Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

Expert Rev Precis Med Drug Dev. Author manuscript; available in PMC 2020 August 07.

Published in final edited form as:

Expert Rev Precis Med Drug Dev. 2019 ; 4(2): 95–104. doi:10.1080/23808993.2019.1599685.

Privacy and ethical challenges in next-generation sequencing

Nicole Martinez-Martin, **David Magnus**

Stanford Center for Biomedical Ethics, Stanford University, Stanford, CA, USA

Abstract

Introduction: Next-generation sequencing (NGS) is expected to revolutionize health care. NGS allows for sequencing of the whole genome more cheaply and quickly than previous techniques. NGS offers opportunities to advance medical diagnostics and treatments, but also raises complicated ethical questions that need to be addressed.

Areas considered: This article draws from the literature on research and clinical ethics, as well as next-generation sequencing, in order to provide an overview of the ethical challenges involved in next-generation sequencing. This article includes a discussion of the ethics of NGS in research and clinical contexts.

Expert opinion: The use of NGS in clinical and research contexts has features that pose challenges for traditional ethical frameworks for protecting research participants and patients. NGS generates massive amounts of data and results that vary in terms of known clinical relevance. It is important to determine appropriate processes for protecting, managing and communicating the data. The use of machine learning for sequencing and interpretation of genomic data also raises concerns in terms of the potential for bias and potential implications for fiduciary obligations. NGS poses particular challenges in three main ethical areas: privacy, informed consent, and return of results.

Keywords

Next generation sequencing; ethics; privacy

1. Introduction

Next-generation sequencing (NGS) allows rapid and relatively inexpensive sequencing of the entire genome [1 2] NGS refers to various types of sequencing platforms that can sequence millions of fragments of DNA in parallel, such as whole genome sequencing (WGS))3]. With NGS, an entire human genome can be sequenced in under 24 h[3]. NGS is being used in clinical settings, for diagnosis of hereditary and immune disorders, as well as

CONTACT Nicole Martinez-Martin nicolemz@stanford.edu Stanford Center for Biomedical Ethics, Stanford University, Stanford, CA, USA.

Declaration of interest

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

Reviewer disclosures

Peer reviewers on this manuscript have no relevant financial or other relationships to disclose.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

non-invasive prenatal diagnosis and treatment decisions for somatic cancers. [4–6], NGS offers tremendous opportunities to advance diagnosis and treatment, but the implementation of NGS in clinical care raises complex questions regarding how to appropriately manage, communicate and protect the resulting genomic data. NGS generates a substantially larger amount of potentially important personal data, and the breadth and depth of the coding variant data present a departure from previous human genetics research. Moreover, with advances in technology and the increase in avenues for accessing personal data, it has become more challenging to address risks to privacy. For these reasons, NGS in research and clinical contexts has raised concerns regarding data protection, informed consent, and return of results.

Institutional structures and ethical frameworks for privacy, communicating risks and benefits, and returning results will need to be carefully considered[7]. The amount of genetic data generated by NGS presents challenges for protecting privacy in genomic data. The application of Big Data approaches to data collection and analysis, as well as the trends toward increased data sharing, have also created challenges for providing patients and research with adequate information regarding how their data may be shared and used, as well as maintaining privacy protections. The use of NGS also presents difficulties for existing ethical frameworks for the return of results and informed consent. NGS grants access to the entire genetic profile of an individual, while, at the same time, many questions remain regarding the appropriate interpretation of this information [8,9]. Genetic variants available from NGS analysis can be grouped into four main categories: classical pathogenic mutations of known clinical relevance, mutations of probable clinical relevance, genetic variants of unknown relevance, and harmless polymorphisms [10]. A crucial area of concern is how to determine what genetic information generated by NGS needs to be conveyed to individuals and how to do so appropriately.

The distinction between tests used for diagnosis and screening has often informed discussions regarding how clinicians should inform patients of test results-This distinction can become blurred when it comes to clinical uses of NGS-NGS tests may be recommended for patients who are not symptomatic, and thus those patients will be likely to receive results that are unrelated to their initial reason for seeking genomic data or that may not be currently actionable or of known significance. The blurring of this boundary between diagnostics and screening has implications for privacy, informed consent, and decisions regarding how and which results should be returned. Furthermore, genetic results have implications not just for the individual being tested, but potentially for family members or groups with genetic ties to that individual[11]. Thus, efforts to address ethical concerns regarding privacy or informed consent may need to take into account interests beyond the individual.

