Am J Transl Res 2023;15(10):6255-6263
www.ajtr.org /ISSN:1943-8141/AJTR0152443

Review Article

Arguments for and against the whole-genome sequencing of newborns

Csaba Szalai[1,2]

[1]*Department of Genetics, Cell and Immunobiology, Semmelweis University, 1089 Budapest, Hungary;* [2]*Heim Pál Children's Hospital, 1089 Budapest, Hungary*

Received July 26, 2023; Accepted October 11, 2023; Epub October 15, 2023; Published October 30, 2023

Abstract: Recent decades have brought enormous progress in both genetics and genomics, as well as in information technology (IT). The sequence of the human genome is now known, and although our knowledge is far from complete, great progress has been made in understanding how the genome works. With the developments in storage capacity, artificial intelligence, and learning algorithms, we are now able to learn and interpret complex systems such as the human genome in a very short time. Perhaps the most important goal of learning about the human genome is to understand diseases better: how they develop; how their processes can be prevented or slowed down; and after diseases have developed, how they can be cured or their symptoms alleviated. The vast majority of diseases have a genetic background, i.e., genes, sequence variations, and gene-gene interactions play a role in most diseases to a greater or lesser extent. Accordingly, the first step is to discover which genes, or genomic variants, cause or contribute to the development of a particular disease in a given patient. Given that an individual's genome remains virtually unchanged throughout their life (with one or two exceptions, such as in the case of cancer, which is caused by somatic mutations), it might be considered advantageous to sequence the genome of every person at birth. In this paper, we set out to show the possible benefits of sequencing the entire genome of every human being at birth, while also discussing the main arguments against it.

Keywords: Whole-genome screening, newborn, personal health record, artificial intelligence, monogenic diseases, multifactorial diseases, pharmacogenomics, decision support system

Arguments for whole-genome sequencing of newborns

It is already technically possible to sequence the entire genome of a newborn infant. The Novaseq instrument, for example, can sequence one genome per hour [1]. Currently, the net price of sequencing a human genome is between $600-800, but several biotechnology companies have indicated that a price of $100/genome is feasible in the not-too-distant future. However, while sequencing the genome is now relatively straightforward, evaluating the huge amount of data obtained from sequencing is a major challenge. A human genome consists of more than 3 billion nucleotides, while each person, has two genomes (one maternal and one paternal). The total sequencing data are 60-160 GB in size. Nevertheless, there are already several software programs that can extract a lot of useful information from the data [2].

This set of data would form a key part of the personal health record of the newborn, accessible to specialists and physicians throughout the life of that individual. The consequence of this is that, as genomics and informatics develop, more and more things can be learned from the data, pertinent to the treatment and prevention of diseases.

Moreover, with their genome sequenced, it would be possible to find out immediately if the newborn has a known hereditary disease. One study suggests that there are currently 388 hereditary diseases that develop in childhood for which effective treatment is already available, provided the disease is detected in time [3]. According to some estimates, this number will increase to 1,000 by 2030, but with recently discovered and continuously developing genome editing technologies, it can be assumed that most genetic errors causing hereditary dis-

eases can be corrected with gene therapy or genome editing in the future.

Hereditary diseases are rare individually, but with more than 7,000 so-called rare diseases (the vast majority of which are monogenic), it is estimated that 3.5-5.9% of the human population are affected, i.e., 263-446 million people worldwide suffer from one hereditary disease or another [4]. This high number shows that, taken together, hereditary diseases are in fact quite prevalent.

There are various diseases for which early detection is crucial for successful therapy. Spinal muscular atrophy (SMA), for example, which has a reported prevalence at birth of 9.4 per 100,000 live births [5], can be treated effectively with gene therapy but the treatment cannot reverse the nerve damage already caused. The earlier SMA is recognized and therapy is started, the less damage SMA does to the patient.

Most monogenic, inherited diseases develop in childhood, but there are also genetic diseases that only develop in adulthood. The American College of Medical Genetics (ACMG) has identified 73 genes whose defects can cause late-onset diseases, but if these genetic defects are identified, it is possible to prevent or slow down the development of the associated diseases [6]. Examples include high cholesterol, which is caused by defects in several genes; atherosclerosis and myocardial infarction, which may be caused by defects in *LDLR* or *APOB*; or breast and ovarian cancer caused by a defect in the *BRCA1* gene. Whole genome sequencing (WGS) of individuals from the UK Biobank found that 4.1% of the 149,960 people sequenced carried an actionable genotype in one of the 73 genes [7].

