Horizon 2020 Research and Innovation Program under grant agreements No 733006 (PACE) and 779293 (HIPGEN) and 754995 (EU-TRAIN) and 101095635 (PROTO). J.N.U. and

T.S. received funding from Charité and M.H.B. (G.M./I.H.). We acknowledge financial support from the Open Access Publication Fund of Charité Universitätsmedizin Berlin and the DFG. O.C.M.'s, D.L.M.F.'s, and I.S.F.'s contributions were made possible by The São Paulo Research Foundation (FAPESP 2018/18886-9, 2020/01688-0, and 2020/07069-0 to O.C.M. and 2020/16246-2 and 2023/133356-0 to D.L.M.F. and 2023/07806-2 to I.S.F.) and the National Council for Scientific and Technological Development (CNPq) Brazil (Grant: 309482/2022-4 to OCM). J.N.U. was supported by the Coordination of Superior Level Staff Improvement under Academic Excellence Program (CAPES/PROEX; Ref. No. 88887.917898/2023-00) and the German Academic Exchange Service (DAAD; Ref. No. 91898528). O.R. was supported by grants from the Swedish Cancer Society and was a recipient of a Distinguished Professor Award from Karolinska Institutet.

Conflicts of interest

The authors declare that the research was conducted without any commercial or financial relationships that could potentially create a conflict of interest.

References

1. Pittenger MF, Mackay AM, Beck SC, et al. Multilineage potential of adult human mesenchymal stem cells. *Science*. 1999; 284:143-147.
2. Pittenger MF, Discher DE, Péault BM, Phinney DG, Hare JM, Caplan AI. Mesenchymal stem cell perspective: cell biology to clinical progress. *NPJ Regen Med*. 2019;4:22. <https://doi.org/10.1038/s41536-019-0083-6>
3. De Luca M, Aiuti A, Cossu G, Parmar M, Pellegrini G, Robey PG. Advances in stem cell research and therapeutic development. *Nat Cell Biol*. 2019;21:801-811. <https://doi.org/10.1038/s41556-019-0344-z>
4. Zakrzewski W, Dobrzyński M, Szymonowicz M, Rybak Z. Stem cells: past, present, and future. *Stem Cell Res Ther*. 2019;10:68. <https://doi.org/10.1186/s13287-019-1165-5>
5. Ringdén O, Moll G, Gustafsson B, Sadeghi B. Mesenchymal stromal cells for enhancing hematopoietic engraftment and treatment of graft-versus-host disease, hemorrhages and acute respiratory distress syndrome. *Front Immunol*. 2022;13:839844. <https://doi.org/10.3389/fimmu.2022.839844>
6. Galipeau J, Sensebe L. Mesenchymal stromal cells: clinical challenges and therapeutic opportunities. *Cell Stem Cell*. 2018;22:824-833. <https://doi.org/10.1016/j.stem.2018.05.004>
7. Levy O, Kuai R, Siren EMJ, et al. Shattering barriers toward clinically meaningful MSC therapies. *Sci Adv*. 2020;6:eaba6884. <https://doi.org/10.1126/sciadv.aba6884>
8. Moll G, Ankrum JA, Kamhiel-Milz J, et al. Intravascular mesenchymal stromal/stem cell therapy product diversification: time for new clinical guidelines. *Trends Mol Med*. 2019;25:149-163. <https://doi.org/10.1016/j.molmed.2018.12.006>
9. Moll G, Hoogduijn MJ, Ankrum JA. Editorial: safety, efficacy and mechanisms of action of mesenchymal stem cell therapies. *Front Immunol*. 2020;11:243. <https://doi.org/10.3389/fimmu.2020.00243>
10. Moll G, Ankrum JA, Olson SD, Nolte JA. Improved MSC minimal criteria to maximize patient safety: a call to embrace tissue factor and hemocompatibility assessment of MSC products. *Stem Cells Transl Med*. 2022;11:2-13. <https://doi.org/10.1093/stcltm/szab005>
11. Dolgin E. Stealthy stem cells to treat disease. *Nature*. 2024. <https://doi.org/10.1038/d41586-024-00590-y>
12. Silva-Sousa T, Usuda JN, Al-Arawe N, et al. The global evolution and impact of systems biology and artificial intelligence in stem

- cell research and therapeutics development: a scoping review. *Stem Cells*. 2024;42:929-944. <https://doi.org/10.1093/stmcls/sxae054>
13. Bian Q, Cahan P. Computational tools for stem cell biology. *Trends Biotechnol*. 2016;34:993-1009. <https://doi.org/10.1016/j.tibtech.2016.05.010>
 14. Kinney MA, Vo LT, Frame JM, et al. A systems biology pipeline identifies regulatory networks for stem cell engineering. *Nat Biotechnol*. 2019;37:810-818. <https://doi.org/10.1038/s41587-019-0159-2>
 15. Cahan P, Treutlein B. A conversation with ChatGPT on the role of computational systems biology in stem cell research. *Stem Cell Rep*. 2023;18:1-2. <https://doi.org/10.1016/j.stemcr.2022.12.009>
 16. Del Sol A, Jung S. The importance of computational modeling in stem cell research. *Trends Biotechnol*. 2021;39:126-136. <https://doi.org/10.1016/j.tibtech.2020.07.006>
 17. Heidt A. AI for research: the ultimate guide to choosing the right tool. *Nature*. 2025;640:555-557. <https://doi.org/10.1038/d41586-025-01069-0>
 18. Goldsobel G, von Herrath C, Schlickeiser S, et al. RESTORE survey on the public perception of advanced therapies and ATMPs in Europe-why the European Union should invest more! *Front Med (Lausanne)*. 2021;8:739987. <https://doi.org/10.3389/fmed.2021.739987>
 19. Hort S, Herbst L, Bäckel N, et al. Toward rapid, widely available autologous CAR-T cell therapy - artificial intelligence and automation enabling the smart manufacturing hospital. *Front Med (Lausanne)*. 2022;9:913287. <https://doi.org/10.3389/fmed.2022.913287>
 20. Bäckel N, Hort S, Kis T, et al. Elaborating the potential of artificial intelligence in automated CAR-T cell manufacturing. *Front Mol Med*. 2023;3:1250508. <https://doi.org/10.3389/fmmed.2023.1250508>
 21. Elsallab M, Maus MV. Expanding access to CAR T cell therapies through local manufacturing. *Nat Biotechnol*. 2023;41:1698-1708. <https://doi.org/10.1038/s41587-023-01981-8>
 22. Schett G, Mackensen A, Mougiakakos D. CAR T-cell therapy in autoimmune diseases. *Lancet*. 2023;402:2034-2044. [https://doi.org/10.1016/S0140-6736\(23\)01126-1](https://doi.org/10.1016/S0140-6736(23)01126-1)
 23. Cappell KM, Kochenderfer JN. Long-term outcomes following CAR T cell therapy: what we know so far. *Nat Rev Clin Oncol*. 2023;20:359-371. <https://doi.org/10.1038/s41571-023-00754-1>
 24. Passweg JR, Baldomero H, Ciceri F, et al. Hematopoietic cell transplantation and cellular therapies in Europe 2022. CAR-T activity continues to grow; transplant activity has slowed: a report from the EBMT. *Bone Marrow Transplant*. 2024;59:803-812. <https://doi.org/10.1038/s41409-024-02248-9>
 25. Cadinanos-Garai A, Flugel CL, Cheung A, Jiang E, Vaissé A, Abou-El-Enin M. High-dimensional temporal mapping of CAR T cells reveals phenotypic and functional remodeling during manufacturing. *Mol Ther*. 2025;33:2291-2309. <https://doi.org/10.1016/j.ymthe.2025.04.006>
 26. Qiu S, Chen J, Wu T, et al. CAR-toner: an AI-driven approach for CAR tonic signaling prediction and optimization. *Cell Res*. 2024;34:386-388. <https://doi.org/10.1038/s41422-024-00936-1>
 27. Di Cerbo V, Song HW, Herbst L, et al. Artificial intelligence, machine learning, and digitalization systems in the cell and gene therapy sector: a guidance document from the ISCT industry committees. *Cytotherapy*. 2025;27:903-909. <https://doi.org/10.1016/j.jcyt.2025.05.003>
 28. Caplan H, Olson SD, Kumar A, et al. Mesenchymal stromal cell therapeutic delivery: translational challenges to clinical application. *Front Immunol*. 2019;10:1645. <https://doi.org/10.3389/fimmu.2019.01645>