Certain aspects of NGS applications also present challenges for the traditional boundary between research and clinical care. NGS technology moved relatively quickly from the domain of research to clinical application. Whole genome sequence data often moves from clinical to research contexts, as well as back again, from research to clinical settings[12]. NGS sequencing is often conducted at research laboratories, which have different standards and regulations than for laboratories that produce results intended for clinical use. Genetic

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

testing for clinical use must meet the standards of the Clinical Laboratory Improvement Amendments (CLIA)[13], while genetic testing for research purposes does not. NGS analysis that is not conducted at a CLIA-certified laboratory may not be as reliable as other conventional techniques for DNA sequencing[14]. NGS applications include microbiome analysis, diagnosis of infectious diseases, pathogen discovery and public health monitoring. While this article focuses on NGS of the human genome, it should be noted that the issue of establishing appropriate guidelines for NGS test validation is of concern for a wide range of NGS testing that has clinical applications[15]. Establishing standards and guidelines for the validation of NGS testing for different assays is necessary for transitioning NGS tests into clinical laboratories. [6,16] The demand for genomic data for clinical use has meant that some research data may reach clinical care without being properly validated and quality controlled, which can increase the potential for false positives or negatives. Research and clinical care also entail different relationships and obligations between researcher and subject, clinician and patient, and thus have differences in the associated ethical obligations. The complexities of interpreting NGS results will necessitate the development of standardized guidelines and procedures to ensure that patients and clinicians can be confident in the accuracy of results and minimize risks posed by the technology.

These attributes of NGS applications pose challenges for the current ethical frameworks for informed consent, data protection)and return of results)particularly when viewed in connection with trends favoring data sharing and Big Data approaches to research. The use of machine learning for genetic sequencing and interpretation also poses ethical concerns, particularly in terms of the potential for bias and implications for fiduciary obligations.

2. Privacy and data protection

Privacy and confidentiality have long been core tenets for ethical medical practice. Physicians need patients to disclose sensitive and personal information in order to provide appropriate medical care for them. In turn, patients need to be able to trust that these disclosures will not lead to harm[17]. ‘Confidentiality refers to the obligations of professionals, such as physicians, not to disclose information entrusted to them unless authorized to do so[18]. Privacy generally refers to the rights of individuals to make decisions over what personal information is shared and how. Privacy remains a concept that is notoriously difficult to define, particularly as it encompasses a tangle of associated rights and obligations that are dependent upon context 19]. In medical settings, and particularly when it comes to genetic privacy, the type of privacy at stake has been described as informational privacy [20]. Information privacy is ‘defined by how much personal information is available from sources other than the individual to whom it pertains. [9] Informational privacy involves decisions about the communication of personal information and the use of that data, as well as how information is stored, maintained, disclosed and protected[21]. NGS generates a substantial amount of genetic information, including almost all adverse protein-coding alleles in the genome for each participant[22]. The probabilistic nature of genetic information and evolving understandings of genetic data also pose challenges for stakeholders making decisions regarding how to approach and address privacy. While confidentiality obligations apply to NGS results, patients need to be

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

adequately informed of how practices such as data sharing and electronic health records may impact their data and pose risks to privacy.

The Health Insurance Portability and Accountability Act (HIPAA) regulates the privacy of patient data[23]. Title II establishes national standards of electronic health care for providers, insurance companies and employers. The Privacy Rule under HIPAA establishes guidelines and regulations for the disclosure of protected health information (PHI))24]. The Health Information Technology for Economic and Clinical Health Act (HITECH) modifies the HIPAA Privacy Rule, increasing privacy protections for genomic data by requiring confidentiality of genetic information and prohibiting disclosure of genomic information to insurance companies[25]. The Genetic Information Nondiscrimination Act (GINA) also prohibits discrimination by health insurance and employers for genetic information[26]. The push in recent years for data sharing, as well as the trends toward digital health information and Big Data methodologies, pose particular challenges for addressing privacy in NGS data [9,27]. Even though health data is protected under HIPAA, one logistical challenge for compliances is that clinical NGS datasets are massive. They also may need to be transferred into the cloud for further processing[1]. Because there is potential for misuse of NGS data by commercial entities, it is crucial for institutions implementing NGS technology to ensure that adequate data security measures, including assessment of the effectiveness of the institutional firewall, are in place[16]. Furthermore, institutions need to be aware of the potentially high cost of HIPAA-compliant security for large quantities of data[28]. Even with HIPAA in place, patients and even physicians may not sufficiently understand the implications and consequences of placing NGS data in a medical record. The use of electronic health records (EHRs) can exacerbate privacy concerns, as patients may consent to third-party health providers or even employers having access to health records, without realizing that their clinical genomic data could be in those records and thus accessible to third-parties [29].