Each person carries > 100 defective genes that do not cause disease, but if their spouse also has one of these defective genes, then their child has a 25% chance of suffering from a serious monogenic disease [8]. If it is established that the odds of a genetic disease are high, then in-vitro fertilization and preimplantation genetic diagnosis can allow the implantation of a healthy embryo, resulting in a healthy newborn. In the long term, such preventive measures could also mean a reduction in the number of harmful mutations carried by mankind. Currently, with the development of medical science, this number is increasing, as many people who would not have reached reproductive age in the past due to a genetic defect are having children and passing on harmful mutations.

There are also genes that influence the response to certain medications and treatments. Some genetic variations can result in certain medicines not being effective in the given individual; others are associated with serious side effects. With knowledge of the genome, the appropriate IT software, and appropriate decision support, clinicians could provide the optimal therapy. This science of pharmacogenomics is also in intensive development. About 40% of medicines in clinical trials could be classified as precision therapeutics, i.e., the treatment depends on the genetic background of the individual. In oncology, this percentage rises to 75% [9].

If each person were sequenced, the results could be stored in a personal database. These databases could also be continuously updated with additional parameters throughout the life of the individual. For example, any diseases suffered by that person, the results of individual medical examinations, personal parameters, all information on smoking, alcohol consumption, diet, education, infections, etc. Artificial intelligence could integrate the data of all individuals using a learning algorithm. It could then analyze the genetic background and phenotypes of the individuals along with various other measurements. These could then be used to predict with increasing accuracy the likely diseases, therapeutic responses and even the optimal diet and characteristics of any given individual or demographic, based on the various genomes and other data.

The most common diseases are not caused by a single gene defect, but are rather the result of the interaction of many genes and the environment. These are the so-called polygenic, complex, or multifactorial diseases. These include atherosclerosis (which is the leading cause of death in the developed world), diabetes, hypertension, obesity, Alzheimer's disease, depression, allergies, and asthma. For these diseases, it is very difficult to find a usable gene-disease relationship, as the disease often develops as a result of the interaction of several hundred

genes, each with its own small effects, and the environment. There is already a polygenic risk score (PRS) that can be calculated for several diseases - and for the given person, based on his or her genetic background - which can be used to screen out those at high genetic risk [10]. Its clinical utility is currently small: many high-risk people do not get sick, while many people with low scores do get sick. This is a reflection of the limitations of our current knowledge; however, with the help of the data collection and learning algorithms described in the previous paragraph, the polygenic risk score can be expected to predict, with ever-increasing precision, who is at risk of developing certain diseases, and to indicate countermeasures accordingly. In addition, knowing the genetic backgrounds of patients can also speed the search for new effective therapies.

One suggested alternative to whole genome sequencing (WGS) is to take a selective approach and look only at genes whose defects cause a known disease with a childhood manifestation and whose development can be prevented or symptoms alleviated. However, this selective approach is severely limited as it does not detect disease-causing mutations in regulatory regions, non-coding RNA genes, or as-yet unknown genes whose defects also cause disease. With newborn WGS, all genetic variations can be detected, and later discoveries can be immediately exploited.

The current efficiency of WGS is demonstrated by the fact that the 100,000 Genome Project in the UK recently screened 4,660 patients with unknown rare diseases. These are patients who could not otherwise have been diagnosed with traditional diagnostic and targeted genetic tests. WGS revealed that 35% of the patients most likely had a monogenic disease, while 11% had a complex background [11]. It is likely that this ratio and the reliability of the result will improve in the future.

Since these data, together with a sophisticated decision support system, would be available on the computer of every treating physician, the given physician would be able to apply optimal therapy for most diseases, combining the patient's genomic information with their previous medical history and current symptoms, and thus greatly increasing the likelihood of a rapid and successful treatment.

Two projects are currently underway to test the feasibility of full genome screening in newborns, and whether it can be offered to all newborns in the future. The Newborn Genomes Program, conducted by Genomics England, will involve 200,000 newborns with parental consent. While everything will be tested, parents will only be informed about the detection of 200 known genetic variations that lead to diseases that develop before the age of 5. All can be treated, with treatments ranging from vitamin supplementation to bone marrow transplantation. It is estimated that, of the 200,000 studied, 500 affected children will be identified [12]. If the same proportion is extrapolated to the UK and the rest of the world, it means 3,000 and 360,000 sick but treatable children a year, respectively. This would result in earlier treatment and better prognosis for these children, and improved quality of life for both the children and their parents/caregivers. The second project is the New York City project, which was launched by Columbia University in September 2022, with a plan to sequence 100,000 newborns over 4 years. Parents will be informed about 160 treatable genetic diseases and, if they wish, about 100 additional neurodevelopmental disorders that cannot be cured, but can be alleviated with speech and physical therapy [13].

