Current Hypertension Reports (2020) 22: 70
https://doi.org/10.1007/s11906-020-01068-8

TELEMEDICINE AND TECHNOLOGY (HB BOSWORTH, SECTION EDITOR)

AI (Artificial Intelligence) and Hypertension Research

Franco B. Mueller[1]

Published online: 27 August 2020
© Springer Science+Business Media, LLC, part of Springer Nature 2020

Abstract

Purpose of Review This review a highlights that to use artificial intelligence (AI) tools effectively for hypertension research, a new foundation to further understand the biology of hypertension needs to occur by leveraging genome and RNA sequencing technology and derived tools on a broad scale in hypertension.

Recent Findings For the last few years, progress in research and management of essential hypertension has been stagnating while at the same time, the sequencing of the human genome has been generating many new research tools and opportunities to investigate the biology of hypertension. Cancer research has applied modern tools derived from DNA and RNA sequencing on a large scale, enabling the improved understanding of cancer biology and leading to many clinical applications. Compared with cancer, studies in hypertension, using whole genome, exome, or RNA sequencing tools, total less than 2% of the number cancer studies. While true, sequencing the genome of cancer tissue has provided cancer research an advantage, DNA and RNA sequencing derived tools can also be used in hypertension to generate new understanding how complex protein network, in non-cancer tissue, adapts and learns to be effective when for example, somatic mutations or environmental inputs change the gene expression profiles at different network nodes. The amount of data and differences in clinical condition classification at the individual sample level might be of such magnitude to overwhelm and stretch comprehension. Here is the opportunity to use AI tools for the analysis of data streams derived from DNA and RNA sequencing tools combined with clinical data to generate new hypotheses leading to the discovery of mechanisms and potential target molecules from which drugs or treatments can be developed and tested.

Summary Basic and clinical research taking advantage of new gene sequencing-based tools, to uncover mechanisms how complex protein networks regulate blood pressure in health and disease, will be critical to lift hypertension research and management from its stagnation. The use of AI analytic tools will help leverage such insights. However, applying AI tools to vast amounts of data that certainly exist in hypertension, without taking advantage of new gene sequencing-based research tools, will generate questionable results and will miss many new potential molecular targets and possibly treatments. Without such approaches, the vision of precision medicine for hypertension will be hard to accomplish and most likely not occur in the near future.

Keywords Artificial intelligence · Deep machine learning algorithms · Whole genome and RNA sequencing · Hypertension treatment · Gene and protein networks · Target molecules · Cancer and hypertension research publications

Introduction

Recent calls to transform how we approach one of the pressing global health problems, hypertension, summarize in detail current shortcomings in detection, control, appropriate treatment, and management of the condition [1•]. Progress in hypertension has stalled, and it appears that over the last decade ,very few new insights have been generated to better understand the biology and the pathophysiology of the condition. At the same time, new research tools to better understand the biology of conditions derived from DNA and RNA sequencing are now widely available. The premise of this article is that broad application of these new tools of molecular biology and genetics, which include RNA and whole genome sequencing with associated protein-omics and metabolomics in hypertension research, has the potential to help lift hypertension research and management from its stagnation, and ultimately lead to a more precise therapeutic management and care at the individual (single) patient level (N-of-1). The

This article is part of the Topical Collection on *Telemedicine and Technology*

✉ Franco B. Mueller
francobm6@gmail.com

[1] Division of Nephrology and Hypertension, Department of Medicine, New York Presbyterian Hospital-Weill Cornell Medical College, 318 West 100th Street, Box 8D, New York, NY 10025, USA

individual patient (N-of-1) is what physicians treat in daily practice, and all too often patients poorly fit treatment guidelines, with elderly patients, women, ethnic minorities and or patients with severe co-morbidities underrepresented or excluded from randomized controlled trials (RCTs) to establish such guidelines.