 29. Moll G, Drzeniek N, Kamhieh-Milz J, Geissler S, Volk HD, Reinke P. MSC therapies for COVID-19: importance of patient coagulopathy, thromboprophylaxis, cell product quality and mode of delivery for treatment safety and efficacy. *Front Immunol*. 2020;11:1091. <https://doi.org/10.3389/fimmu.2020.01091>
 30. Cottle C, Porter AP, Lipat A, et al. Impact of cryopreservation and freeze-thawing on therapeutic properties of mesenchymal stromal/stem cells and other common cellular therapeutics. *Curr Stem Cell Rep*. 2022;8:72-92. <https://doi.org/10.1007/s40778-022-00212-1>
 31. Couto PS, Al-Arawe N, Filgueiras IS, et al. Systematic review and meta-analysis of cell therapy for COVID-19: global clinical trial landscape, published safety/efficacy outcomes, cell product manufacturing and clinical delivery. *Front Immunol*. 2023;14:1200180. <https://doi.org/10.3389/fimmu.2023.1200180>
 32. Hackel A, Vollmer S, Bruderek K, Lang S, Brandau S. Immunological priming of mesenchymal stromal/stem cells and their extracellular vesicles augments their therapeutic benefits in experimental graft-versus-host disease via engagement of PD-1 ligands. *Front Immunol*. 2023;14:1078551. <https://doi.org/10.3389/fimmu.2023.1078551>
 33. Wanjek C. *Systems Biology as Defined by NIH*. NIH Intramural Research Program. Updated May 2, 2022. <https://irp.nih.gov/catalyst/19/6/systems-biology-as-defined-by-nih>
 34. Long Q, Zhang P, Ou Y, Li W, Yan Q, Yuan X. Single-cell sequencing advances in research on mesenchymal stem/stromal cells. *Hum Cell*. 2024;37:904-916. <https://doi.org/10.1007/s13577-024-01076-9>
 35. Kabat M, Bobkov I, Kumar S, Grumet M. Trends in mesenchymal stem cell clinical trials 2004-2018: is efficacy optimal in a narrow dose range? *Stem Cells Transl Med*. 2020;9:17-27. <https://doi.org/10.1002/sctm.19-0202>
 36. Tavassoly I, Goldfarb J, Iyengar R. Systems biology primer: the basic methods and approaches. *Essays Biochem*. 2018;62:487-500. <https://doi.org/10.1042/ebc20180003>
 37. Jakhar D, Kaur I. Artificial intelligence, machine learning and deep learning: definitions and differences. *Clin Exp Dermatol*. 2020;45:131-132. <https://doi.org/10.1111/ced.14029>
 38. Edley. *Artificial Intelligence vs. Machine Learning vs. Deep Learning: What Is the Difference?* 2024. <https://edley.de/insights/unterschied-artificial-intelligence-machine-deep-learning/>
 39. Mazalan M, Do T-D, Zaman WSWK, Ramlan EI. Machine learning approaches for stem cells. *Curr Stem Cell Rep*. 2023;9:43-56. <https://doi.org/10.1007/s40778-023-00228-1>
 40. Chen RJ, Lu MY, Williamson DFK, et al. Pan-cancer integrative histology-genomic analysis via multimodal deep learning. *Cancer Cell*. 2022;40:865-878.e866. <https://doi.org/10.1016/j.ccell.2022.07.004>
 41. Cabral-Marques O, Marques A, Giil LM, et al. GPCR-specific autoantibody signatures are associated with physiological and pathological immune homeostasis. *Nat Commun*. 2018;9:5224. <https://doi.org/10.1038/s41467-018-07598-9>
 42. Cabral-Marques O, Moll G, Catar R, et al. Autoantibodies targeting G protein-coupled receptors: an evolving history in autoimmunity. Report of the 4th international symposium. *Autoimmun Rev*. 2023;22:103310. <https://doi.org/10.1016/j.autrev.2023.103310>
 43. Cabral-Marques O, Schimke LF, Moll G, et al. Advancing research on regulatory autoantibodies targeting GPCRs: insights from the 5th international symposium. *Autoimmun Rev*. 2025;24:103855. <https://doi.org/10.1016/j.autrev.2025.103855>
 44. Cabral-Marques O, Bastos V, Pacheco da Silva V, et al. Neuroimmunology of rabies: new insights into an ancient disease. *J Med Virol*. 2023;95:e29042.
 45. Sotzny F, Filgueiras IS, Kedor C, et al. Dysregulated autoantibodies targeting vaso- and immunoregulatory receptors in post COVID syndrome correlate with symptom severity. *Front Immunol*. 2022;13:981532. <https://doi.org/10.3389/fimmu.2022.981532>
 46. Prado CADs, Fonseca DLM, Singh Y, et al. Integrative systems immunology uncovers molecular networks of the cell cycle that stratify COVID-19 severity. *J Med Virol*. 2023;95:e28450. <https://doi.org/10.1002/jmv.28450>
 47. Baiocchi GC, Vojdani A, Rosenberg AZ, et al. Cross-sectional analysis reveals autoantibody signatures associated with COVID-19 severity. *J Med Virol*. 2023;95:e28538. <https://doi.org/10.1002/jmv.28538>

48. Usuda JN, Plaça DR, Fonseca DLM, et al. Interferome signature dynamics during the anti-dengue immune response: a systems biology characterization. *Front Immunol.* 2023;14:1243516. <https://doi.org/10.3389/fimmu.2023.1243516>
49. Plaça DR, Fonseca DLM, Marques AHC, et al. Immunological signatures unveiled by integrative systems vaccinology characterization of dengue vaccination trials and natural infection. *Front Immunol.* 2024;15:1282754. <https://doi.org/10.3389/fimmu.2024.1282754>
50. Fonseca DLM, Filgueiras IS, Marques AHC, et al. Severe COVID-19 patients exhibit elevated levels of autoantibodies targeting cardiolipin and platelet glycoprotein with age: a systems biology approach. *NPJ Aging.* 2023;9:21. <https://doi.org/10.1038/s41514-023-00118-0>
51. Fonseca DLM, Jäpel M, Gyamfi MA, et al. Dysregulated autoantibodies targeting AGTR1 are associated with the accumulation of COVID-19 symptoms. *NPJ Syst Biol Appl.* 2025;11:7. <https://doi.org/10.1038/s41540-025-00488-z>
52. Vandereyken K, Sifrim A, Thienpont B, Voet T. Methods and applications for single-cell and spatial multi-omics. *Nat Rev Genet.* 2023;24:494-515. <https://doi.org/10.1038/s41576-023-00580-2>
53. Miura T, Kouno T, Takano M, et al. Single-Cell RNA-Seq reveals LRRC75A-expressing cell population involved in VEGF secretion of multipotent mesenchymal stromal/stem cells under ischemia. *Stem Cells Transl Med.* 2023;12:379-390. <https://doi.org/10.1093/stcltm/szad029>
54. Schwarz CS, Bucher CH, Schlundt C, et al. Spatio-temporal bone remodeling after hematopoietic stem cell transplantation. *Int J Mol Sci.* 2020;22:267. <https://doi.org/10.3390/ijms22010267>
55. Rao A, Barkley D, França GS, Yanai I. Exploring tissue architecture using spatial transcriptomics. *Nature.* 2021;596:211-220. <https://doi.org/10.1038/s41586-021-03634-9>
56. Palla G, Fischer DS, Regev A, Theis FJ. Spatial components of molecular tissue biology. *Nat Biotechnol.* 2022;40:308-318. <https://doi.org/10.1038/s41587-021-01182-1>
57. Bressan D, Battistoni G, Hannon GJ. The dawn of spatial omics. *Science.* 2023;381:eabq4964. <https://doi.org/10.1126/science.abq4964>
58. Marx U, Akabane T, Andersson TB, et al. Biology-inspired micro-physiological systems to advance patient benefit and animal welfare in drug development. *Altex.* 2020;37:365-394. <https://doi.org/10.14573/altex.2001241>
59. Marx U, Accastelli E, David R, et al. An individual patient's "body" on chips-how organismoid theory can translate into your personal precision therapy approach. *Front Med (Lausanne).* 2021;8:728866. <https://doi.org/10.3389/fmed.2021.728866>
60. Cao UMN, Zhang Y, Chen J, Sayson D, Pillai S, Tran SD. Microfluidic organ-on-a-chip: a guide to biomaterial choice and fabrication. *Int J Mol Sci.* 2023;24:3232. <https://doi.org/10.3390/ijms24043232>
61. Huang S. Systems biology of stem cells: three useful perspectives to help overcome the paradigm of linear pathways. *Philos Trans R Soc Lond B Biol Sci.* 2011;366:2247-2259. <https://doi.org/10.1098/rstb.2011.0008>