Of course, understanding the functioning of the genome, as well as developing methods necessary to identify the gene-disease relationships, are not negligible challenges. A significant part of the genome cannot be sequenced with the cheap, fast, and relatively accurate second-generation sequencing currently used for WGS. It was only in 2021 that the remaining 8% of the human genome was identified with the newly developed third-generation or long-read sequencing. This method is currently more expensive and less accurate than the second-generation sequencing, but it is much better at sequencing repetitive regions, detecting epigenetic modifications, and identifying haplotypes [14]. These methods are currently still under development. There is still a lot of work to be done to understand topologically associated domains and inter-chromosomal interactions, as well as the 3D structure of the genome in gene expression and the possible causes of disease. However, these all seem to be obstacles that can be overcome in the future, and

the introduction of methods currently used in newborn WGS would already represent a significant step forward in reducing the number of people with genetic diseases, slowing the development of genetic diseases in patients, and alleviating their symptoms.

Arguments against neonatal whole-genome sequencing

As related in the previous section, whole-genome sequencing in newborns has many advantages; however, it is currently not recommended anywhere, and is often even prohibited by law. In this section, we present the arguments against whole-genome sequencing in newborns.

The legal prohibition of WGS in some places has several primarily ethical reasons. First, there is the issue of privacy and data protection. Genetic data are, by their nature, extremely personal and sensitive and are therefore vulnerable to abuse. Accordingly, there are legitimate concerns about the proposition of a centralized database by which physicians and others would have access to newborns' WGS data. Currently, it is practically impossible to totally prevent these data from being leaked, e.g., by hackers. Indeed, there are many examples where even the most protected data, with access restricted to only a few individuals, has been leaked or held to ransom. The proposed WGS database would be accessed by numerous physicians and other specialists, across different departments and locations, thus greatly increasing the risk of a data leak. Leaked genomic data can lead to social stigma, be misused by insurance companies, and negatively affect employment.

The second ethical problem is related to the results for parents and, in the long term, for the children concerned. Whole genome screening will provide information on all hereditary diseases based on our current knowledge. As mentioned previously, there are currently several hundred diseases that develop in childhood where meaningful intervention is possible, in cases where the gene defect is causing the disease; additionally, there are 73 identified genes that can cause disease later on, for which there are methods to reduce the risk. On the other hand, there are currently > 7,000 known hereditary (monogenic) diseases, meaning that the vast majority of diseases currently have no effective therapy [15, 16]. Knowing that a child has an untreatable disease, especially one that manifests later, can present a serious psychological burden for the family and the person affected. The ethical question is raised: would the affected person(s) have a better quality of life without knowing of this risk?

There are two ways to mitigate this problem. One is to investigate everything, but report to the parents only those findings where meaningful intervention is possible. The other is to investigate only genes with early manifestation and possible intervention. Each of these approaches has weaknesses. First, a declaration of informed consent is required for all genetic tests. This means that in the case that a newborn is tested for all known hereditary diseases but only limited results are shared with the parents, the ethical problem would have to be explained to the parents, and they would be required to understand it in order to make the legally-required 'informed decision'. The problem here is that signing this 'informed consent' demands a level of knowledge and understanding about both genetics and ethics, that a significant proportion of the population will find challenging. Indeed, due to a lack of genetic knowledge, a significant proportion of the population has reservations about genetic testing. In a Dutch study, for example, married couples were asked if they would like to be tested for 50 hereditary diseases when they have children. Only 34% of the 504 respondents said yes [17].

The alternative proposal is the targeted testing approach. The disadvantage of targeted genetic testing is that as only a fraction of hereditary diseases are tested, a negative result does not rule out the possibility that a sick child will be born. Of course, it must be noted here that even testing for all known hereditary diseases cannot completely rule out the possibility that a sick child will be born, as our current knowledge is far from complete; however, the probability is much lower with WGS, and continues to decrease as our knowledge increases.

The situation is further complicated by the fact that a defect in a known disease-causing gene may or may not cause a disease in a person; moreover, in the case that a disease is caused, it is not certain how severe the disease will be.