It is tempting to assume, with today's availability of huge financial medical claims data, clinical records, public health data sources, and databases enabling the generation of vast data inputs, that analysis of such data streams by artificial intelligence algorithms will find better treatment solutions for individual patients and thereby lead to better population blood pressure control and outcome. Ultimately, artificial intelligence (AI) algorithms that incorporate learning-processes prompted by inheritance and interactions between biological entities and environment will help make precision medicine possible. However, before such algorithms can be applied successfully in hypertension on a large scale, inputs need to add data and knowledge gained by the broad use DNA and RNA sequencing and related technologies.

Current Antihypertensive Drug Treatment was Built on a Large Foundation of Basic and Clinical Research

Starting in the 1940′s, giants of hypertension research and treatment, using established sciences of physics, chemistry, biochemistry, and biology, generated the data and knowledge that ultimately led to seven or more pharmaceutical drug classes that successfully lower blood pressure long-term at relatively low risk and cost, thereby providing physicians with an effective armamentarium to reduce the sequelae of hypertension: heart, kidney disease, and stroke. Initial pathophysiological insights enabled separation of primary forms of hypertension from secondary ones, such as hypertension due to hyperaldosteronism, the obstruction of the renal artery with excessive renin and angiotensin II production, or tumors of the adrenergic nervous system. Understanding of these secondary forms, together with mechanisms discovered in primary hypertension, such as renal salt and water metabolism, the renin-angiotensin aldosterone system, the functioning of pre- and postsynaptic receptors of the adrenergic nervous system, and mechanisms of calcium entry into vascular smooth muscle cells, led to antihypertensive drug development based on molecular targets identified by such research. Early on, using the understanding of these mechanisms in primary hypertension, attempts were made to personalize treatment at the individual patient level [2, 3•]. However, such treatment approaches had little chance to succeed given the powerful financial interests of pharmaceutical manufacturers who advocated every newly approved drug to reduce blood pressure in every hypertensive patient.

Figure 1 shows the human clinical trials publication history of antihypertensive drugs from seven different pharmacotherapeutic classes [4]. In many of these classes, multiple different active ingredients with different safety, side effect, and efficacy profiles haven been tested and approved for treatment. The black line in Fig. 1 represents the sum of all trials smoothed by the moving average which is peaking in the late 1980s and early 90s, thereafter, flattening and declining