Data sharing and secondary use of genomic data collected for research purposes have become common practices in the research community. The 21st Century Cures Act supported initiatives to facilitate the sharing of genomic and other data for research and clinical uses[30]. Many institutions have developed policies requiring data sharing for funded research projects. This secondary use of genomic data for research can reduce the cost and effort of collecting data in a research setting. The Privacy Rule states that de-identified data can be used and disclosed by covered entities without limitation [31]. De-identified data are not categorized as human subjects research and thus are not covered by current Federal Regulations for research. Data are considered ‘de-identified when all 18 of the identifying elements specified under HIPAA are removed from the data or if there is an expert determination that there is a very small risk that the data could be used to identify an individual[32]. It should be noted that there are several techniques that have been demonstrated to allow individuals to be identified through their genomic data [33], which will be discussed more below. There have therefore been efforts to find an appropriate balance between the sharing of genomic data for scientific purposes, and minimizing privacy risks for individuals, such as limiting the proportion of the genome that is released or statistically degrading the data[34]. Although there are not requirements for sharing clinical genomic data, clinical genomic data could potentially be used for research if the identifiers

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

are removed. Research consent forms often address the use of de-identified data, but clinical NGS consent forms do not necessarily require or provide a discussion of the use of deidentified data. EHRs, which may contain NGS reports and data, are widely used for research purposes and thus pose another potential avenue through which a person s genomic data might be shared without them being specifically aware[12].

What are the potential risks to a person whose genomic data is made available to third parties? Genomic data could be used to make inferences regarding a person s current and future health. Such information could potentially have negative repercussions for a person in terms of insurance, employment or legal concerns. NGS data does not only pose privacy risks for individuals but also can raise issues in relation to groups and family members. Research on a small portion of a given group could be used to generalize about the larger group, leading to overgeneralizations or even stigma [35]. Some members ofthat group may feel violated at the thought that others feel like they ‘know information about the members of that group [36].

The forensic use of NGS also presents challenges and risks regarding privacy. DNA analysis is an important tool in forensic science, such as child custody cases and identification of individuals for legal and criminal justice purposes. The majority of current DNA forensic tests analyze the variation in short tandem repeat (STR) markers[37]. NGS improves on current methods by allowing investigators to detect multiple STR loci on autosome and sex chromosomes, providing data on loci across the genome and address a wider range of questions in a single run, including generating potential physical descriptions of suspects[38]. A primary problem with current forensic tests is that the available sample may be limited in amount or degraded. As the amount of DNA input required for preparation of NGS libraries decreases, it is thought that nearly any sample could be sequenced, maximizing the information that can be obtained from any biological sample. [39] Law enforcement agencies have also increased efforts to share DNA information, in turn prompting international databases to expand their analyses of loci. These improvements to locus databases have improved the efficiency of law enforcement investigations[40]. At the same time, U.S. law enforcement has been using techniques that connect the DNA held in forensic databases to the DNA libraries gathered by consumer and genealogy databases in order to locate suspects, often through identification of relatives in the database[41]. For example, part of the work that led to the identification of the Golden State Killer, a serial killer active in California, decades after his crimes, involved utilizing GEDMatch, a genealogy website, to build a family tree for the killer [42]. While such measures can help to resolve crimes, they can also raise privacy concerns. It is estimated that with NGS and the growth of databases like GEDMatch in a few years it will be possible to identify almost anyone in the US from a DNA sample, even if they had not voluntarily placed their genetic information in the public domain. The Fourth Amendment of the U.S. Constitution is meant to provide protection against government search, but how it applies to this type of law enforcement practice that allows people to be identified through long-range familial searches has yet to be determined.