These two concepts are called penetrance and expressivity in genetics, respectively. Penetrance concerns the probability that a given mutation will cause a disease, while expressivity concerns the probability that a given mutation will cause a severe disease in one carrier and a milder disease in another. There are mutations with a penetrance of 100%, but there are mutations with a lower penetrance. It is important to know that although monogenic diseases are caused by a single defective gene, the genome functions as a system, i.e., genes can influence each other's functions. It is possible that a variation in another gene may neutralize, limit, or enhance the effect of the disease-causing mutation. In addition, environmental factors can also influence the development of the disease. This can happen, for example, in the most common recessive inherited disease, cystic fibrosis. Homozygosity of the deltaF508 mutation in the *CFTR* gene results in a penetrance of 100%, but there are other mutations in this gene that have lower penetrance, meaning that certain carriers will develop cystic fibrosis, and others will not. This is called CFSPID (CF screen positive, inconclusive diagnosis) [18, 19]. Presently, little is known about the psychological implications of such inconclusive diagnoses on families.

Another similar example is Krabbe disease, a severe neurodegenerative disorder. Because it is difficult to predict whether patients with positive results will develop clinical symptoms, there are children with positive results who must simply wait for the disease to manifest. Many such parents developed depression or were severely upset upon receiving the positive results, and discovering that their child might develop a devastating neurodegenerative illness [20]. However, in recent years, there has been rapid progress in understanding the pathophysiology of the disease, predicting its severity, and developing new therapies (including gene therapy).

A similar problem is the so-called variants of uncertain significance, or VUS, meaning that it is not possible to say for many variations whether they cause a problem or are neutral. It should be noted that the whole-genome screening, data collection, and learning algorithms discussed earlier would significantly reduce the number of variants of such uncertain effects in the long term.

The problem of late-onset diseases poses similar psychological burdens. Huntington's disease is a common example. It is a dominantly inherited disease, i.e., the majority of patients inherit the defective gene from one parent. The disease begins to develop around the age of 40, with gradual mental and physical deterioration, which is accompanied by, among other things, jerking movements. Patients usually die within 10-15 years. As it is a dominantly inherited disease, if one parent is affected, there is a 50% chance that their child will also develop the disease. Disease penetrance is reduced for the *HTT* gene with a CAG repeat number of 36-39, whereas it is 100% for patients with repeat number above 40 [21]. Obviously, it is a serious psychological burden when someone knows that he or she carries the mutation that leads to the disease, and that their child will endure prolonged suffering. Currently, it is not possible to prevent the development of the disease. For this reason, this *HTT* gene is not even investigated, because if the carrier is discovered, it is a serious dilemma whether to tell the person concerned, as it can lead to serious depression. It must be added, however, that a similar psychological burden is imposed if the person is not aware of the presence of the mutation but knows that he or she has a 50% chance of developing it. There is currently some hope (based on animal experiments) that the development of the disease can be successfully prevented with gene therapy in the future [22, 23]. However, such therapies would only be feasible if the person were known to be a carrier. For these genetic variants, the solution may be that the affected person is told only if there is already a successful therapy for the disease.

In addition to ethical problems, there are other difficulties associated with whole-genome screening in newborns, although these problems may in principle be solved in the future, and are mainly a matter of social will. One such challenge is the enormous storage capacity required to store so many data on each person, for an extended period, i.e., for the rest of their lives at least. In one study, the genome data of a single individual required an average of 156 GB of space [3, 24]. Although this can be reduced by cleaning the data, it can also be increased significantly with a detailed annotation of the genome. Meanwhile, the considerable computing capacity required for the evalu-

ation of millions of genomes and other patient data using learning algorithms, as well as the sophisticated decision support systems that help physicians, would need to be made available to everyone concerned.

Another difficulty is that more than 80% of monogenic diseases have a prevalence of less than 1 in 1 million, and there are diseases that affect only a few families [4]. Most of these diseases could potentially be eliminated in the future, e.g., with gene therapy or genome editing; however, therapy development can cost billions of dollars, and pharmaceutical companies naturally want to recoup this cost. For example, a recently approved gene therapy treatment for the relatively common hemophilia B (prevalence = 1/30,000 persons) costs $3.5 million. The treatment appears to permanently cure the patient, and to save $5-5.8 million in the US, compared to a weekly treatment for the lifetime of the patient [25]. However, for rarer diseases, the cost of the therapy would be shared between fewer patients, making it prohibitively high. This means that even if the patient is diagnosed and the genetic background of the disease is known, it would theoretically be possible to develop a therapy for it, but, due to the prohibitive price, it will not be funded. Nevertheless, the genetic diagnosis can still prove useful, as it allows clinicians to identify any pathological mutations that might run in the family, and, in theory, to prevent the birth of sick children. Furthermore, if the disease is discovered immediately after birth, then parents have the opportunity to prepare for the illness of the child, and in many cases, it is possible to alleviate the symptoms of the disease with various conservative methods (e.g., diet, physical therapy, traditional medicines, etc.).