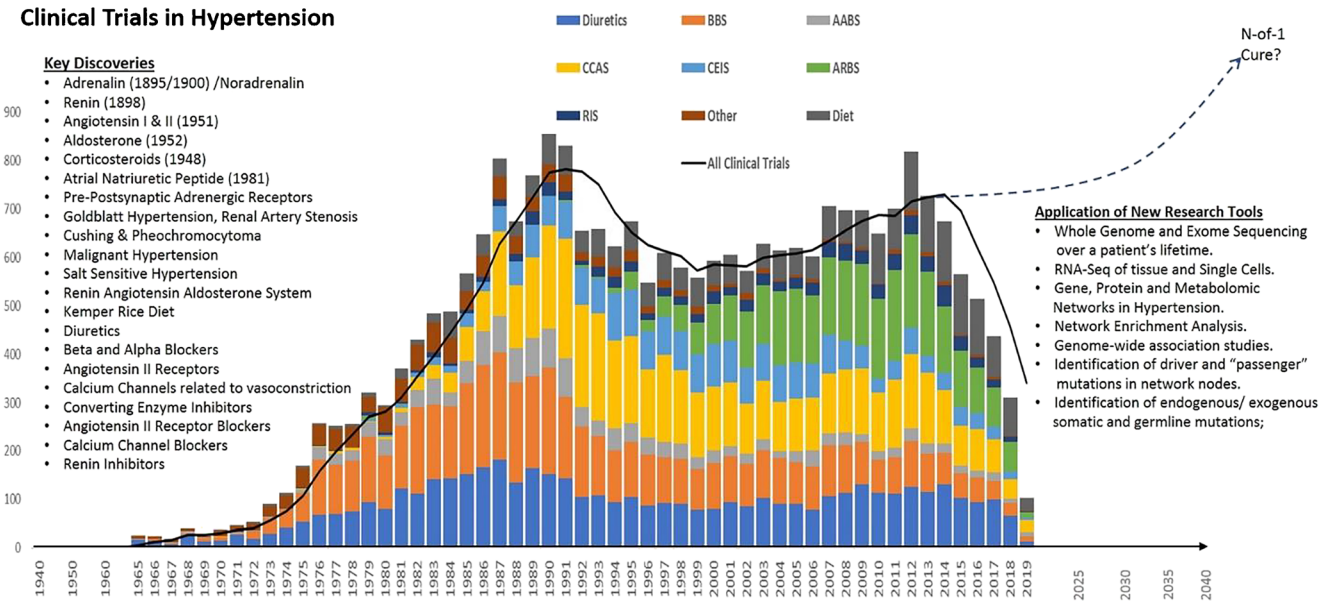

Fig. 1 History of human clinical trials in hypertension; Diuretic—blue, BBS (Betablockers)—orange; AABS (alpha blockers)—gray; CCA (calcium channel blockers)—yellow; CEIS (converting enzyme inhibitors)—light blue; ARBS (angiotensin receptors blockers)—green; RIS (renin inhibitors) —dark blue; Other (clonidine, hydralazine, and methyldopa)—brown; Diet—gray; Smoothed line of the sum of all clinical trial by moving average: black. Source: PubMed accessed January 15, 2020

over the last 5–7 years even when including dietary trials. Except for atrial natriuretic peptide, discovered in 1981 [5], all pharmacotherapeutic principles and molecular targets were discovered and tested starting in the late 40s throughout the 60s, 70s, and early 80s. This basic research was fundamental to generating the seven pharmacotherapeutic drug classes and high public interest in hypertension. As many physicians who specialize in hypertension treatment will attest, the antihypertensive medications of these drug classes enable effective blood pressure control in almost all patients (maybe as high as 95%) with very few patients resistant to treatment, as documented by the difficulties of finding such patients for the enrollment in renal denervation studies. Of course, as Dzau and Balatbat remind us, in practice, large gaps exist in detection, blood pressure control, and treatment compliance [1•].

Two Innovative Discoveries that Shape Biology and Computing

Two fundamental innovative discoveries, one shaping our understanding of living biology, the other our scientific approach, need to be incorporated into research of blood pressure biology. Large scale gene sequencing and algorithmic computation, the basis of all AI programs, will greatly impact the development of precision treatment, thereby leading to better blood pressure control and treatment compliance.

In 1953, Watson and Crick discovered the double-stranded helical structure of the DNA, consisting of a sequence of nucleobases of four different chemicals containing the blueprint for every living cell and inherited by its offspring [6••]. Although they did not, as they believed, "unlock all secrets of life", they showed that the information is stored in DNA as a sequence of base symbols (A, G, T, C) in a fixed alphabet, thus providing evidence that biology is computational [7]. DNA sequences define the level of protein expression in a cell, and these in turn cause other proteins to be expressed according to the interdependencies specified in the protein expression network [7].

In 2003, the first human genome was sequenced costing roughly $2.7 billion over a 10-year period. Although the hope that knowing the sequence would vastly accelerate the development of new treatment approaches for many different diseases, this did not occur in the expected timeframe. However, the project-generated technologies that now enable sequencing of a whole genome in 26 h for $1000 or less. We are now able to sequence whole exomes and genomes in virtually any tissue and most recently, in single cells of tissues, possibly multiple times over a patient's lifetime. The engineering progress in sequencing is outpacing computer hardware development. Plotting the reduction of sequencing cost, as a measure of sequencing technology progress over time, and comparing it to the curve of Moore's Law, which plots a long-term trend in the computer hardware industry that shows the doubling of "compute power" every 2 years [8], highlights a large gap between the two curves (Fig. 2a). One is tempted to speculate and ask if current computer hardware technology would be able to handle the vast amount of data and analysis were whole gene sequencing to be applied today in the clinical setting on a large scale.