Considering the number of studies demonstrating the ease with which de-identified genomic data can be re-identified [43,44], the use of de-identified data does not ameliorate privacy

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

concerns. As early as 2004, Lin et al. demonstrated that an individual could be uniquely identified by having access to single-nucleotide polymorphisms (SNPs) from that person 34]. The availability of genomic data online, through genomic data repositories and databases, combined with advances in data analysis techniques, make it increasingly difficult to eliminate the potential for reidentification of genomic data[45]. The risk for re-identification becomes greater when there is more information available regarding mutations in an individual s genome [46]. There has been debate regarding what the proper threshold for identifiability is when it comes to genomic data. Some have suggested that the distinction between identifiability and non-identifiability may no longer be a useful construct for addressing expectations about information privacy, while some have suggested that identifiability should be located along a continuum rather than as a binary. [47 48],

With NGS moving increasingly into clinical applications, patients and institutions will need to consider how to inform patients of the privacy risks that might arise from reidentification of the data and the sharing of clinical genomic data for research purposes. Some have advocated that genomic data should not be treated as de-identified information[22]. At the same time, in the modern data environment, it may be impossible to eliminate the possibility that data can be re-identified. Sheri Alpert suggests that the appropriate focus for privacy concerns is minimizing the potential for misuse of the data, rather than stopping access[15]. In order to address the potential for misuse, efforts to increase transparency are important, as well as options that allow individuals more control over their data and ability to consent to downstream uses of their data. More broadly, institutions will need to coordinate efforts to minimize potentially harmful downstream data linkage and identifications of individuals[49]. Efforts to understand the risks and benefits of big data research in healthcare and formulate ethical standards should be harmonized with the formulation of ethical guidelines for NGS applications! 44]. Furthermore, the limitations of privacy protections should be clearly communicated to patients. Clinical repositories that turn over genomic data to researchers in ‘de-identified fashion could insist that no re-identification will take place as a condition of access unless there is consent. It will also be necessary to implement privacy practices in ways that do not place overly restrictive burdens on scientific research.

3. Informed consent

Informed consent is meant to ensure that an individual understands the purpose of a procedure or treatment, the risks, and benefits of the procedure, and alternative options[50]. For genetic testing in clinical and research settings, the overall goal of the informed consent process is to let people know what type of information the genetic test may reveal so that they can make decisions regarding whether to take the test and how to prepare for potential results. Generally speaking, genetic results contain complexities that are difficult for the average person to comprehend. NGS does not just yield a large amount of data, but also different types of results that range in what is known regarding the genetic variant and the probability that a given condition will occur. Therefore, there are considerable challenges in deciding on an informed consent process that adequately addresses providing relevant information in a reasonable amount of time[51]. There are also concerns regarding how to effectively communicate such information about genetic variants and the limitations of particular tests 52].

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

There are differences in the principles and obligations guiding informed consent in research and clinical contexts. The goal of the research is to produce generalizable data that advances science and benefits society, while clinical contexts are focused on providing care that will benefit the individual patient. This distinction leads to some differences in how NGS testing is framed for informed consent purposes. Because the participant will not receive direct benefit from the procedure itself, it is considered particularly important in a research setting to provide an understanding of the risks and benefits of participation, so that the participant may make an informed and voluntary consent. Informed consent also is used to help avoid the therapeutic misconception, in which a research participant inaccurately believes that a procedure conducted for research purposes is meant to have a therapeutic effect for the individual[53]. NGS, though, is being increasingly used for research on diseases that are rare or difficult to diagnose, where the research goals include identifying a genetic abnormality in order to improve treatments, thus complicating the traditional distinction between research and clinical testing[54]. When the division between research and clinical care is not clear, it can increase the risk of therapeutic misconception and that non-relevant genetic variants are over-reported. While clinical consent covers many of the same issues as research consent, it involves different obligations. Informed consent for the clinical use of NGS involves a balance between providing sufficient information for informed decision-making, while not overwhelming the person with information that is not actionable or of unknown significance. Informed consent in clinical settings also needs to convey how the testing is expected to provide benefit for the care of the patient.