The next problem is the extremely increased demand for specialists. This problem seems insurmountable at the moment, although it can in principle be solved with adequate financial investment and training. The widespread, routine use of WGS and related analysis would require, for example, a large number of well-trained computer scientists, statisticians, psychologists, laboratory workers, and, above all, clinical geneticists who can interpret and explain the results to parents in a comprehensible manner. In addition, it is not only the clinical geneticist who must be able to interpret the results, but practically all physicians, as in such a system, the genetic background of the patient concerned should be taken into account in almost all decisions. A good decision support system can obviously help with this, but the final decision must be made by the physician. This would require considerable training, and even represents something of a paradigmatic shift in medicine.

Currently, the practice is to perform targeted genetic tests (i.e., testing for only suspected genes) when other tests (clinical, biochemical, ultrasound) indicate that the newborn may have some genetic abnormality. The problem is that the vast majority of genetic diseases cannot be screened for with this method. First, it is often not possible to tell from clinical symptoms which gene defect is the cause; moreover, the symptoms often appear only later. Duchenne muscular dystrophy, for example, is relatively common in boys, with the first symptoms appearing around the age of 3 years, and the final diagnosis made on average 2 years later, i.e., at the age of 5, when the symptoms are often severe, difficult, or irreversible [26]. However, gene therapy is available and, if started in time, can significantly improve the symptoms of the patient.

Professional organizations are currently mostly in favor of introducing targeted genetic screening, i.e., they recommend testing only for genes that cause diseases that start in childhood and can be treated with existing therapies [3, 15, 18, 27, 28]. However, as mentioned above, there are significant obstacles to the introduction of targeted genetic screening, including the increased demand for specialists and the difficulties around informed parental consent. Finally, there is some minority support for the introduction, or at least the trialing, of whole-genome screening in the neonatal period, based on the many benefits outlined in the first section; however, as we have seen, the problems here are even more difficult to solve [12-14].

Table 1 summarizes the main advantages, problems, difficulties and possible solutions of WGS in newborns.

Conclusion

A comparison of pros and cons of whole-genome screening in newborns shows that the

Table 1. Main advantages, problems, difficulties, and possible solutions of WGS in newborns

Process	Advantages, positive outcomes	Problems, difficulties	Possible solutions for the problems
Baby is born			
WGS within days		Increased cost and personnel.	Social and political will, training.
Personal database for each newborn	Genomic data is part of the personal health record, accessible to specialists and physicians throughout the life of that individual. As genomics and informatics develop, more and more things can be learned from the data, pertinent to the treatment and prevention of diseases.	Increased storage capacity and cost, risk for data leakage, misuse of genomic data. Due to a lack of genetic knowledge, a significant proportion of the population have reservations about genetic testing.	Social and political will. Development of IT can reduce the risk of data leakage. Training, AI could help in informed consent.
Evaluation of the genomic data	Knowledge about potential hereditary and risk of polygenic diseases of the newborns. It can be assumed that most genetic errors causing hereditary diseases can be corrected with gene therapy and other methods in the future. Various diseases for which early detection is crucial for successful therapy. Increasing number of late-onset diseases whose development can be prevented or slowed down. Polygenic risk score can be expected to predict, with ever-increasing precision, who is at risk of developing certain diseases, and to indicate countermeasures accordingly. It would help couples with known genomic backgrounds to avoid children with hereditary diseases. Reduction in the number of harmful mutations carried by mankind.	The vast majority of genetic diseases currently have no effective therapy. For rare diseases, the cost of therapy can be prohibitively high. Hereditary diseases with reduced penetrance and variable expressivity. Serious psychological burden for the family and the person affected. Extremely increased demand for specialists and overall genetic and genomic knowledge. Even testing for all known hereditary diseases cannot completely rule out the possibility that a sick child will be born, as our current knowledge is far from complete. Problems with variants of uncertain significance (VUS).	Continuously improving genome editing techniques and understanding of disease mechanisms. AI could integrate the genetic and environmental data of all individuals using a learning algorithm. Continuously updated decision support system could be used to predict with increasing accuracy the likely diseases, therapeutic responses and even the optimal diet and characteristics of any given individual or demographic, based on the various genomes and other data. WGS, data collection, and learning algorithms would significantly reduce the number of VUS.