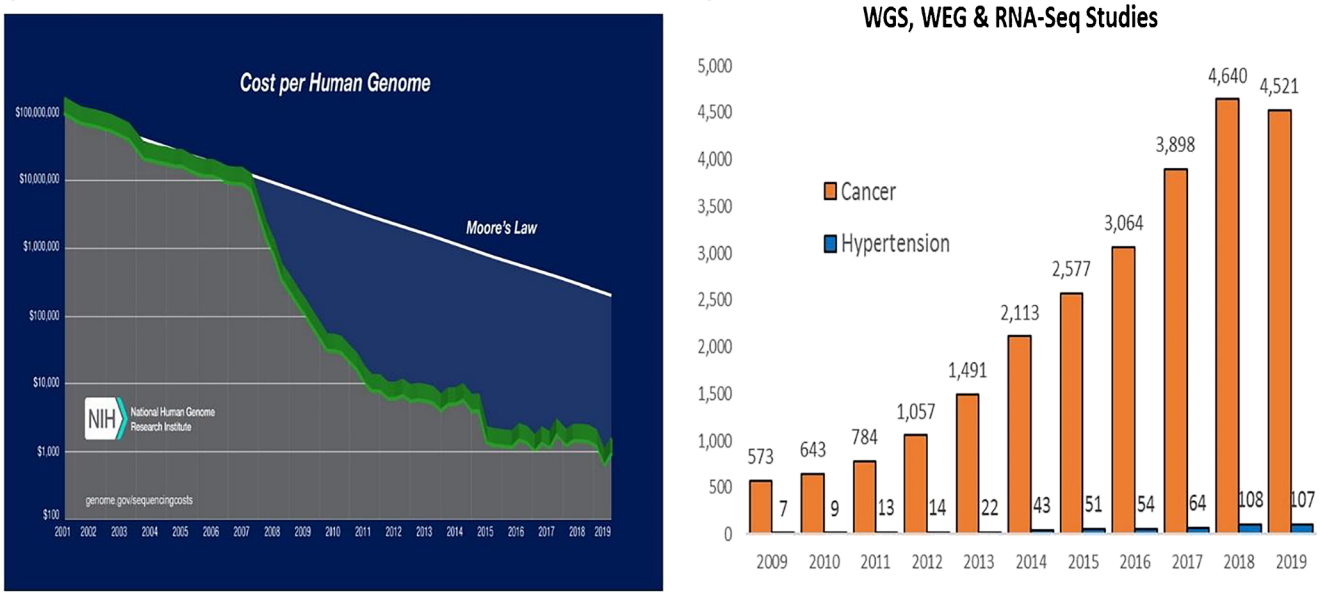

Fig. 2 **a** Reduction in DNA sequencing costs compared to Moore's Law, which describes a long-term trend in the computer hardware industry that involves the doubling of "compute power" every2 years. The first human genome cost roughly $2.7 billion in 2003. In 2016, a whole human genome can be sequenced for $1000 or as low as $699. Source: Wetterstrand KA. DNA Sequencing Costs: Data from the NHGRI Genome Sequencing Program (GSP) Available at: www.genome.gov/sequencingcostsdata accessed January 15, 2020. **b** Whole genome (WGS), exome, and RNA sequencing studies in cancer (orange) and hypertension (blue). Source: PubMed accessed January 15, 2020

The capability to sequence the whole genome and RNA has vastly improved our understanding of living biology in terms of complex computational interaction networks that act in cells and between cells based on DNA. It encodes information for more than 20,000 proteins with instructions controlled by regulatory mechanisms, also coded by DNA and RNA, that specify how much new protein of each kind is produced or not. As an example, the KEGG pathway database, one of multiple such databases which try to capture our current knowledge of molecular gene expression interactions, reactions, and relation networks, lists more than 500 network pathways with 29,000 genes [9]. However, the unique number of genes in the database is only 7394, indicating that for more than 12,000 protein coding genes, our knowledge of how these genes interact within the network of human biology is still lacking. A new understanding is emerging that these biological networks are shaped by continuous molecular learning processes. Interactions between the biological entity and its environment adapt the network to optimize outcomes for the biological entity. Learning in this context can be understood as a computational process.