62. McNamara LE, Turner L-A, Burgess KV. Systems biology approaches applied to regenerative medicine. *Curr Pathobiol Rep.* 2015;3:37-45. <https://doi.org/10.1007/s40139-015-0072-4>
63. Hamet P, Tremblay J. Artificial intelligence in medicine. *Metabolism.* 2017;69s:S36-s40. <https://doi.org/10.1016/j.metabol.2017.01.011>
64. Rajpurkar P, Chen E, Banerjee O, Topol EJ. AI in health and medicine. *Nat Med.* 2022;28:31-38. <https://doi.org/10.1038/s41591-021-01614-0>
65. Ouyang JF, Chothani S, Rackham OJL. Deep learning models will shape the future of stem cell research. *Stem Cell Reports.* 2023;18:6-12. <https://doi.org/10.1016/j.stemcr.2022.11.007>
66. O'Donnell MA, Jain B. How systems biology can help solve the enigma of aging. *Nature Aging.* 2021;1:750-752. <https://doi.org/10.1038/s43587-021-00115-6>
67. Bartal A, Jagodnik KM. Progress in and opportunities for applying information theory to computational biology and bioinformatics. *Entropy (Basel).* 2022;24:925. <https://doi.org/10.3390/e24070925>
68. Fouché A, Zinovyev A. Omics data integration in computational biology viewed through the prism of machine learning paradigms. *Front Bioinform.* 2023;3:1191961. <https://doi.org/10.3389/fbinf.2023.1191961>
69. Rai S, Raj U, Varadwaj PK. Systems biology: a powerful tool for drug development. *Curr Top Med Chem.* 2018;18:1745-1754. <https://doi.org/10.2174/1568026618666181025113226>
70. Loewa A, Feng JJ, Hedtrich S. Human disease models in drug development. *Nat Rev Bioeng.* 2023;1:1-559. <https://doi.org/10.1038/s44222-023-00063-3>
71. Bacakova L, Zarubova J, Travnickova M, et al. Stem cells: their source, potency and use in regenerative therapies with focus on adipose-derived stem cells - a review. *Biotechnol Adv.* 2018;36:1111-1126. <https://doi.org/10.1016/j.biotechadv.2018.03.011>
72. Blau HM, Brazelton TR, Weimann JM. The evolving concept of a stem cell: entity or function? *Cell.* 2001;105:829-841. [https://doi.org/10.1016/s0092-8674\(01\)00409-3](https://doi.org/10.1016/s0092-8674(01)00409-3)
73. Blau HM, Daley GQ. Stem cells in the treatment of disease. *N Engl J Med.* 2019;380:1748-1760. <https://doi.org/10.1056/NEJMra1716145>
74. Fuchs E, Blau HM. Tissue stem cells: architects of their niches. *Cell Stem Cell.* 2020;27:532-556. <https://doi.org/10.1016/j.stem.2020.09.011>
75. Du P, Wu J. Hallmarks of totipotent and pluripotent stem cell states. *Cell Stem Cell.* 2024;31:312-333. <https://doi.org/10.1016/j.stem.2024.01.009>
76. Parolini O, Alviano F, Bagnara GP, et al. Concise review: isolation and characterization of cells from human term placenta: outcome of the first international workshop on placenta derived stem cells. *Stem Cells.* 2008;26:300-311. <https://doi.org/10.1634/stemcells.2007-0594>
77. Silini AR, Di Pietro R, Lang-Olip I, et al. Perinatal derivatives: where do We stand? A roadmap of the human placenta and consensus for tissue and cell nomenclature. *Front Bioeng Biotechnol.* 2020;8:610544. <https://doi.org/10.3389/fbioe.2020.610544>
78. Moll G, Ignatowicz L, Catar R, et al. Different procoagulant activity of therapeutic mesenchymal stromal cells derived from bone marrow and placental decidua. *Stem Cells Dev.* 2015;24:2269-2279. <https://doi.org/10.1089/scd.2015.0120>
79. Davenport T, Kalakota R. The potential for artificial intelligence in healthcare. *Future Healthc J.* 2019;6:94-98. <https://doi.org/10.7861/futurehosp.6-2-94>
80. Mukherjee S, Yadav G, Kumar R. Recent trends in stem cell-based therapies and applications of artificial intelligence in regenerative medicine. *World J Stem Cells.* 2021;13:521-541. <https://doi.org/10.4252/wjsc.v13.i6.521>
81. Kufel J, Bargiel-Łączek K, Kocot S, et al. What is machine learning, artificial neural networks and deep learning?-examples of practical applications in medicine. *Diagnostics (Basel, Switzerland).* 2023;13:2582. <https://doi.org/10.3390/diagnostics13152582>
82. Nosrati H, Nosrati M. Artificial intelligence in regenerative medicine: applications and implications. *Biomimetics (Basel).* 2023;8:442. <https://doi.org/10.3390/biomimetics8050442>
83. Sharma NS. Patient centric approach for clinical trials: current trend and new opportunities. *Perspect Clin Res.* 2015;6:134-138. <https://doi.org/10.4103/2229-3485.159936>
84. Bohr A, Memarzadeh K. The rise of artificial intelligence in healthcare applications. *Artificial Intelligence Healthc.* 2020;25-60.
85. Heumos L, Schaar AC, Lance C, et al. Single-Cell Best Practices Consortium. Best practices for single-cell analysis across modalities. *Nat Rev Genet.* 2023;24:550-572. <https://doi.org/10.1038/s41576-023-00586-w>
86. Li X, Wang C-Y. From bulk, single-cell to spatial RNA sequencing. *Int J Oral Sci.* 2021;13:36. <https://doi.org/10.1038/s41368-021-00146-0>

87. Pinu FR, Beale DJ, Paten AM, et al. Systems biology and multi-omics integration: viewpoints from the metabolomics research community. *Metabolites*. 2019;9:76. <https://doi.org/10.3390/metabo9040076>
88. van Kampen AH, Moerland PD. Taking bioinformatics to systems medicine. *Methods Mol Biol*. 2016;1386:17-41. https://doi.org/10.1007/978-1-4939-3283-2_2
89. Gupta MK, Misra K. A holistic approach for integration of biological systems and usage in drug discovery. *Netw Model Anal Health Inform Bioinform*. 2016;5:4. <https://doi.org/10.1007/s13721-015-0111-4>
90. Shahzad K, Loor JJ. Application of top-down and bottom-up systems approaches in ruminant physiology and metabolism. *Curr Genomics*. 2012;13:379-394. <https://doi.org/10.2174/138920212801619269>
91. Duwal S, von Kleist M. Top-down and bottom-up modeling in system pharmacology to understand clinical efficacy: an example with NRTIs of HIV-1. *Eur J Pharm Sci*. 2016;94:72-83. <https://doi.org/10.1016/j.ejps.2016.01.016>
92. Tylutki Z, Polak S, Wiśniowska B. Top-down, bottom-up and middle-out strategies for drug cardiac safety assessment via modeling and simulations. *Curr Pharmacol Rep*. 2016;2:171-177. <https://doi.org/10.1007/s40495-016-0060-3>
93. Brereton RG. Statistical experimental design. *J Chemom*. 2017;31:e2902. <https://doi.org/10.1002/cem.2902>
94. Eicher T, Kinnebrew G, Patt A, et al. Metabolomics and multi-omics integration: a survey of computational methods and resources. *Metabolites*. 2020;10:202. <https://doi.org/10.3390/metabo10050202>
95. Jobjörnsson S, Schaak H, Musshoff O, Friede T. Improving the statistical power of economic experiments using adaptive designs. *Exp Econ*. 2023;26:357-382. <https://doi.org/10.1007/s10683-022-09773-8>
96. Gille C, Bölling C, Hoppe A, et al. HepatoNet1: a comprehensive metabolic reconstruction of the human hepatocyte for the analysis of liver physiology. *Mol Syst Biol*. 2010;6:411. <https://doi.org/10.1038/msb.2010.62>
97. Dressler F, Fischer S. Connecting in-body nano communication with body area networks: challenges and opportunities of the internet of nano things. *Nano Commun Netw*. 2015;6:29-38. <https://doi.org/10.1016/j.nancom.2015.01.006>
98. Kuestner A, Stratmann L, Wendt R, Fischer S, Dressler F. A simulation framework for connecting in-body nano communication with out-of-body devices. In: *Proceedings of the 7th ACM International Conference on Nanoscale Computing and Communication*. Virtual Event, USA: Association for Computing Machinery; 2020: Article12.