arguments in favor center around better health and optimal treatment for individuals, and indeed for humanity as a whole. Counterarguments are largely ethical, technical and economic. The counterarguments currently tip the balance, largely because of the extent of the unknowns involved. The field is still evolving, and general public understanding of genetics and the ethical issues involved is currently low. IT data security is in itself a big question, which may be answered in the near future by technological development. In the long term, it will be an important task for researchers, physicians, clinical geneticists, psychologists, and bioinformaticians to solve the problems associated with newborn genome screening, as it has such clear promise for the hundreds of millions of people worldwide who suffer from preventable or treatable genetic (mono- and polygenic) diseases. Serious family and personal tragedies and unfathomable suffering could be prevented or alleviated through newborn genome screening. Solving these problems is still a long way off, with a range of complex questions yet to be answered, but the continued development of knowledge and technology suggests that we can be hopeful.

Disclosure of conflict of interest

None.

Address correspondence to: Dr. Csaba Szalai, Department of Genetics, Cell and Immunobiology, Semmelweis University, H-1089 Budapest, Nagyvárad tér 4, Hungary. Tel: +36-20-3917-041; E-mail: szalai.csaba@med.semmelweis-univ.hu; szalaics@gmail.com

References

[1] Abeyrama DK, Boyle B and Burg TM. Comparison of genotyping by sequencing procedures to determine population genetic structure. Funct Integr Genomics 2022; 23: 9.

[2] Tanjo T, Kawai Y, Tokunaga K, Ogasawara O and Nagasaki M. Practical guide for managing large-scale human genome data in research. J Hum Genet 2021; 66: 39-52.

[3] Kingsmore SF, Smith LD, Kunard CM, Bainbridge M, Batalov S, Benson W, Blincow E, Caylor S, Chambers C, Del Angel G, Dimmock DP, Ding Y, Ellsworth K, Feigenbaum A, Frise E, Green RC, Guidugli L, Hall KP, Hansen C, Hobbs CA, Kahn SD, Kiel M, Van Der Kraan L, Krilow C, Kwon YH, Madhavrao L, Le J, Lefebvre S, Mardach R, Mowrey WR, Oh D, Owen MJ,

Powley G, Scharer G, Shelnutt S, Tokita M, Mehtalia SS, Oriol A, Papadopoulos S, Perry J, Rosales E, Sanford E, Schwartz S, Tran D, Reese MG, Wright M, Veeraraghavan N, Wigby K, Willis MJ, Wolen AR and Defay T. A genome sequencing system for universal newborn screening, diagnosis, and precision medicine for severe genetic diseases. Am J Hum Genet 2022; 109: 1605-19.

[4] Nguengang Wakap S, Lambert DM, Olry A, Rodwell C, Gueydan C, Lanneau V, Murphy D, Le Cam Y and Rath A. Estimating cumulative point prevalence of rare diseases: analysis of the Orphanet database. Eur J Hum Genet 2020; 28: 165-73.

[5] Lally C, Jones C, Farwell W, Reyna SP, Cook SF and Flanders WD. Indirect estimation of the prevalence of spinal muscular atrophy Type I, II, and III in the United States. Orphanet J Rare Dis 2017; 12: 175.

[6] Miller DT, Lee K, Gordon AS, Amendola LM, Adelman K, Bale SJ, Chung WK, Gollob MH, Harrison SM, Herman GE, Hershberger RE, Klein TE, McKelvey K, Richards CS, Vlangos CN, Stewart DR, Watson MS and Martin CL; ACMG Secondary Findings Working Group. Recommendations for reporting of secondary findings in clinical exome and genome sequencing, 2021 update: a policy statement of the American College of Medical Genetics and Genomics (ACMG). Genet Med 2021; 23: 1391-8.

[7] Halldorsson BV, Eggertsson HP, Moore KHS, Hauswedell H, Eiriksson O, Ulfarsson MO, Palsson G, Hardarson MT, Oddsson A, Jensson BO, Kristmundsdottir S, Sigurpalsdottir BD, Stefansson OA, Beyter D, Holley G, Tragante V, Gylfason A, Olason PI, Zink F, Asgeirsdottir M, Sverrisson ST, Sigurdsson B, Gudjonsson SA, Sigurdsson GT, Halldorsson GH, Sveinbjornsson G, Norland K, Styrkarsdottir U, Magnusdottir DN, Snorradottir S, Kristinsson K, Sobech E, Jonsson H, Geirsson AJ, Olafsson I, Jonsson P, Pedersen OB, Erikstrup C, Brunak S, Ostrowski SR; DBDS Genetic Consortium, Thorleifsson G, Jonsson F, Melsted P, Jonsdottir I, Rafnar T, Holm H, Stefansson H, Saemundsdottir J, Gudbjartsson DF, Magnusson OT, Masson G, Thorsteinsdottir U, Helgason A, Jonsson H, Sulem P and Stefansson K. The sequences of 150,119 genomes in the UK Biobank. Nature 2022; 607: 732-40.