In 1936, Alan Turing discovered and described that the execution of step-by-step procedures for processing information can be defined, captured, and studied systematically. He showed how to design a universal Turing machine that can execute every possible mechanical procedure. He also proved that not all well-defined mathematical problems can be solved mechanically [10••]. These step-by-step procedures, now commonly known as algorithms, follow computational laws that are as striking as physical laws [11]. Turing understood that these computational laws do not only apply to programming digital computers (machines) but to biology, and that natural phenomena can be understood as computational processes or algorithms. He described how such algorithms apply to biology in "The Chemical Basis of Morphogenesis" [12••] and predicted in "Intelligent Machinery, A Heretical Theory", that if a machine is to be intelligent, then it will need to learn by experience, and he states "…why one should not start from a comparatively simple machine, and, by subjecting it to a suitable range of 'experience' transform it into one which was more elaborate and was able to deal with a far greater range of contingencies" [13]. Turing thereby describes the foundation of artificial intelligence as a learning process. Today, the foundation of most, if not all programs of artificial intelligence, is based on machine learning algorithms.

Cancer Research Leading the Way Using Gene Sequencing Tools

Cancer research has applied modern tools derived from DNA and RNA sequencing on a large scale, enabling the improved understanding of cancer biology and leading to clinical applications such as improved diagnosis and prognosis, identification of new therapeutic targets, decision support for therapeutic choices, and application in disease monitoring [14•]. Sequencing cancerous tissue obtained from tumors alongside samples of normal tissue, usually blood, allows genetic variants to be identified and classified in either somatic mutations only found in tumor samples or inherited (germline) polymorphisms. Analyses of data from tens of thousands of patients in such studies are generating wide-ranging insights in cancer biology. Somatic mutations arise both from endogenous and exogenous mutational processes. Exogenous mutations can arise from chemicals such as tobacco, aflatoxin B, chemotherapeutic agents, ionizing radiation, and ultraviolet light, all of which damage DNA, generating mutations when damaged bases are incorrectly repaired or copied. Cell intrinsic processes (endogenous mutations), such as errors that occur during DNA replication and or impaired DNA repair due to activity of viruses, or increased amounts of reactive oxygen species that may be generated by long-term exposure to certain dietary behaviors, occur at a constant rate throughout life and accumulate with increasing age [14•]. Growing evidence in cancer research further suggests that, among the thousands of mutations acquired by a cancer cell, only a handful instruct the cell to act as an autonomous clone. These are called driver mutations, and the remaining are termed "passenger" mutations [15]. These driver mutations are of high interest because they are causative, and drugs that target the function of resulting proteins can be therapeutic. Cancer research is developing tools to identify driver mutations either by identifying mutation frequency, which requires to estimate the background mutation frequency, or by function-based approaches.

Figure 2b shows the number of cancer publications responsible for driving the improved understanding of cancer biology and treatment using whole genome, exome, or RNA sequencing. In comparison, studies in hypertension, excluding pulmonary hypertension, using whole genome, exome, or RNA sequencing, total less than 2% of the number cancer studies [16]. One can argue that cancer research has the advantage of being able to obtain and sequence cancerous tissue alongside samples of normal tissue.

What Can Hypertension Research Learn from Cancer Research?

True, sequencing the genome of cancer tissue has provided cancer research an advantage and is helping to understand cancer biology, but tools derived from DNA and RNA sequencing can also be used in hypertension research. We can identify the gene and RNA transcriptome not only in blood and urine specimen, kidney, and possibly vascular tissue, but also at the single cell level of these tissues. Combing RNA with whole genome sequencing will enable research into

control mechanisms of extended and complex protein networks of blood pressure control that still must be discovered. Let us not forget that the discovery of how the renin angiotensin aldosterone system interacts with sodium excretion and blood pressure control is one of the earliest descriptions of a protein network acting within different cells and across tissues and organs in medicine. Understanding how complex protein network, in non-cancer tissue adapt and learn to be effective when somatic mutations change the gene expression profiles at different network nodes, is one of the challenges to be uncovered. Initial studies of somatic mutations in hypertension follow the blueprint of cancer research and have identified such mutations in secondary hypertension of primary aldosteronism [17], paraganglion and pheochromocytoma [18], and Cushing's disease [19] but also in voltage-gated calcium channels [20], vascular disease formation [21], and pulmonary hypertension [22].