Informed consent for genetic testing in clinical contexts generally involves informing patients of the risks and benefits of the testing, the implications of positive or negative results, limitations, alternatives, protections for information, future use of data and follow-up after testing[55]. Types of issues that would need to be covered in an informed consent process for NGS would be the (1) range of seriousness of each condition, (2) the range of penetrance, (3) the range of certainty, and (4) false-positive rates. When a patient misunderstands the information that they are given, it can potentially lead to worse outcomes than if they had not been given the data. False positives are another area of potential concern, because of anxiety over the results or worse outcomes resulting from a patient acting on results without first confirming them. The informed consent process should address what type of information the patient wishes to be informed of because unwanted knowledge can cause distress. For example, someone may receive results that reveal false paternity or traits that will cause distress but cannot be acted upon. One of the risks of NGS that needs to be accounted for in the informed consent process involves the challenges that come from interpreting and communicating information about genetic variants, particularly those of unknown significance. Furthermore, because certain minority groups have been less likely to be participants in genomic research, they are more likely to have variants of unknown significance[56]. This means that the benefits of NGS may be fewer for members of these groups, and it will be important for patients to be informed of this issue as part of their decision-making regarding the risks and benefits of testing.

NGS is being increasingly recommended for patients before they exhibit symptoms of a disorder. These kinds of uses of NGS are more akin to screening than diagnosis. Diagnostic tests have different trade-offs than screening. Traditionally, diagnostic tests occur within the

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

context of clinical care for an individual, with a goal of understanding and addressing symptoms that the patient is already experiencing. In contrast, the use of NGS for asymptomatic individuals places an emphasis on proactive preventive care. The difference in clinical utility for these different approaches to NGS testing has implications for the risks and benefits of testing and therefore may need to be considered as part of the informed consent process.

Genetic information derived from NGS will likely have broader implications for how clinicians make treatment recommendations and allocate resources for patients. It will, therefore, be important for patients to be informed of the potential for NGS results to impact different aspects of their long-term care. For example, clinicians caring for critically ill children with congenital heart disease (CHD) anticipate that whole genome sequencing results will be used in ways that have an impact on declarations of futility, withdrawal of care and rationing of care for critically ill children[57]. They anticipated such use despite uncertainty about the accuracy of the testing and adequacy of testing validation. Clinicians use genetic finding regarding developmental delay, cognitive impairment, and poor prognosis to make high-stakes decisions for infants and children. As genetic psychiatric testing expands, clinicians also expect that genetically discoverable mental disorders will potentially be used to justify rationing of resources in acute care decisions[58]. For example, 25% of the children with DiGeorge syndrome develop schizophrenia, and if clinicians are able to identify which of these children are more at risk of developing schizophrenia, rightly or wrongly, it could influence their decision-making regarding recommending a risky surgery during that child s infancy[59]. It will be vital to ensure that clinicians receive appropriate education regarding the limitations of the genetic tests as well as support for understanding what implications specific results have for the quality of life and prognosis for the patient. Additional empirical research is needed to understand the implications that NGS results may have for different types of patient care decisions. Consent procedures will need to inform parents of pediatric patients, as well as adult patients, that NGS results may influence their treatment over their lifetime in ways that go beyond the initial reason that motivated the NGS testing.

Patients who have had their genome sequenced for clinical testing may be asked to consent for their data to be shared and used for research purposes[60]. Individuals who have had their whole genome sequenced may also opt to have their data re-analyzed periodically, so that test results that initially were reported as inconclusive may be updated to reflect more recent research. Thus, NGS data that is acquired for research purposes may also be applied to the treatment and diagnosis of patients or NGS research may reveal the significance of a clinical test that had previously been considered inconclusive. Such developments have contributed to the sense that NGS for clinical care has blurred the boundary between research and clinical care. [61,62], While there is potentially a wide range of benefits from the use of clinical data for research, it is important to include the implications of data sharing for privacy in the informed consent process [63].

Standardized approaches to the informed consent process are needed. As institutions implement NGS into clinical care, there have been two main approaches to informed consent. In some approaches, the ordering physician conducts the informed consent process

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

as well as pre- and post-test counseling for the patient. Often these programs offer education and support resources as ancillary services. In contrast, some programs offer extensive informed consent performed in-person by a genetic counselor. Because NGS will include too many types of traits for a fully informed consent process to be accomplished in a reasonable time frame, alternative practices for providing informed consent should be considered, such as staging the information, or using video and online interactions to conduct aspects of providing information to patients.