99. Figueroa RL, Zeng-Treitler Q, Kandula S, Ngo LH. Predicting sample size required for classification performance. *BMC Med Inform Decis Mak*. 2012;12:8. <https://doi.org/10.1186/1472-6947-12-8>
100. Jobson JD. Multiple linear regression. In: *Applied Multivariate Data Analysis: Regression and Experimental Design*. Springer New York; 1991:219-398.
101. C Sammut, GI Webb, eds. Logistic regression. In: *Encyclopedia of Machine Learning*. Springer US; 2010: 631-631.
102. Rey D, Neuhauser M. Wilcoxon-signed-rank test. In: M Lovric, ed. *International Encyclopedia of Statistical Science*. Springer Berlin Heidelberg; Berlin, Heidelberg; 2011:1658-1659.
- *103. Faul F, Erdfelder E, Lang AG, Buchner A. GPower 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods*. 2007;39:175-191. <https://doi.org/10.3758/bf03193146>
104. Freire PP, Marques AH, Baiocchi GC, et al. The relationship between cytokine and neutrophil gene network distinguishes SARS-CoV-2-infected patients by sex and age. *JCI Insight*. 2021;6:e147535. <https://doi.org/10.1172/jci.insight.147535>
105. Bender R, Lange S. Adjusting for multiple testing—when and how? *J Clin Epidemiol*. 2001;54:343-349. [https://doi.org/10.1016/S0895-4356\(00\)00314-0](https://doi.org/10.1016/S0895-4356(00)00314-0)
106. González-Plaza JJ, Furlan C, Rijavec T, et al. Advances in experimental and computational methodologies for the study of microbial-surface interactions at different omics levels. *Front Microbiol*. 2022;13:1006946. <https://doi.org/10.3389/fmicb.2022.1006946>
107. Krassowski M, Das V, Sahu SK, Misra BB. State of the field in multi-omics research: from computational needs to data mining and sharing. *Front Genet*. 2020;11:610798. <https://doi.org/10.3389/fgene.2020.610798>
108. Jendoubi T. Approaches to integrating metabolomics and multi-omics data: a primer. *Metabolites*. 2021;11:184. <https://doi.org/10.3390/metabo11030184>
109. Chen W, Zhao Y, Chen X, et al. A multicenter study benchmarking single-cell RNA sequencing technologies using reference samples. *Nat Biotechnol*. 2021;39:1103-1114. <https://doi.org/10.1038/s41587-020-00748-9>
110. Madeira F, Park YM, Lee J, et al. The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res*. 2019;47:W636-W641. <https://doi.org/10.1093/nar/gkz268>
111. Heather JM, Chain B. The sequence of sequencers: the history of sequencing DNA. *Genomics*. 2016;107:1-8. <https://doi.org/10.1016/j.ygeno.2015.11.003>
112. Gupta N. DNA extraction and polymerase chain reaction. *J Cytol*. 2019;36:116-117. https://doi.org/10.4103/joc.Joc_110_18
113. Verbeelen T, Van Houdt R, Leys N, Ganigué R, Mastroleo F. Optimization of RNA extraction for bacterial whole transcriptome studies of low-biomass samples. *iScience*. 2022;25:105311. <https://doi.org/10.1016/j.isci.2022.105311>
114. Andrzejewska A, Catar R, Schoon J, et al. Multi-Parameter analysis of biobanked human bone marrow stromal cells shows little influence for donor age and mild comorbidities on phenotypic and functional properties. *Front Immunol*. 2019;10:2474. <https://doi.org/10.3389/fimmu.2019.02474>
115. Houseley J, Tollervey D. The many pathways of RNA degradation. *Cell*. 2009;136:763-776. <https://doi.org/10.1016/j.cell.2009.01.019>
116. Moll G, Geißler S, Catar R, et al. Cryopreserved or fresh mesenchymal stromal cells: only a matter of taste or key to unleash the full clinical potential of MSC therapy? *Adv Exp Med Biol*. 2016;951:77-98. https://doi.org/10.1007/978-3-319-45457-3_7
117. Catar R, Moll G, Kamhieh-Milz J, et al. Expanded hemodialysis therapy ameliorates uremia-induced systemic microinflammation and endothelial dysfunction by modulating VEGF, TNF- α and AP-1 signaling. *Front Immunol*. 2021;12:774052. <https://doi.org/10.3389/fimmu.2021.774052>
118. Drzeniek NM, Mazzocchi A, Schlickeiser S, et al. Bio-instructive hydrogel expands the paracrine potency of mesenchymal stem cells. *Biofabrication*. 2021;13:10.1088/1758-5090/ac0a32. <https://doi.org/10.1088/1758-5090/ac0a32>
119. Drzeniek NM, Kahwaji N, Schlickeiser S, et al. Immuno-engineered mRNA combined with cell adhesive niche for synergistic modulation of the MSC secretome. *Biomaterials*. 2023;294:121971. <https://doi.org/10.1016/j.biomaterials.2022.121971>
120. Duong VA, Lee H. Bottom-up proteomics: advancements in sample preparation. *Int J Mol Sci*. 2023;24:5350. <https://doi.org/10.3390/ijms24065350>
121. Dagher G. Quality matters: international standards for biobanking. *Cell Prolif*. 2022;55:e13282. <https://doi.org/10.1111/cpr.13282>
122. Doorn J, Moll G, Le Blanc K, van Blitterswijk C, de Boer J. Therapeutic applications of mesenchymal stromal cells: paracrine effects and potential improvements. *Tissue Eng Part B Rev*. 2012;18:101-115. <https://doi.org/10.1089/ten.teb.2011.0488>
123. Singer NG, Caplan AI. Mesenchymal stem cells: mechanisms of inflammation. *Annu Rev Pathol*. 2011;6:457-478. <https://doi.org/10.1146/annurev-pathol-011110-130230>
124. ISO/DIS 5649(en)—Medical laboratories—concepts and specifications for the design, development, implementation, and use of laboratory-developed tests. <https://www.iso.org/obp/ui/es/#iso:std:81506:en>. Accessed April 30, 2024.

125. Streitz M, Miloud T, Kapinsky M, et al. Standardization of whole blood immune phenotype monitoring for clinical trials: panels and methods from the ONE study. *Transplant Res.* 2013;2:17. <https://doi.org/10.1186/2047-1440-2-17>
126. Schlickeiser S, Boes D, Streitz M, Sawitzki B. The use of novel diagnostics to individualize immunosuppression following transplantation. *Transpl Int.* 2015;28:911-920. <https://doi.org/10.1111/tri.12527>
127. Japp AS, Hoffmann K, Schlickeiser S, et al. Wild immunology assessed by multidimensional mass cytometry. *Cytometry A.* 2017;91:85-95. <https://doi.org/10.1002/cyto.a.22906>
128. Roemhild A, Otto NM, Moll G, et al. Regulatory T cells for minimising immune suppression in kidney transplantation: phase I/IIa clinical trial. *BMJ.* 2020;371:m3734. <https://doi.org/10.1136/bmj.m3734>
129. Sawitzki B, Harden PN, Reinke P, et al. Regulatory cell therapy in kidney transplantation (the ONE study): a harmonised design and analysis of seven non-randomised, single-arm, phase 1/2A trials. *Lancet.* 2020;395:1627-1639. [https://doi.org/10.1016/S0140-6736\(20\)30167-7](https://doi.org/10.1016/S0140-6736(20)30167-7)
130. Böttcher C, van der Poel M, Fernández-Zapata C, et al. Single-cell mass cytometry reveals complex myeloid cell composition in active lesions of progressive multiple sclerosis. *Acta Neuropathol Commun.* 2020;8:136. <https://doi.org/10.1186/s40478-020-01010-8>
131. Schulte-Schrepping J, Reusch N, Paclik D, et al. Severe COVID-19 is marked by a dysregulated myeloid cell compartment. *Cell.* 2020;182:1419-1440.e1423. <https://doi.org/10.1016/j.cell.2020.08.001>
132. Nikolaou C, Muehle K, Schlickeiser S, et al. High-dimensional single cell mass cytometry analysis of the murine hematopoietic system reveals signatures induced by ageing and physiological pathogen challenges. *Immun Ageing.* 2021;18:20. <https://doi.org/10.1186/s12979-021-00230-3>
133. Oppizzi L, Hosszu K, Prockop S, et al. Immune monitoring after cell therapy and hematopoietic cell transplantation: guidelines by the ISCT stem cell engineering committee. *Cytotherapy.* 2025;27:888-902. <https://doi.org/10.1016/j.jcyt.2025.04.069>
134. Gligorijević V, Pržulj N. Methods for biological data integration: perspectives and challenges. *J R Soc Interface.* 2015;12:20150571. <https://doi.org/10.1098/rsif.2015.0571>
135. Shi Y, Inoue H, Wu JC, Yamanaka S. Induced pluripotent stem cell technology: a decade of progress. *Nat Rev Drug Discov.* 2017;16:115-130. <https://doi.org/10.1038/nrd.2016.245>
136. Petzschner FH. Practical challenges for precision medicine. *Science.* 2024;383:149-150. <https://doi.org/10.1126/science.adm9218>
137. Chekroud AM, Hawrilenko M, Loh H, et al. Illusory generalizability of clinical prediction models. *Science.* 2024;383:164-167. <https://doi.org/10.1126/science.adg8538>
138. Bersanelli M, Mosca E, Remondini D, et al. Methods for the integration of multi-omics data: mathematical aspects. *BMC Bioinformatics.* 2016;17:15. <https://doi.org/10.1186/s12859-015-0857-9>
139. Starruß J, de Back W, Bruschi L, Deutsch A. Morpheus: a user-friendly modeling environment for multiscale and multicellular systems biology. *Bioinformatics.* 2014;30:1331-1332. <https://doi.org/10.1093/bioinformatics/btt772>
140. Zhou G, Soufan O, Ewald J, Hancock REW, Basu N, Xia J. NetworkAnalyst 3.0: a visual analytics platform for comprehensive gene expression profiling and meta-analysis. *Nucleic Acids Res.* 2019;47:W234-W241. <https://doi.org/10.1093/nar/gkz240>
141. Afgan E, Baker D, Batut B, et al. The galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res.* 2018;46:W537-W544. <https://doi.org/10.1093/nar/gky379>
142. Salgado RC, Fonseca DLM, Marques AHC, et al. The network interplay of interferon and toll-like receptor signaling pathways in the anti-candida immune response. *Sci Rep.* 2021;11:20281. <https://doi.org/10.1038/s41598-021-99838-0>
143. Schimke LE, Marques AHC, Baiocchi GC, et al. Severe COVID-19 shares a common neutrophil activation signature with other acute inflammatory states. *Cells.* 2022;11:847. <https://doi.org/10.3390/cells11050847>
144. ComBat. Adjust for batch effects using an empirical Bayes framework. <https://rdrr.io/bioc/sva/man/ComBat.html>. Accessed April 30, 2024.
