

*Perspective*

## Transforming Drug Discovery for Aging Research with AI-Based Integrated Approaches

**Quentin Vanhaelen**

Insilico Medicine, Rm 8, Unit 111-113, ICT Co-Working Centre, 1/F Enterprise Place, 5 Science Park West Avenue, Hong Kong;

Email: vanhaelen@insilicomedicine.com

---

### ABSTRACT

Aging is a complex biological process which is a consequence of an unbalance between cellular dynamical equilibriums, environmental constraints and, from a larger perspective, evolutionary constraints. The key characteristic of aging is that it propagates through the body to progressively affect all organ systems. Because of this systemic nature, the study of aging-related diseases and the design of appropriate drugs and treatments have been shown to be challenging. Thanks to recent technological developments, it is now possible to make a more efficient and systematic use of the large amount of biological and patient data which has been generated and accumulated over the years. All these different data types provide insights into the mechanisms of aging from different and complementary perspectives and designing methods to extract as much information as possible from them is of paramount importance. Advanced multimodal deep learning methods are transforming aging research and drug discovery. AI based platforms can be deployed to identify promising targets for aging-related diseases and AI can be used for the generation of novel drug-like compounds with good activity profiles and suitable chemical properties. Those new computational methods have the potential to revolutionize drug discovery in oncology, a field currently facing an increasing pressure to deliver new drugs in a faster and more efficient ways.

**KEYWORDS:** deep learning; drug discovery; artificial intelligence; target identification; generative adversarial network; reinforcement learning; de novo molecular generation; aging

---

### Open Access

Received: 16 December 2019

Accepted: 05 March 2020

Published: 09 March 2020

Copyright © 2020 by the author(s). Licensee Hapres, London, United Kingdom. This is an open access article distributed under the terms and conditions of [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

### ABBREVIATIONS

ARD, age-related-disease; ARG, age-related-gene; NME, New Molecular Entity; AI, artificial intelligence; GAN, generative adversarial network; RL, reinforcement learning; VAE, variational autoencoders

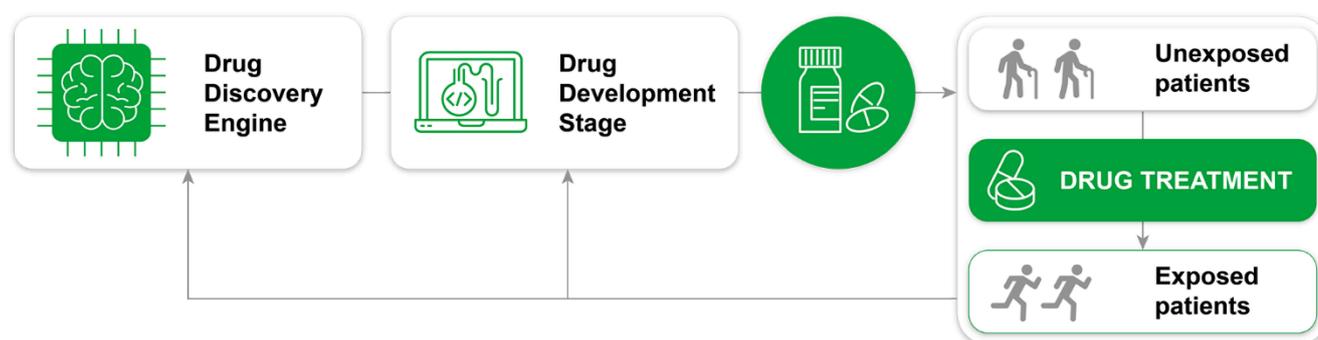
There are ongoing debates about the definition of aging. Nonetheless, over the last decades, large amounts of experimental data was generated and further used to expand our biological knowledge and understanding