In cancer research, many algorithms have been developed for detecting somatic mutations. However, compared to proliferating cancer cells, in which the mutations confer a selective advantage, the frequency of DNA molecules carrying functionally relevant somatic mutations in normal cells and therefore possibly in hypertension is much lower [23•], but new approaches are emerging to characterize somatic mutations in normal and non-tumor tissue in bulk and single cell sequencing [23•]. While whole genome sequencing can detect most types of somatic mutations, including structural variants, in non-cancer tissue, the detection is limited to high frequency mutations, as the sequencing depth required to detect such low frequency mutation today remains prohibitively expensive. A solution overcoming this problem might be integrating bulk and single cell sequencing approaches to detect somatic mutations in non-cancer tissue.

The New Frontier of Hypertension Research

Generating new insights into the complex gene and protein interactions and regulatory networks that adapt to mutational changes over a patient's lifetime is critical to gain a deeper understanding of hypertension biology and to find the next generation of treatment tailored to the individual patient. How will we be able to understand the complex network of gene and protein interaction involved in hypertension at the individual patient level? Are there somatic mutations that change the hypertension gene and protein network at critical network nodes, and if so, will we be able to identify them? Are there somatic mutations of genes that are capable of interfering with the blood pressure set point in individual patients? These, and many more such questions, are the new challenges of hypertension research. Data from patient tissue samples, organ fluids, bloodwork, and single cells can be gathered to study genomics, proteomics, metabolomics, and lipidomics and to generate new hypotheses that ultimately will lead to new treatment targets, but there are additional challenges. The amount of data and differences in clinical condition classification at the individual sample level might be of such magnitude to overwhelm and stretch comprehension. Furthermore, researchers' efforts often tend to be based on previous efforts and hypotheses which can bias outcomes and overly restrict the research direction and assessment of molecular markers and targets that are correlated with the condition but ultimately do not prove to be causative.

Given such challenges, there is an opportunity to apply artificial intelligence or, a more appropriate term, deep machine learning tools, to the analysis of data streams generated by DNA and RNA sequencing tools and clinical data to generate new hypotheses that will lead, if confirmed, to actional insights, such as discovery of potential mechanisms and target molecules for which drugs or treatment can be developed and tested. These tools do not work by having expert-developed analytical techniques programmed into them; rather, users feed them sample problems (e.g., a network of gene expression) and solutions (how the gene expressions ultimately relate to different clinical conditions) so that the software can develop its own computational approaches for producing the same solutions [24, 25]. Numerous tools, such as deep autoencoders, deep belief, and deep neural networks, are already being researched and tested for cancer diagnosis, gene selection/classification, variants identification, and numerous other applications [26•]. Deep learning algorithms are run with unsupervised and supervised training methodologies. In supervised learning, labeled data (e.g., hypertension yes/no) are used to train, and weights are used to minimize the error to predict a target value for classification or regression, whereas unsupervised learning is usually used for clustering, feature extraction, or dimensionality reduction [24,26•]. The principle aim of these tools will be not to uncover the cause of hypertension but to generate new hypotheses and to develop diagnostic classifications not only in broad terms but ultimately at the patient level as well. Results can then be tested, confirmed, or discarded by additional mechanistic studies and, when applied, treatment effectiveness.

In conclusion, the use of deep machine learning tools will be critical to generate new hypotheses and clinical classifications about the biology of hypertension. They will help leverage potential actional insights generated by DNA and RNA sequencing, related techniques, and the vast clinical knowledge already available. However, to obtain such new insights in hypertension biology, broad use of new DNA and RNA sequencing and related technologies needs to generate an additional foundation of data and knowledge. This will be critical for the generation of new hypotheses and molecular understanding of the condition and will help lift hypertension research, detection, management, and treatment out of its stagnation. Applying AI tools to vast amounts of data, that