145. RStudio. *Open Source & Professional Software for Data Science Teams*. RStudio. <https://www.rstudio.com/>. Accessed April 30, 2024.
146. Mullender S. X Python Reference Manual. <https://x-python.sourceforge.net/Doc/x/index.html>. Accessed April 30, 2024.
147. Bash Reference Manual. <https://www.gnu.org/software/bash/manual/bash.html>. Accessed April 30, 2024.
148. Bioconductor. <https://www.bioconductor.org/>. Accessed April 30, 2024.
149. Liu YYF, Lu Y, Oh S, Conduit GJ. Machine learning to predict mesenchymal stem cell efficacy for cartilage repair. *PLoS Comput Biol.* 2020;16:e1008275. <https://doi.org/10.1371/journal.pcbi.1008275>
150. Firoozbakht F, Yousefi B, Schwikowski B. An overview of machine learning methods for monotherapy drug response prediction. *Brief Bioinform.* 2022;23:bbab408. <https://doi.org/10.1093/bib/bbab408>
151. Ban C, Yang F, Wei M, et al. Integrative analysis of gene expression through One-Class logistic regression machine learning identifies stemness features in multiple myeloma. *Front Genet.* 2021;12:666561. <https://doi.org/10.3389/fgene.2021.666561>
152. Jung H, Jung H-U, Baek EJ, et al. Integration of risk factor polygenic risk score with disease polygenic risk score for disease prediction. *Commun Biol.* 2024;7:180. <https://doi.org/10.1038/s42003-024-05874-7>
153. Wu S, Wells A, Griffith LG, Lauffenburger DA. Controlling multipotent stromal cell migration by integrating “course-graining” materials and “fine-tuning” small molecules via decision tree signal-response modeling. *Biomaterials.* 2011;32:7524-7531. <https://doi.org/10.1016/j.biomaterials.2011.06.050>
154. Moslehi S, Rabiei N, Soltanian AR, Mamani M. Application of machine learning models based on decision trees in classifying the factors affecting mortality of COVID-19 patients in Hamadan, Iran. *BMC Med Inform Decis Mak.* 2022;22:192. <https://doi.org/10.1186/s12911-022-01939-x>
155. Imai Y, Kanie K, Kato R. Morphological heterogeneity description enabled early and parallel non-invasive prediction of T-cell proliferation inhibitory potency and growth rate for facilitating donor selection of human mesenchymal stem cells. *Inflamm Regen.* 2022;42:8. <https://doi.org/10.1186/s41232-021-00192-5>
156. Huang BF, Boutros PC. The parameter sensitivity of random forests. *BMC Bioinformatics.* 2016;17:331. <https://doi.org/10.1186/s12859-016-1228-x>
157. Lien C-Y, Chen T-T, Tsai E-T, et al. Recognizing the differentiation degree of human induced pluripotent stem cell-derived retinal pigment epithelium cells using machine learning and deep learning-based approaches. *Cells.* 2023;12:211. <https://doi.org/10.3390/cells12020211>
158. Cozzetto D, Minneci F, Currant H, Jones DT. FFPred 3: feature-based function prediction for all gene ontology domains. *Sci Rep.* 2016;6:31865. <https://doi.org/10.1038/srep31865>
159. Mota SM, Rogers RE, Haskell AW, et al. Automated mesenchymal stem cell segmentation and machine learning-based phenotype classification using morphometric and textural analysis. *J Med Imaging (Bellingham).* 2021;8:014503. <https://doi.org/10.1117/1.Jmi.8.1.014503>
160. Lee JY, Styczynski MP. NS-kNN: A Modified K-Nearest Neighbors Approach for Imputing Metabolomics Data.
161. Wood A, Shpilrain V, Najarian K, Kahrobaei D. Private naive bayes classification of personal biomedical data: application in cancer data analysis. *Comput Biol Med.* 2019;105:144-150. <https://doi.org/10.1016/j.combiomed.2018.11.018>

162. Kong Y, Ao J, Chen Q, et al. Evaluating differentiation status of mesenchymal stem cells by Label-Free microscopy system and machine learning. *Cells*. 2023;12:1524.
163. Buggenthin F, Buettner F, Hoppe PS, et al. Prospective identification of hematopoietic lineage choice by deep learning. *Nat Methods*. 2017;14:403-406. <https://doi.org/10.1038/nmeth.4182>
164. Marzec-Schmidt K, Ghosheh N, Stahlschmidt SR, Küppers-Munther B, Synnergren J, Ulfenborg B. Artificial intelligence supports automated characterization of differentiated human pluripotent stem cells. *Stem Cells*. 2023;41:850-861. <https://doi.org/10.1093/stmcls/sxad049>
165. Atashgahi Z, Pieterse J, Liu S, Mocanu DC, Veldhuis R, Pechenizkiy M. A brain-inspired algorithm for training highly sparse neural networks. *Mach Learn*. 2022;111:4411-4452. <https://doi.org/10.1007/s10994-022-06266-w>
166. Khouj Y, Dawson J, Coad J, Vona-Davis L. Hyperspectral imaging and K-means classification for histologic evaluation of ductal carcinoma in situ. *Front Oncol*. 2018;8:17. <https://doi.org/10.3389/fonc.2018.00017>
167. He Z, Shen X, Zhou Y, Wang Y. Application of K-means clustering based on artificial intelligence in gene statistics of biological information engineering. In: *Proceedings of the 2024 4th International Conference on Bioinformatics and Intelligent Computing*. Beijing, China: Association for Computing Machinery; 2024:468-473.
168. Zimmermann R, Nitschke M, Magno V, et al. Discriminant principal component analysis of ToF-SIMS spectra for deciphering compositional differences of MSC-secreted extracellular matrices. *Small Methods*. 2023;7:e2201157. <https://doi.org/10.1002/smt.202201157>
169. Taherkhani A, Cosma G, McGinnity TM. AdaBoost-CNN: an adaptive boosting algorithm for convolutional neural networks to classify multi-class imbalanced datasets using transfer learning. *Neurocomputing*. 2020;404:351-366.
170. Rapakoulia T, Lopez Ruiz De Vargas S, Omgha PA, Laupert V, Ulitsky I, Vingron M. Centre: a gradient boosting algorithm for cell-type-specific ENhancer-Target pREdiction. *Bioinformatics*. 2023;39:btad687. <https://doi.org/10.1093/bioinformatics/btad687>
171. Li K, Yao S, Zhang Z, et al. Efficient gradient boosting for prognostic biomarker discovery. *Bioinformatics*. 2022;38:1631-1638. <https://doi.org/10.1093/bioinformatics/btab869>
172. Koza JR, Bennett FH, Andre D, Keane MA. Automated design of both the topology and sizing of analog electrical circuits using genetic programming. In: JS Gero, F Sudweeks, eds. *Artificial Intelligence in Design '96*. Springer Netherlands; 1996:151-170.