of aging thanks to computational methods making possible to analyze the behavior and the dynamics of biological systems from various perspectives. Qualifying aging as a complex phenomenon would be an understatement. The systems where aging takes place, i.e., living organisms, are among the most complex systems ever been studied by scientists. The human body is a multi-levels complex system whose components (cells, tissues, organs and systems) are organized in a hierarchical manner while interacting between each other to perform multiple tasks and functionalities. Due to this systemic nature, dysfunctions affecting all kinds of processes within cells of one or several organs propagate to all parts of the cells, organs and systems leading to a progressive disruption of the general homeostasis. From this perspective, one considers aging as a systemic process which translates into a time-dependent decline of functional capacity and stress resistance, associated with increased risk of morbidity and mortality. The external symptoms of aging are the apparent disparate damages leading to the onset of ARDs. ARDs are diseases which appear or have a higher probability to appear as individuals get older. Studies showed that ARDs are related to genes which undergo change in expression with age, the so-called ARG, i.e., genes which are beneficial at young age can, for whatever reason, become harmful at older age. The time dependence of aging is also apparent when examining the local topologies of protein-protein interaction networks. Indeed they elicit significant modifications with age. More precisely, local topologies around only a subset of proteins in the networks do change with age and interestingly those which undergo major modifications in their interactions are related to the ARGs and ARDs. There is an overlap between functions and diseases that are enriched in their aging-related predictions and those that are enriched in well-known aging-related data. Aging could be the consequence of local topological changes in the vicinity of ARGs which ultimately affect the stability of the different systems and organs within the body. Mechanistically, aging is a disruption of the homeostasis established between cellular processes and age-related-diseases are the external symptoms of this disruption [1]. The systemic nature of this process should not be neglected when looking for treatment.

The pharmacology industry faces important challenges to deliver new treatments for ARDs (Parkinson's disease, Alzheimer's disease, cardiovascular diseases, etc.) which are on the rise due to an increasingly aging population. Indeed, the drug discovery industry is facing financial pressures driven by the increasing cost of bringing a drug to market [2]. This is due to a combination of factors including increased safety requirements by regulatory authorities, demands for larger clinical trials and greater overall development costs. It costs up to \$2.6 Billion and takes up to 10 years to bring the new NME to the market [3,4]. Despite the many advances and availability of new technologies continuously integrated within drug discovery and development pipeline, the average

number of NME released on the market has been on the decline since the 1950s.

The key to optimize drug discovery and allow healthcare to face the tremendous challenges coming ahead is to accelerate, through the design of effective multimodal approaches, the convergence of big data, chemistry, biology and medicine together with the capabilities offered by the recent development in AI [5]. This would allow a better characterization of the interplay between molecular structures, properties, alteration in biological samples and drug response which is needed for the development of more effective treatments (Figure 1). For target identification and early drug discovery, this goal can be achieved with the deployment of automated drug discovery AI engines for optimizing target identification, lead selection and optimization. Indeed, improving the early stages of the drug discovery process is paramount to reduce the high attrition rate faced during clinical trials [6–8].



**Figure 1.** Integrated AI-based technologies can make drug discovery pipeline more efficient. At the early phase of drug discovery, AI-based technologies is used to optimize the identification of drug-target pairs. This reduces the attrition rates encountered during further phases of development and clinical trials. AI Technologies can be used to develop medical devices which monitor the health status of individuals. Patient and health data obtained from patient experiences might be used to further optimize the design of personalized treatments through a feedback loop between experimental validation and computational methods.

AI-based algorithms are used to identify novel targets in a variety of diseases using massive -omics data and text repositories. Compounds for those novel targets are obtained through generative chemistry. Generative chemistry started to emerge in 2015 as a new area of application of deep learning where generative models are used to propose novel molecules with the desired activity, specificity and novelty as well as many other generation conditions [9]. The generated molecules can be further validated and annotated using in silico virtual screening tools if needed. Among the numerous architectures and algorithms developed so far, GANs and RL, have encountered great successes [10,11]. Others modes such as VAEs, and autoregressive models are also rapidly growing in popularity for molecular de novo design and molecular feature extraction [12].

At first, many computational chemists and medicinal chemists in the pharmaceutical industry were skeptical about this novel approach. The molecules that we generated using the generative models were not diverse enough or easy to synthesize and the targets were too easy with the available structures and hits and it was not obvious to assess the true benefit of using AI-base technologies over other standard methods. Nevertheless, recent results obtained using generative chemistry where potent and novel DDR1 inhibitor were generated, synthesized and went through in vitro and in vivo testing in around 45 days showed that when correctly integrated within drug discovery pipelines, this technology offer promising results [13]. This study and other subsequent results demonstrate the capabilities of the generative reinforcement learning generator technology in terms of novel molecule design. The fact that these molecules were synthesized and tested in multiple experiments including animals is a valuable experimental validation of this technology.