certainly exist in hypertension, without taking advantage of these new gene sequencing-based research tools, will generate questionable results and will miss many new potential molecular targets. Without such approaches, the vision of precision medicine for hypertension will be hard to accomplish and most likely not occur in the near future. Furthermore, many might be tempted to argue, given that many of the modern hypertensive drugs are available as generics at relatively cheap rates, that the quest for treatments to more precisely target the pathophysiology at the individual patient level might significantly increase the financial burden of hypertension. While possibly true short term, such financial considerations must be weighed against possible better outcomes such as reaching blood pressure targets or improved compliance especially in patient groups underrepresented in past trials. The current Covid-19 pan epidemic underscores the importance of linking basic gene sequencing and molecular and clinical research. Angiotensin converting enzyme 2 (ACE2), discovered with the help of an expressed sequence tag database in 2000, is an additional pivotal component of the renin angiotensin system [27, 28]. SARS-CoV-2 uses ACE2 in the lung as a receptor to enter cells and questions emerged over the benefits or harms of RAS blockade during SARS-Cov-2 exposure [29]. As of May 2020, there are at least seven dedicated studies to investigating the effect of anti-hypertensive regimen on the course of COVID-19 in patients with hypertension.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

Human and Animal Rights and Informed Consent This article does not contain any studies with human or animal subjects performed by the author.

References

Papers of particular interest, published recently, have been highlighted as:
- • Of importance
- •• Of major importance

1• Dzau VJ, Balatbat CA. Future of hypertension. Hypertension. 2019;74(3):450–7. https://doi.org/10.1161/HYPERTENSIONAHA.119.13437 **This paper highlights the slow progress in hypertension control and despite hypertension being studied recently in a large number of clinical trials, and recent research has not yielded mayor advances in hypertension.**

2. Bühler FR, Hulthén UL, Bolli P. Calcium channel inhibitors for identification of mechanisms and treatment of hypertension. J Cardiovasc Pharmacol. 1982;4(Suppl 3):S267–8.

3• Mueller FB, Laragh JH. Issues. goals and guidelines for choosing first-line and combination antihypertensive drug therapy. In: Laragh JH, Brenner BM, editors. Hypertension: Pathophysiology, Diagnosis and Management, vol. 2. New York: Raven Press; 1990. **Early approach of individualized treatment of hypertension**.

4. PubMed Search of human clinical trial publications for each drug class. Accessed 1/15/2020.

5. De Bold AJ. Atrial natriuretic factor: a hormone produced by the heart. Science. 1985;230(4727):767–70. https://doi.org/10.1126/science.2932797.

6•• Watson J, Crick F. Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. Nature. 1953;171:737–8. https://doi.org/10.1038/171737a0 **Key paper by Watson and Crick describing the DNA.**

7. Valiant L. Probably approximately correct: nature's algorithms for learning and prospering in a complex world. Basic Books. 2013 p 52.

8. Wetterstrand KA. DNA sequencing costs: data from the NHGRI Genome Sequencing Program (GSP) Available at: www.genome.gov/sequencingcostsdata. Accessed 1/15/2020.

9. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Mao T. KEGG as a reference resource for gene and protein annotation. Nucleic Acids Res. 2016;44(D1):D457–62. https://doi.org/10.1093/nar/gkv1070.

10•• Turing A. On computable numbers, with an application to the Entscheidungsproblem (1936). In: Copeland BJ. The Essential Turing. Oxford: Oxford University Press; 2013: Chapter 1; p. 58–90. **Alan Turing's paper in which he describes algorithmic computing.**

11. Valiant L. Probably approximately correct: nature's algorithms for learning and prospering in a complex world. Basic Books, 2013 p 20.

12•• Turing A. The chemical basis of morphogenesis (1952). In: Copeland BJ. The Essential Turing. Oxford: Oxford University Press; 2013: Chapter 15; p. 519–561**. Key paper of Alan Turing in which he describes that algorithmic computing applies to biology.**

13. Turing A. Intelligent machinery, a heretical theory (c. 1951). In: Copeland BJ, editor. The Essential Turing. Oxford: Oxford University Press; 2013: Chapter 12. p. 465–4861.

14• Nangalia J, Campbell PJ. Genome sequencing during a patients journey through cancer. N Engl J Med. 2019;381:2145–56. https://doi.org/10.1056/NEJMra1910138 **In-depth review of biologic and clinical insights from cancer genomes and research**.