[8] 1000 Genomes Project Consortium, Abecasis GR, Altshuler D, Auton A, Brooks LD, Durbin RM, Gibbs RA, Hurles ME and McVean GA. A map of human genome variation from population-scale sequencing. Nature 2010; 467: 1061-73.

[9] Akhoon N. Precision medicine: a new paradigm in therapeutics. Int J Prev Med 2021; 12: 12.

[10] Mars N, Lindbohm JV, Della Briotta Parolo P, Widén E, Kaprio J and Palotie A; FinnGen, Ripatti S. Systematic comparison of family history and polygenic risk across 24 common diseases. Am J Hum Genet 2022; 109: 2152-62.

[11] 100,000 Genomes Project Pilot Investigators, Smedley D, Smith KR, Martin A, Thomas EA, McDonagh EM, Cipriani V, Ellingford JM, Arno G, Tucci A, Vandrovcova J, Chan G, Williams HJ, Ratnaike T, Wei W, Stirrups K, Ibanez K, Moutsianas L, Wielscher M, Need A, Barnes MR, Vestito L, Buchanan J, Wordsworth S, Ashford S, Rehmström K, Li E, Fuller G, Twiss P, Spasic-Boskovic O, Halsall S, Floto RA, Poole K, Wagner A, Mehta SG, Gurnell M, Burrows N, James R, Penkett C, Dewhurst E, Gräf S, Mapeta R, Kasanicki M, Haworth A, Savage H, Babcock M, Reese MG, Bale M, Baple E, Boustred C, Brittain H, de Burca A, Bleda M, Devereau A, Halai D, Haraldsdottir E, Hyder Z, Kasperaviciute D, Patch C, Polychronopoulos D, Matchan A, Sultana R, Ryten M, Tavares ALT, Tregidgo C, Turnbull C, Welland M, Wood S, Snow C, Williams E, Leigh S, Foulger RE, Daugherty LC, Niblock O, Leong IUS, Wright CF, Davies J, Crichton C, Welch J, Woods K, Abulhoul L, Aurora P, Bockenhauer D, Broomfield A, Cleary MA, Lam T, Dattani M, Footitt E, Ganesan V, Grunewald S, Compeyrot-Lacassagne S, Muntoni F, Pilkington C, Quinlivan R, Thapar N, Wallis C, Wedderburn LR, Worth A, Bueser T, Compton C, Deshpande C, Fassihi H, Haque E, Izatt L, Josifova D, Mohammed S, Robert L, Rose S, Ruddy D, Sarkany R, Say G, Shaw AC, Wolejko A, Habib B, Burns G, Hunter S, Grocock RJ, Humphray SJ, Robinson PN, Haendel M, Simpson MA, Banka S, Clayton-Smith J, Douzgou S, Hall G, Thomas HB, O'Keefe RT, Michaelides M, Moore AT, Malka S, Pontikos N, Browning AC, Straub V, Gorman GS, Horvath R, Quinton R, Schaefer AM, Yu-Wai-Man P, Turnbull DM, McFarland R, Taylor RW, O'Connor E, Yip J, Newland K, Morris HR, Polke J, Wood NW, Campbell C, Camps C, Gibson K, Koelling N, Lester T, Németh AH, Palles C, Patel S, Roy NBA, Sen A, Taylor J, Cacheiro P, Jacobsen JO, Seaby EG, Davison V, Chitty L, Douglas A, Naresh K, McMullan D, Ellard S, Temple IK, Mumford AD, Wilson G, Beales P, Bitner-Glindzicz M, Black G, Bradley JR, Brennan P, Burn J, Chinnery PF, Elliott P, Flinter F, Houlden H, Irving M, Newman W, Rahman S, Sayer JA, Taylor JC, Webster AR, Wilkie AOM, Ouwehand WH, Raymond FL, Chisholm J, Hill S, Bentley D, Scott RH, Fowler T, Rendon A and Caulfield M. 100,000 genomes pilot on rare-disease diagnosis in health care - preliminary report. N Engl J Med 2021; 385: 1868-80.