173. Lever J, Krzywinski M, Altman N. Principal component analysis. *Nat Methods*. 2017;14:641-642. <https://doi.org/10.1038/nmeth.4346>
174. Breiman L. Random forests. *Machine Learn*. 2001;45:5-32. <https://doi.org/10.1023/A:1010933404324>
175. Huang S, Cai N, Pacheco PP, Narrandes S, Wang Y, Xu W. Applications of support vector machine (SVM) learning in cancer genomics. *Cancer Genomics Proteomics*. 2018;15:41-51. <https://doi.org/10.21873/cgp.20063>
176. Karatzoglou A, Smola A, Hornik K, Zeileis A. kernlab - an S4 package for kernel methods in R. *J Stat Soft*. 2004;11:1-20. <https://doi.org/10.18637/jss.v011.i09>
177. Singh K, Xie M. Bootstrap method. In: P Peterson, E Baker, B McGaw, eds. *International Encyclopedia of Education*. 3rd ed.Elsevier; 2010:46-51.
178. Durap A. A comparative analysis of machine learning algorithms for predicting wave runup. *Anthropocene Coasts*. 2023;6:17. <https://doi.org/10.1007/s44218-023-00033-7>
179. Medrano-Trochez C, Chatterjee P, Pradhan P, et al. Single-cell RNA-seq of out-of-thaw mesenchymal stromal cells shows tissue-of-origin differences and inter-donor cell-cycle variations. *Stem Cell Res Ther*. 2021;12:565. <https://doi.org/10.1186/s13287-021-02627-9>
180. Chinnadurai R, Viswanathan S, Moll G. Editorial: next generation MSC therapy manufacturing, potency and mechanism of action analysis. *Front Immunol*. 2023;14:1192636. <https://doi.org/10.3389/fimmu.2023.1192636>
181. Giri J, Moll G. MSCs in space: mesenchymal stromal cell therapeutics as enabling technology for long-distance manned space travel. *Curr Stem Cell Rep*. 2022;8:1-13. <https://doi.org/10.1007/s40778-022-00207-y>
182. Hoogduijn MJ, de Witte SFH, Luk F, et al. Effects of freeze-thawing and intravenous infusion on mesenchymal stromal cell gene expression. *Stem Cells Dev*. 2016;25:586-597. <https://doi.org/10.1089/scd.2015.0329>
183. Moll G, Le Blanc K. Engineering more efficient multipotent mesenchymal stromal (stem) cells for systemic delivery as cellular therapy. *ISBT Sci Ser*. 2015;10:357-365. <https://doi.org/10.1111/vox.12133>
184. Moll G, Hult A, von Bahr L, et al. Do ABO blood group antigens hamper the therapeutic efficacy of mesenchymal stromal cells? *PLoS One*. 2014;9:e85040. <https://doi.org/10.1371/journal.pone.0085040>
185. Moll G, Alm JJ, Davies LC, et al. Do cryopreserved mesenchymal stromal cells display impaired immunomodulatory and therapeutic properties? *Stem Cells*. 2014;32:2430-2442. <https://doi.org/10.1002/stem.1729>
186. Chinnadurai R, Garcia MA, Sakurai Y, et al. Actin cytoskeletal disruption following cryopreservation alters the biodistribution of human mesenchymal stromal cells in vivo. *Stem Cell Reports*. 2014;3:60-72. <https://doi.org/10.1016/j.stemcr.2014.05.003>
187. Chinnadurai R, Copland IB, Garcia MA, et al. Cryopreserved mesenchymal stromal cells are susceptible to T-Cell mediated apoptosis which is partly rescued by IFN γ licensing. *Stem Cells*. 2016;34:2429-2442. <https://doi.org/10.1002/stem.2415>
188. Gramlich OW, Burand AJ, Brown AJ, Deutsch RJ, Kuehn MH, Ankrum JA. Cryopreserved mesenchymal stromal cells maintain potency in a retinal ischemia/reperfusion injury model: toward an off-the-shelf therapy. *Sci Rep*. 2016;6:26463. <https://doi.org/10.1038/srep26463>
189. Galipeau J. Letter to the editor: response to "function of cryopreserved MSCs with and without IFN γ pre-licensing is context dependent" by Ankrum et al. *Stem Cells*. 2017;35:1440-1441. <https://doi.org/10.1002/stem.2526>
190. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS*. 2012;16:284-287. <https://doi.org/10.1089/omi.2011.0118>
191. Kuleshov MV, Jones MR, Rouillard AD, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res*. 2016;44:W90-W97. <https://doi.org/10.1093/nar/gkw377>
192. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA*. 2005;102:15545-15550. <https://doi.org/10.1073/pnas.0506580102>
193. Shevade K, Peddada S, Mader K, Przybyla L. Functional genomics in stem cell models: considerations and applications. *Front Cell Dev Biol*. 2023;11:1236553. <https://doi.org/10.3389/fcell.2023.1236553>
194. Barata T, Duarte I, Futschik ME. Integration of stemness gene signatures reveals core functional modules of stem cells and potential novel stemness genes. *Genes (Basel)*. 2023;14:745. <https://doi.org/10.3390/genes14030745>
195. Tiklová K, Nolbrant S, Fiorenzano A, et al. Single cell transcriptomics identifies stem cell-derived graft composition in a model of Parkinson's disease. *Nat Commun*. 2020;11:2434. <https://doi.org/10.1038/s41467-020-16225-5>
196. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol*. 2018;36:411-420. <https://doi.org/10.1038/nbt.4096>

197. Tran HTN, Ang KS, Chevrier M, et al. A benchmark of batch-effect correction methods for single-cell RNA sequencing data. *Genome Biol.* 2020;21:12. <https://doi.org/10.1186/s13059-019-1850-9>
198. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;15:550. <https://doi.org/10.1186/s13059-014-0550-8>
199. Bielekova B, Vodovotz Y, An G, Hallenbeck J. How implementation of systems biology into clinical trials accelerates understanding of diseases. *Front Neurol.* 2014;5:102. <https://doi.org/10.3389/fneur.2014.00102>
200. AlMusawi S, Ahmed M, Nateri AS. Understanding cell-cell communication and signaling in the colorectal cancer microenvironment. *Clin Transl Med.* 2021;11:e308. <https://doi.org/10.1002/ctm2.308>
201. Panés J, García-Olmo D, Van Assche G, et al. Expanded allogeneic adipose-derived mesenchymal stem cells (Cx601) for complex perianal fistulas in Crohn's disease: a phase 3 randomised, double-blind controlled trial. *Lancet.* 2016;388:1281-1290. [https://doi.org/10.1016/S0140-6736\(16\)31203-X](https://doi.org/10.1016/S0140-6736(16)31203-X)
202. Francois M, Galipeau J. New insights on translational development of mesenchymal stromal cells for suppressor therapy. *J Cell Physiol.* 2012;227:3535-3538. <https://doi.org/10.1002/jcp.24081>
203. Galipeau J. The mesenchymal stromal cells dilemma-does a negative phase III trial of random donor mesenchymal stromal cells in steroid-resistant graft-versus-host disease represent a death knell or a bump in the road? *Cytotherapy.* 2013;15:2-8. <https://doi.org/10.1016/j.jcyt.2012.10.002>
204. Galipeau J. Mesenchymal stromal cells for graft-versus-host disease: a trilogy. *Biol Blood Marrow Transpl.* 2020;26:e89-e91. <https://doi.org/10.1016/j.bbmt.2020.02.023>
205. Martin I, Galipeau J, Kessler C, Le Blanc K, Dazzi F. Challenges for mesenchymal stromal cell therapies. *Sci Transl Med.* 2019;11:eaat2189. <https://doi.org/10.1126/scitranslmed.aat2189>
206. Winkler T, Perka C, von Roth P, et al. Immunomodulatory placental-expanded, mesenchymal stromal cells improve muscle function following hip arthroplasty. *J Cachexia Sarcopenia Muscle.* 2018;9:880-897. <https://doi.org/10.1002/jcsm.12316>
207. Winkler T, Costa ML, Ofir R, et al. HIPGEN: a randomized, multicentre phase III study using intramuscular PLacenta-eXpanded stromal cells therapy for recovery following hip fracture arthroplasty: a study design. *Bone Jt Open.* 2022;3:340-347. <https://doi.org/10.1302/2633-1462.34.BJO-2021-0156.R1>
208. Norgren L, Weiss N, Nikol S, et al. PACE: randomized, controlled, multicentre, multinational, phase III study of PLX-PAD for critical limb ischaemia in patients unsuitable for revascularization: randomized clinical trial. *Br J Surg.* 2024;111:znad437. <https://doi.org/10.1093/bjs/znad437>
209. Avivar-Valderas A, Martín-Martín C, Ramírez C, et al. Dissecting allo-sensitization after local administration of human allogeneic adipose mesenchymal stem cells in perianal fistulas of Crohn's disease patients. *Front Immunol.* 2019;10:1244. <https://doi.org/10.3389/fimmu.2019.01244>
210. Gotts JE, Matthay MA. Cell-based therapy in sepsis. A step closer. *Am J Respir Crit Care Med.* 2018;197:280-281. <https://doi.org/10.1164/rccm.201710-2068ED>
211. Laffey JG, Matthay MA. Fifty years of research in ARDS. Cell-based therapy for acute respiratory distress syndrome. Biology and potential therapeutic value. *Am J Respir Crit Care Med.* 2017;196:266-273. <https://doi.org/10.1164/rccm.201701-0107CP>
212. Sagoo P, Perucha E, Sawitzki B, et al. Development of a cross-platform biomarker signature to detect renal transplant tolerance in humans. *J Clin Invest.* 2010;120:1848-1861. <https://doi.org/10.1172/JCI39922>
213. Abou-El-Enein M, Bauer G, Reinke P. The business case for cell and gene therapies. *Nat Biotechnol.* 2014;32:1192-1193. <https://doi.org/10.1038/nbt.3084>
214. Abou-El-Enein M, Elsanhoury A, Reinke P. Overcoming challenges facing advanced therapies in the EU market. *Cell Stem Cell.* 2016;19:293-297. <https://doi.org/10.1016/j.stem.2016.08.012>
215. Abou-El-Enein M, Volk HD, Reinke P. Clinical development of cell therapies: setting the stage for academic success. *Clin Pharmacol Ther.* 2017;101:35-38. <https://doi.org/10.1002/cpt.523>
216. Fraunhofer-MEOS-Center-for-Microelectronic-and-Optical-Systems-for-Biomedicine. Multimodal data fusion and analysis. <https://www.meos.fraunhofer.de/en/research-development/multimodal-data-fusion-analysis.html>. Accessed May 20, 2024.