4. Return of results

The return of results involves complex and difficult ethical considerations. NGS technologies produce a substantial amount of information. How should that information be conveyed to individuals, and are there types of information that should not be returned as results? First, we consider some of the primary arguments for returning peoples genetic results to them. Patients are said to have a ‘right to know their medical information. Many individuals say they would want to know all of their results, including ambiguous and unvali-dated results[64]. Knowing about potential genetic conditions can lead patients to make positive changes for a healthier lifestyle. For research participants, the return of results may also provide recognition of their contribution to the research enterprise. At the same time, being presented with unvali-dated or confusing genetic information can present risks for the patient. Different patients may have very different responses to the genetic information they receive. 65–67], Receiving genomic information for certain conditions could lead to undue anxiety or negatively impact an individual s health behaviors or even physiology [68,69]. If patients are given information about a late-onset condition that is not actionable, the information may not be perceived as a benefit.

Indeed, some patients may not want to know certain information, which has sometimes been referred to as a right to ‘not know. Thus, a return of the results process should include options for a patient to refuse certain information. The right to not know also presents additional considerations for parental informed consent and return of results for children. NGS can provide information about children before they are born and reveal the potential for late-onset conditions. Return of results for children thus raises questions regarding parental rights to choose the information that their child will know and whether certain information should be withheld until the child can make the choice whether to receive the information[70].

Given the range of data that NGS can provide, there are a number of questions to address regarding the quantity and type of NGS information that should be returned to patients. For example, should patients receive the raw data from their NGS? Raw data needs to first be analyzed and interpreted in order to yield a report that is meaningful for a patient[71]. Patients may view raw data as something to which they have a right, as their property. However, clinicians may worry that releasing raw data can lead to patients becoming unnecessarily worried or requesting costly or unnecessary care.

This section focuses on the return of clinical results, but, as noted above, most genetic data has been acquired in a research context. Research and clinical laboratories are governed by

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

different standards for analyzing and interpreting findings [72]. Whether findings come from a research lab or clinic can have ethical import[73]. There may be concerns regarding the quality of the data obtained from research contexts. There are critical questions that arise regarding whether some of the genetic tests used in research have sufficient scientific foundation for being applied to clinical purposes. There are related decisions regarding what information from these analyses to include in results. Additionally, there is concern that applying population results for complex traits to patients in a clinical context may not be appropriate.

Moreover, different NGS programs have independently developed strategies for variant interpretation and reporting. NGS programs use different internal and external sources to identify genetic variants and their significance, such as the Single Nucleotide Polymorphism Database, Sorts Intolerant From Tolerant amino acid substitutions prediction tool, the Human Gene Mutation Database. NGS programs also use varied strategies, such as different expert opinions or automated or manual systems, to make determinations regarding the pathological significance of variants[28]. This lack of standardization regarding database sources and strategies for identifying genetic variants underscores the complexity of providing uniform recommendations regarding the return of results.

The range of information revealed through NGS findings includes genetic variants of known and probable clinical relevance, as well as those that are of unknown relevance. There are different implications for a patient between a finding that identifies a highly penetrant gene and a predictive gene for a particular disorder and genes that indicate a susceptibility for a condition, such as cancer or heart disease, where multiple genes may play a role or where environmental factors may exert more influence on risk. Another question is the extent to which variants of unknown significance (VUS) should be considered results that should be reported to patients. This question is further complicated by the fact that certain minority populations, such as African Americans, have been underrepresented in genetic research and thus also are more likely to have genetic variants of unknown relevance 64]. Thus, decisions about how to address variants of unknown relevance in clinical care can have a different impact on minority populations. As genomic databases improve, some VUS will become better understood. Currently, there is no consensus regarding whether there is an obligation to provide follow-up communication and contact patients or families regarding a change in interpretation for a VUS. Moreover, the practicalities of reanalysis and contact could be cumbersome[74].

There are also questions regarding the return of results that are ‘incidental findings, sometimes described as the ‘inciden-talome. [75] Historically, ‘incidental findings refers to research findings that have potential relevance for health but are unrelated to the motivating reason for the research or diagnostic test[76]. In the context of whole-genome sequencing, some argue that the term ‘incidental findings is not suitable, because even when NGS is ordered for diagnostic purposes, one would expect to generate information regarding unrelated variants of clinical utility or with health implications. For that reason, terms such as ‘unsolicited results [77] or ‘secondary findings [78] are sometimes used in place of incidental findings. The underlying idea that such findings are not incidental, in the traditional sense, has potential implications for disclosure obligations[79]. Some have

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

recommended that incidental findings should not be reported within clinical contexts [80]; on the other side, some argue that all variations in genes associated with a disease are medically relevant and should be disclosed[81]. Surveys of public attitudes indicate that most people would want to know genetic information of unclear risk, even if the results involve a condition that does not have a clear treatment or prevention available [82,83].