[12] Spiekerkoetter U, Bick D, Scott R, Hopkins H, Krones T, Gross ES and Bonham JR. Genomic newborn screening: are we entering a new era of screening? J Inherit Metab Dis 2023; 46: 778-795.
[13] Chung WK, Berg JS, Botkin JR, Brenner SE, Brosco JP, Brothers KB, Currier RJ, Gaviglio A, Kowtoniuk WE, Olson C, Lloyd-Puryear M, Saarinen A, Sahin M, Shen Y, Sherr EH, Watson MS and Hu Z. Newborn screening for neurodevelopmental diseases: are we there yet? Am J Med Genet C Semin Med Genet 2022; 190: 222-30.
[14] Conlin LK, Aref-Eshghi E, McEldrew DA, Luo M and Rajagopalan R. Long-read sequencing for molecular diagnostics in constitutional genetic disorders. Hum Mutat 2022; 43: 1531-44.
[15] Watson MS, Lloyd-Puryear MA and Howell RR. The progress and future of US newborn screening. Int J Neonatal Screen 2022; 8: 41.
[16] Amberger JS, Bocchini CA, Scott AF and Hamosh A. OMIM.org: leveraging knowledge across phenotype-gene relationships. Nucleic Acids Res 2019; 47: D1038-D1043.
[17] Plantinga M, Birnie E, Abbott KM, Sinke RJ, Lucassen AM, Schuurmans J, Kaplan S, Verkerk MA, Ranchor AV and van Langen IM. Population-based preconception carrier screening: how potential users from the general population view a test for 50 serious diseases. Eur J Hum Genet 2016; 24: 1417-23.
[18] Remec ZI, Trebusak Podkrajsek K, Repic Lampret B, Kovac J, Groselj U, Tesovnik T, Battelino T and Debeljak M. Next-generation sequencing in newborn screening: a review of current state. Front Genet 2021; 12: 662254.
[19] Cystic Fibrosis Foundation, Borowitz D, Parad RB, Sharp JK, Sabadosa KA, Robinson KA, Rock MJ, Farrell PM, Sontag MK, Rosenfeld M, Davis SD, Marshall BC and Accurso FJ. Cystic fibrosis foundation practice guidelines for the management of infants with cystic fibrosis transmembrane conductance regulator-related metabolic syndrome during the first two years of life and beyond. J Pediatr 2009; 155 Suppl: S106-16.
[20] Ghabash G, Wilkes J, Barney BJ and Bonkowsky JL. Hospitalization burden and incidence of krabbe disease. J Child Neurol 2022; 37: 12-9.
[21] Kay C, Collins JA, Miedzybrodzka Z, Madore SJ, Gordon ES, Gerry N, Davidson M, Slama RA and Hayden MR. Huntington disease reduced penetrance alleles occur at high frequency in the general population. Neurology 2016; 87: 282-8.
[22] Morelli KH, Wu Q, Gosztyla ML, Liu H, Yao M, Zhang C, Chen J, Marina RJ, Lee K, Jones KL, Huang MY, Li A, Smith-Geater C, Thompson LM, Duan W and Yeo GW. An RNA-targeting CRISPR-Cas13d system alleviates disease-related phenotypes in Huntington's disease models. Nat Neurosci 2023; 26: 27-38.
[23] Thomson SB, Stam A, Brouwers C, Fodale V, Bresciani A, Vermeulen M, Mostafavi S, Petkau TL, Hill A, Yung A, Russell-Schulz B, Kozlowski P, MacKay A, Ma D, Beg MF, Evers MM, Vallès A and Leavitt BR. AAV5-miHTT-mediated huntingtin lowering improves brain health in a Huntington's disease mouse model. Brain 2023; 146: 2298-2315.
[24] Ding S and Han L. Newborn screening for genetic disorders: current status and prospects for the future. Pediatr Investig 2022; 6: 291-8.
[25] Naddaf M. Researchers welcome $3.5-million haemophilia gene therapy - but questions remain. Nature 2022; 612: 388-9.
[26] Ciafaloni E, Kumar A, Liu K, Pandya S, Westfield C, Fox DJ, Caspers Conway KM, Cunniff C, Mathews K, West N, Romitti PA and McDermott MP. Age at onset of first signs or symptoms predicts age at loss of ambulation in Duchenne and Becker muscular dystrophy: data from the MD STARnet. J Pediatr Rehabil Med 2016; 9: 5-11.
[27] Moore AM and Richer J. Genetic testing and screening in children. Paediatr Child Health 2022; 27: 243-53.
[28] Veneruso I, Di Resta C, Tomaiuolo R and D'Argenio V. Current updates on expanded carrier screening: new insights in the omics era. Medicina (Kaunas) 2022; 58: 455.