217. Qi W, Su H, Fan K, et al. Multimodal data fusion framework enhanced robot-assisted minimally invasive surgery. *Trans Instit Meas Control.* 2022;44:735-743. <https://doi.org/10.1177/0142331220984350>
218. Graphite-Note. *How Much Data Is Needed For Machine Learning?* <https://graphite-note.com/how-much-data-is-needed-for-machine-learning/>. Accessed April 10, 2024.
219. Smolic H. *How Much Data Do You Need for Machine Learning?* <https://graphite-note.com/how-much-data-is-needed-for-machine-learning/>. Accessed 10 April, 2024.
220. Subbiah V. The next generation of evidence-based medicine. *Nat Med.* 2023;29:49-58. <https://doi.org/10.1038/s41591-022-02160-z>
221. Srinivasan M, Thangaraj SR, Ramasubramanian K, Thangaraj PP, Ramasubramanian KV. Exploring the current trends of artificial intelligence in stem cell therapy: a systematic review. *Cureus.* 2021;13:e20083. <https://doi.org/10.7759/cureus.20083>
222. Zaman WSWK, Karman SB, Ramlan EI, Tukimin SNB, Ahmad MYB. Machine learning in stem cells research: application for biosafety and bioefficacy assessment. *IEEE Access.* 2021;9:25926-25945. <https://doi.org/10.1109/ACCESS.2021.3056553>
223. Hanahan D, Weinberg R. Hallmarks of cancer: the next generation. *Cell.* 2011;144:646-674. <https://doi.org/10.1016/j.cell.2011.02.013>
224. Hanahan D. Hallmarks of cancer: new dimensions. *Cancer Discov.* 2022;12:31-46. <https://doi.org/10.1158/2159-8290.Cd-21-1059>
225. Marar RI, Abbasi MA, Prathivadhi-Bhayankaram S, et al. Cardio-toxicities of novel therapies in hematologic malignancies: chimeric antigen receptor T-cell therapy and bispecific T-cell engager therapy. *JCO Oncol Pract.* 2023;19:331-342. <https://doi.org/10.1200/op.22.00713>
226. Jing Y, Liu J, Ye Y, et al. Multi-omics prediction of immune-related adverse events during checkpoint immunotherapy. *Nat Commun.* 2020;11:4946. <https://doi.org/10.1038/s41467-020-18742-9>
227. Liang EC, Rejeski K, Fei T, et al. Development and validation of an automated computational approach to grade immune effector cell-associated hematotoxicity. *Bone Marrow Transplant.* 2024;59:910-917. <https://doi.org/10.1038/s41409-024-02278-3>
228. Fischer L, Grieb N, Born P, et al. Cellular dynamics following CAR T cell therapy are associated with response and toxicity in relapsed/refractory myeloma. *Leukemia.* 2024;38:372-382. <https://doi.org/10.1038/s41375-023-02129-y>
229. Tian YE, Di Biase MA, Mosley PE, et al. Evaluation of brain-body health in individuals with common neuropsychiatric disorders. *JAMA Psychiatry.* 2023;80:567-576. <https://doi.org/10.1001/jamapsychiatry.2023.0791>
230. Duarte RRR, Pain O, Bendall ML, et al. Integrating human endogenous retroviruses into transcriptome-wide association studies highlights novel risk factors for major psychiatric conditions. *Nat Commun.* 2024;15:3803. <https://doi.org/10.1038/s41467-024-48153-z>
231. Kaplan DL, Moon RT, Vunjak-Novakovic G. It takes a village to grow a tissue. *Nat Biotechnol.* 2005;23:1237-1239. <https://doi.org/10.1038/nbt1005-1237>
232. Kang H-W, Lee SJ, Ko IK, Kengla C, Yoo JJ, Atala A. A 3D bio-printing system to produce human-scale tissue constructs with structural integrity. *Nat Biotechnol.* 2016;34:312-319. <https://doi.org/10.1038/nbt.3413>

233. Murphy SV, Atala A. 3D bioprinting of tissues and organs. *Nat Biotechnol.* 2014;32:773-785. <https://doi.org/10.1038/nbt.2958>
234. Yamanaka S. Pluripotent stem cell-based cell therapy-promise and challenges. *Cell Stem Cell.* 2020;27:523-531. <https://doi.org/10.1016/j.stem.2020.09.014>
235. Coronello C, Francipane MG. Moving towards induced pluripotent stem cell-based therapies with artificial intelligence and machine learning. *Stem Cell Rev Rep.* 2022;18:559-569. <https://doi.org/10.1007/s12015-021-10302-y>
236. Lubiana T, Lopes R, Medeiros P, et al. Ten quick tips for harnessing the power of ChatGPT in computational biology. *PLoS Comput Biol.* 2023;19:e1011319. <https://doi.org/10.1371/journal.pcbi.1011319>
237. Brody H. Regenerative medicine. *Nature.* 2016;540:S49. <https://doi.org/10.1038/540S49a>
238. Grand-View-Research. Regenerative Medicine Market Size & Trends. <https://www.grandviewresearch.com/industry-analysis/regenerative-medicine-market#>. Accessed June 3, 2024.