An important consideration in decisions regarding the return of results is whether a finding is actionable. The ACMG has issued recommendations that clinicians and laboratories, regardless of the motivating indication for testing, routinely analyze genetic sequences for pathogenic variants '... deemed to be highly medically actionable so as to detect pathogenic variants that may predispose to a severe but preventable outcome. Patients should be informed during the consent process that, if desired, they may opt out of such analysis... [84] Towards that end, the ACMG maintains a list of medically actionable genes that it recommends for return in clinical genomic sequencing.[10s]

There is variability regarding what institutions choose to include in the return of results. A study of early adopting institutions found that a majority reported that they independently developed analytic strategies at their institutions for initial variant calling, determining pathologic or benign variants and identifying VUS[18]. Generally, the most frequently used criteria for evaluating when genetic information is appro-prate to return in clinical practice are 'ACCE: analytic validity, clinical validity and ELSI (ethical, legal and social implications) [85]. The ACCE is a useful tool for decision-making, but decisions regarding which results to reveal and in what contexts remain challenging. As noted above, ACMG has used an approach of listing genetic variants that should be included in the return of results. Another strategy is 'binning, which involves placing different types of results into categories, such as'actionable or'clear benefit or 'possible benefit either as a decision tool or in order to aid discussions with patients regarding what type of results that they would like returned.

In general, however, a foundational recommendation for the return of results is that physicians who utilize NGS will need to undergo ongoing education regarding the existing databases of genetic variants and their clinical status. Furthermore, it is of utmost importance to prepare patients for potent al results, starting with the informed consent process. Preparation may be needed for potential findings, as well as a lack of findings, depending upon the case. Patients and their families may feel great disappointment if there has been a 'diagnostic odyssey, where the patient NGS to try to identify a genomic variant associated with their undiagnosed condition[18]. Counseling support is needed in these situations of managing disappointment in not finding a hoped-for answer.

5. Machine learning and NGS

Artificial intelligence (Al) and machine learning (ML) are expected to change the landscape of medical research and healthcare. [86] As ML is increasingly used to analyze massive data sets, including those associated with NGS, it is necessary to consider how the ethics of AI may impact the implementation of NGS programs. AI refers to the use of machine systems to perform intelligent tasks, such as the interpretation of medical images[87]. ML refers to

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

the methods used to ‘train a computer to recognize patterns in massive datasets, including complex data interactions, and generate the algorithms that enable the AI applications®. Advances in ML have allowed researchers to improve efforts to interpret data obtained through genomic sequencing[88]. ML is being applied to next-generation sequencing for research into phar-macogenomics and tools for genetic screening of newborns 89 90]. Further, applications for NGS include predicting the risk of illness, obesity risk, differentiating low from high confidence variants[91]. The goal is to use ML to make NGS even cheaper, faster, and better. In particular, the use of ML for NGS raises ethical concerns regarding the potential for bias, the need to recognize the limitations of AI tools, and the potential implications for the fiduciary relationship between clinician and patient [92].

The potential for bias can be separated into two primary concerns: the potential for bias in the data used to construct the algorithms and bias in the algorithms themselves. Large datasets are used in order to ‘train ML systems to identify patterns in the data, and the accuracy of the performance of the resulting algorithm depends greatly on the quality of the training and validation datasets[93]. If the dataset does not accurately reflect the population to which it is applied, then bias will be transferred to the outcomes generated by the ML algorithm. This leads to concerns that the findings of ML systems may sometimes reflect and reinforce biases in society. The problem of bias presents a particular concern in applications of ML to genetic sequencing, because of the historical lack of racial and ethnic diversity in human genetic and genomic research[94]. An algorithm generated to predict outcomes from genetic sequences may contain bias if there have not been sufficient genetic studies in relevant populations. These problems limit the benefits that people from underrepresented racial and ethnic populations can receive from NGS and other advances in genomics. Efforts to increase the diversity of populations in genomics research are critical. It is also important to consider how the algorithms themselves may be designed to perform in unethical ways, such as algorithms that may be designed to identify genetic risk for purposes, such as allocation of health-care resources, that would discriminate against certain populations.