15. Pon JR, Marra MA. Driver and passenger mutations in Cancer. Annu Rev Pathol. 2015;10:25–50. https://doi.org/10.1146/annurev-pathol-012414-040312.

16. PubMed Search of whole genome, exome and RNA sequencing studies in cancer and hypertension. Accessed 1/15/2020.

17. Fernandes-Rosa FL, Boulkroun S, Zennaro MC. Somatic and inherited mutations in primary aldosteronism. J Mol Endocrinol. 2017;59(1):R47–63. https://doi.org/10.1530/JME-17-0035.

18. Favier J, Amar L, Gimenez-Roqueplo AP. Paraganglioma and phaeochromocytoma: from genetics to personalized medicine. Nat Rev Endocrinol. 2015;11(2):101–11. https://doi.org/10.1038/nrendo.2014.188.

19. Wanichi IQ, de Paula Mariani BM, Frassetto FP, Siqueira SAC, de Castro Musolino NR, Cunha-Neto MBC, et al. Cushing's disease due to somatic USP8 mutations: a systematic review and meta-analysis. Pituitary. 2019;22(4):435–42. https://doi.org/10.1007/s11102-019-00973-9.

20. Pinggera A, Negro G, Tuluc P, Brown MJ, Lieb A, Striessnig J. Gating defects of disease-causing de novo mutations in Cav1.3 Ca2+ channels. Channels (Austin). 2018;12(1):388–402. https://doi.org/10.1080/19336950.2018.1546518.

21. Weakley SM, Jiang J, Kougias P, Lin PH, Yao Q, Brunicardi FC, et al. Role of somatic mutations in vascular disease formation.

Expert Rev Mol Diagn. 2010;10(2):173–85. https://doi.org/10.1586/erm.10.1.

22. Jones G, Robertson L, Harrison R, Ridout C, Vasudevan P. Somatic mosaicism in ACVRL1 with transmission to several offspring affected with severe pulmonary arterial hypertension. Am J Med Genet A. 2014;164A(8):2121–3. https://doi.org/10.1002/ajmg.a.36568.

23• Dou Y, Gold HD, Luquette LJ, Park PJ. Detecting somatic mutations in normal cells. Trends Genet. 2018;34(7):545–57. https://doi.org/10.1016/j.tig.2018.04.003 **Paper describing how with increasing numbers of samples undergoing genome sequencing, bioinformatic analysis will enable the characterization of somatic mutations and their impact on non-cancer tissues.**

24. Arel I, Rose DC, Karnowski TP. Deep machine learning - a new frontier in artificial intelligence research. IEEE Comput Intell Mag. 2010;5(4):13–8.

25. Freedman DH. Hunting for new drugs with AI. Sci Am. 2020;322:S4–7.

26• Ravi D, Wong C, Deligianni F, et al. Deep learning for health informatics. IEEE J Biomed Health Inform. 2017;21(1):4–21. https://doi.org/10.1109/JBHI.2016.2636665 **This article presents a comprehensive up-to-date review of research employing deep learning in health informatics, providing a critical analysis of the relative merit, and potential pitfalls of the technique as well as its future outlook.**

27. Tipnis SR, Hooper NM, Hyde R, Karran E, Christie G, Turner AJ. A human homolog of angiotensin-converting enzyme. J Biol Chem. 2000;275:33238–43. https://doi.org/10.1074/jbc.M002615200.

28. Donoghue M, Hsieh F, Baronas E, Godbout K, Gosselin M, Stagliano N, et al. A novel angiotensin-converting enzyme-related carboxypeptidase (ACE2) converts angiotensin I to angiotensin 1-9. Circ Res. 2000;87(5):E1–9. https://doi.org/10.1161/01.res.87.5.e1.

29. Malha L, Mueller FB, Pecker MS, Mann SJ, August P, Feig PU. COVID-19 and the renin-angiotensin system. Kidney Int Rep. 2020;5:563–5. https://doi.org/10.1016/j.ekir.2020.03.024.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.