239. Butler D. Translational research: crossing the valley of death. *Nature.* 2008;453:840-842. <https://doi.org/10.1038/453840a>
240. Mullard A. Parsing clinical success rates. *Nat Rev Drug Discov.* 2016;15:447. <https://doi.org/10.1038/nrd.2016.136>
241. Ramsey BW, Nepom GT, Lonial S. Academic, foundation, and industry collaboration in finding new therapies. *N Engl J Med.* 2017;376:1762-1769. <https://doi.org/10.1056/NEJMr1612575>
242. Bates SM, Evans KV, Delsing L, Wong R, Cornish G, Bahjat M. Immune safety challenges facing the preclinical assessment and clinical progression of cell therapies. *Drug Discov Today.* 2024;29:104239. <https://doi.org/10.1016/j.drudis.2024.104239>
243. Viswanathan S, Galipeau J. Hallmarks of MSCs: key quality attributes for pharmacology and clinical use. *Cell Stem Cell.* 2025;32:878-894. <https://doi.org/10.1016/j.stem.2025.05.008>
244. Galipeau J, Krampera M. The challenge of defining mesenchymal stromal cell potency assays and their potential use as release criteria. *Cytotherapy.* 2015;17:125-127. <https://doi.org/10.1016/j.jcyt.2014.12.008>
245. Samsonraj RM, Rai B, Sathiyathan P, et al. Establishing criteria for human mesenchymal stem cell potency. *Stem Cells.* 2015;33:1878-1891. <https://doi.org/10.1002/stem.1982>
246. Galipeau J, Krampera M, Barrett J, et al. International society for cellular therapy perspective on immune functional assays for mesenchymal stromal cells as potency release criterion for advanced phase clinical trials. *Cytotherapy.* 2016;18:151-159. <https://doi.org/10.1016/j.jcyt.2015.11.008>
247. Nolte JA, Galipeau J, Phinney DG. Improving mesenchymal stem/stromal cell potency and survival: proceedings from the international society of cell therapy (ISCT) MSC preconference held in May 2018, palais des congres de montreal, organized by the ISCT MSC scientific committee. *Cytotherapy.* 2020;22:123-126. <https://doi.org/10.1016/j.jcyt.2020.01.004>
248. Cheung TS, Bertolino GM, Giacomini C, Bornhäuser M, Dazzi F, Galleu A. Mesenchymal stromal cells for graft versus host disease: mechanism-based biomarkers. *Front Immunol.* 2020;11:1338. <https://doi.org/10.3389/Fimmu.2020.01338>
249. Chinnadurai R, Rajan D, Qayed M, et al. Potency analysis of mesenchymal stromal cells using a combinatorial assay matrix approach. *Cell Rep.* 2018;22:2504-2517. <https://doi.org/10.1016/j.celrep.2018.02.013>
250. Chinnadurai R, Rajakumar A, Schneider AJ, Bushman WA, Hematti P, Galipeau J. Potency analysis of mesenchymal stromal cells using a Phospho-STAT matrix loop analytical approach. *Stem Cells.* 2019;37:1119-1125. <https://doi.org/10.1002/stem.3035>
251. Giri J, Galipeau J. Mesenchymal stromal cell therapeutic potency is dependent upon viability, route of delivery, and immune match. *Blood Adv.* 2020;4:1987-1997. <https://doi.org/10.1182/bloodadvances.2020001711>
252. Hansen SB, Højgaard LD, Kastrup J, Ekblond A, Follin B, Juhl M. Optimizing an immunomodulatory potency assay for mesenchymal stromal cell. *Front Immunol.* 2022;13:1085312. <https://doi.org/10.3389/fimmu.2022.1085312>
253. Porter AP, Pirlot BM, Dyer K, et al. Conglomeration of T and B cell matrix responses determines the potency of human bone marrow mesenchymal stromal cells. *Stem Cells.* 2022;40:1134-1148. <https://doi.org/10.1093/stmcls/sxac064>
254. Galleu A, Riffo-Vasquez Y, Trento C, et al. Apoptosis in mesenchymal stromal cells induces in vivo recipient-mediated immunomodulation. *Sci Transl Med.* 2017;9:eam7828. <https://doi.org/10.1126/scitranslmed.aam7828>
255. Krampera M, Le Blanc K. Mesenchymal stromal cells: putative microenvironmental modulators become cell therapy. *Cell Stem Cell.* 2021;28:1708-1725. <https://doi.org/10.1016/j.stem.2021.09.006>
256. Galipeau J, Krampera M, Leblanc K, et al. Mesenchymal stromal cell variables influencing clinical potency: the impact of viability, fitness, route of administration and host predisposition. *Cytotherapy.* 2021;23:368-372. <https://doi.org/10.1016/j.jcyt.2020.11.007>
257. Weiss DJ, Filiano A, Galipeau J, et al. An international society for cell and gene therapy mesenchymal stromal cells committee editorial on overcoming limitations in clinical trials of mesenchymal stromal cell therapy for coronavirus disease-19: time for a global registry. *Cytotherapy.* 2022;24:1071-1073. <https://doi.org/10.1016/j.jcyt.2022.07.010>
258. Polley MC. On the quest of risk stratification in HER2-positive breast cancer. *J Natl Cancer Inst.* 2022;114:345-346. <https://doi.org/10.1093/jnci/djab061>
259. Kapil A, Spitzmüller A, Brieu N, et al. HER2 quantitative continuous scoring for accurate patient selection in HER2 negative trastuzumab deruxtecan treated breast cancer. *Sci Rep.* 2024;14:12129. <https://doi.org/10.1038/s41598-024-61957-9>
260. Zhang H, Yang F, Xu Y, et al. Multimodal integration using a machine learning approach facilitates risk stratification in HR+/HER2- breast cancer. *Cell Rep Med.* 2025;6:101924. <https://doi.org/10.1016/j.xcrm.2024.101924>
261. Loupy A, Mengel M, Haas M. Thirty years of the International Banff Classification for Allograft Pathology: the past, present, and future of kidney transplant diagnostics. *Kidney Int.* 2022;101:678-691. <https://doi.org/10.1016/j.kint.2021.11.028>
262. Yoo D, Goutaudier V, Divard G, et al. An automated histological classification system for precision diagnostics of kidney allografts. *Nat Med.* 2023;29:1211-1220. <https://doi.org/10.1038/s41591-023-02323-6>
263. Yoo D, Divard G, Raynaud M, et al. A machine learning-driven virtual biopsy system for kidney transplant patients. *Nat Commun.* 2024;15:554. <https://doi.org/10.1038/s41467-023-44595-z>
264. Hölscher DL, Bouteldja N, Joodaki M, et al. Next-generation morphometry for pathomics-data mining in histopathology. *Nat Commun.* 2023;14:470. <https://doi.org/10.1038/s41467-023-36173-0>
265. Bülow RD, Hölscher DL, Costa IG, Boor P. Extending the landscape of omics technologies by pathomics. *NPJ Syst Biol Appl.* 2023;9:38. <https://doi.org/10.1038/s41540-023-00301-9>
266. Bülow RD, Lan Y-C, Amann K, Boor P. Künstliche intelligenz in der nierentransplantationspathologie. *Pathologie.* 2024;45:277-283. <https://doi.org/10.1007/s00292-024-01324-7>
267. Lu MY, Chen TY, Williamson DFK, et al. AI-based pathology predicts origins for cancers of unknown primary. *Nature.* 2021;594:106-110. <https://doi.org/10.1038/s41586-021-03512-4>
268. Kather JN, Pearson AT, Halama N, et al. Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer. *Nat Med.* 2019;25:1054-1056. <https://doi.org/10.1038/s41591-019-0462-y>

269. Redlich J-P, Feuerhake F, Weis J, et al. Applications of artificial intelligence in the analysis of histopathology images of gliomas: a review. *Npj Imaging*. 2024;2:16. <https://doi.org/10.1038/s44303-024-00020-8>
270. Dembrower K, Crippa A, Colón E, Eklund M, Strand F, ScreenTrust-CAD Trial Consortium. Artificial intelligence for breast cancer detection in screening mammography in Sweden: a prospective, population-based, paired-reader, non-inferiority study. *Lancet Digit Health*. 2023;5:e703-e711. [https://doi.org/10.1016/S2589-7500\(23\)00153-X](https://doi.org/10.1016/S2589-7500(23)00153-X)
271. Ng AY, Oberije CJG, Ambrózay É, et al. Prospective implementation of AI-assisted screen reading to improve early detection of breast cancer. *Nat Med*. 2023;29:3044-3049. <https://doi.org/10.1038/s41591-023-02625-9>
272. Tatullo M. Entropy meets physiology: should we translate aging as disorder? *Stem Cells*. 2024;42:91-97. <https://doi.org/10.1093/stmcls/sxad084>
273. Ball LM, Bernardo ME, Roelofs H, et al. Multiple infusions of mesenchymal stromal cells induce sustained remission in children with steroid-refractory, grade III-IV acute graft-versus-host disease. *Br J Haematol*. 2013;163:501-509. <https://doi.org/10.1111/bjh.12545>
274. Dickersin K, Min YI. NIH clinical trials and publication bias. *Online J Curr Clin Trials*. 1993;Doc No 50:[4967 words; 53 paragraphs].
275. Decullier E, Lhéritier V, Chapuis F. Fate of biomedical research protocols and publication bias in France: retrospective cohort study. *BMJ*. 2005;331:19. <https://doi.org/10.1136/bmj.38488.385995.8F>
276. Song F, Parekh-Bhurke S, Hooper L, et al. Extent of publication bias in different categories of research cohorts: a meta-analysis of empirical studies. *BMC Med Res Methodol*. 2009;9:79. <https://doi.org/10.1186/1471-2288-9-79>