What distinguishes healthcare from other fields when it comes to deploying and validating deep learning technologies is that while the advances in high performance computing and data management made it possible to train the systems very quickly, the time needed to test the output is much more longer than in other industries dealing with pictures, videos, and text for instance. The validation of molecules in cells and animals is time consuming, takes effort and money and requires the continuous collaboration between experts in AI and biological, chemical, and medical sciences. Nevertheless, as the AI field for drug discovery matures, one can expects that the industry continues to move towards the integration of such technologies into the drug discovery pipelines.

#### **CONFLICTS OF INTEREST**

The author is affiliated with Insilico Medicine, a company developing an AI-based end-to-end integrated pipeline for target identification and drug discovery and engaged in aging and cancer research.

#### **REFERENCES**

1. Vanhaelen Q. Aging as an optimization between cellular maintenance requirements and evolutionary constraints. *Curr Aging Sci.* 2015;8(1):110-9.
2. Scannell JW, Blanckley A, Boldon H, Warrington B. Diagnosing the decline in pharmaceutical R&D efficiency. *Nat Rev Drug Discov.* 2012;12:191-200.
3. Vamathevan J, Clark D, Czodrowski P, Dunham I, Ferran E, Lee G, et al. Applications of machine learning in drug discovery and development. *Nat Rev Drug Discov.* 2019;18(6):463-77.
4. DiMasi JA, Grabowski HG, Hansen RW. Innovation in the pharmaceutical industry: New estimates of R&D costs. *J Health Econ.* 2016;47:20-33.

5. Zhavoronkov A, Mamoshina P, Vanhaelen Q, Scheibye-Knudsen M, Moskalev A, Aliper A. Artificial intelligence for aging and longevity research: Recent advances and perspectives. *Ageing Res Rev.* 2019;49:49-66.
6. Harrer S, Shah P, Antony B, Hu J. Artificial Intelligence for Clinical Trial Design. *Trends Pharmacol Sci.* 2019;40(8):577-91.
7. Gayvert KM, Madhukar NS, Elemento O. A Data-Driven Approach to Predicting Successes and Failures of Clinical Trials. *Cell Chem Biol.* 2016;23(10):1294-301.
8. Viceconti M, Henney A, Morley-Fletcher E. In silico clinical trials: how computer simulation will transform the biomedical industry. *Int J Clin Trials.* 2016;3:37.
9. Kadurin A, Aliper A, Kazennov A, Mamoshina P, Vanhaelen Q, Khrabrov K, et al. The cornucopia of meaningful leads: Applying deep adversarial autoencoders for new molecule development in oncology. *Oncotarget.* 2017;8(7):10883-90.
10. Putin E, Asadulaev A, Ivanenkov Y, Aladinskiy V, Sanchez-Lengeling B, Aspuru-Guzik A, et al. Reinforced Adversarial Neural Computer for de Novo Molecular Design. *J Chem Inf Model.* 2018;58:1194-204.
11. Putin E, Asadulaev A, Vanhaelen Q, Ivanenkov Y, Aladinskaya AV, Aliper A, et al. Adversarial Threshold Neural Computer for Molecular de Novo Design. *Mol Pharm.* 2018;15:4386-97.
12. Kadurin A, Nikolenko S, Khrabrov K, Aliper A, Zhavoronkov A. druGAN: An Advanced Generative Adversarial Autoencoder Model for de Novo Generation of New Molecules with Desired Molecular Properties in Silico. *Mol Pharm.* 2017;15:3098-104.
13. Zhavoronkov A, Ivanenkov YA, Aliper A, Veselov MS, Aladinskiy VA, Aladinskaya AV, et al. Deep learning enables rapid identification of potent DDR1 kinase inhibitors. *Nat Biotechnol.* 2019;37(9):1038-40.

How to cite this article:

Vanhaelen Q. Transforming drug discovery for aging research with AI-based integrated approaches. *Adv Geriatr Med Res.* 2020;2(2):e200010. <https://doi.org/10.20900/agmr20200010>