The ML system itself generates the algorithms, leading to a ‘black box issue, in which it is difficult for even the developers themselves to evaluate the specific reasoning behind the out-comes generated 95]. This makes transparency difficult. Furthermore, as industry developers increasingly invest in advancing the use of ML for NGS[96], they are generally reluctant to share information about the workings of their ML systems for proprietary reasons. Recently, there have been calls for the development of ML systems that are also able to explain the reasoning for their findings. ML systems also can be subject to ‘automation bias, where results or findings that arise from an automated tool are perceived as inherently more objective or accurate than other sources of information and thus lead to the limitations of the ML systems being overlooked. Clinicians will need education regarding the ML systems, the data sets, and limitations, including the potential for bias[10]. Institutions will also need to take these limitations into account when formulating policy in the use of such systems and NGS data, particularly for informed consent and return of results in NGS. Finally, while it is important to be aware of the limitations of ML itself, it is also critical that there be attention paid to the systems and processes into which the AI is being integrated, in order to understand and address the ways that the use of AI for NGS may impact fiduciary

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

relationships. For example, if clinicians rely on AI-generated findings for their use of NGS in ways that substitute the judgment of the software for their own, that can have implications for clinical accountability and the physician-patient relationship that will need to be studied and addressed.

6. Conclusion

NGS brings revolutionary opportunities for applying genomic information in clinical contexts. Not only does NGS produce a massive volume of genetic data, but the interpretation of that data is still evolving. The use of NGS raises complicated challenges regarding privacy, informed consent and the return of results. The use of Big Data approaches and machine learning to research and analyze genetic data also add to the complexity of addressing these challenges.

7. Expert opinion

NGS offers rapid and comparatively inexpensive sequencing of the entire genome. Certain features of NGS pose challenges for existing frameworks for research and clinical ethics. In particular, the use of NGS in clinical and research contexts raises ethical concerns in relation to privacy, machine learning techniques, informed consent, and return of results. As NGS is increasingly utilized in clinical contexts, it will be important to develop comprehensive and standardized regulations to address these ethical challenges.

The current research landscape, which encourages the sharing of research data and utilization of Big Data techniques, influences the ethical context for NGS applications. Research collaboration and data sharing are considered critical for advancing scientific knowledge. At the same time, advances in data analysis techniques and the increasing availability of an individual s personal information in online data sets has made it easier to potentially re-identify individuals from their genomic data[97]. The risk of harm to an individual from their genetic information being identified includes repercussions in terms of insurance, social position or relationships. As NGS is increasingly used for forensic applications, there will be additional privacy implications. Institutions that utilize NGS will need to ensure that there are appropriate security and storage standards met for the massive data sets generated by NGS. Some recommendations for privacy protection have focused on how to store genomic data in ways that minimize reidentification risk. However, rather than focusing on eliminating the possibility of identification, it is more useful to consider approaches that can protect individuals from the misuse of their genomic data and potential harms.

Machine learning techniques are being applied to next-generation sequencing. Machine learning applications raise concerns regarding the potential for bias. The data used to construct the machine learning algorithm may be flawed or inaccurate, or crucial data may be missing, which can lead to results that are biased or which reinforce existing biases in society. There is also the potential for algorithms to be designed to perform functions that may be unethical. As machine learning is applied to NGS, it will be important to ensure that stakeholders receive appropriate education regarding the machine learning systems and their

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

limitations. The introduction of machine learning tools will likely also have implications for fiduciary obligations in health care that will need to be studied and addressed.

NGS produces a considerable amount of data, involving different types of information ranging from validity to certainty, and necessitating a complicated assessment of risks and benefits of receiving the available information. These qualities present challenges for informed consent and the return of results. For informed consent, the need to communicate the relevant benefits and risks must be balanced against accomplishing informed consent in a manageable timeframe without over-whelming the recipient with information. The return of results process needs to take into account the individual s preferences regarding receiving all or a selection of their genetic results, while also considering potential harms from information that is uncertain, of unknown significance, and/or inactionable. Institutions have been utilizing different standards and approaches to address these challenges. There remains a need for empirical research and continued efforts to achieve more comprehensive and uniform guidance for informed consent and return of results in research and clinical contexts.

Acknowledgments

Funding

This paper was not funded.

References

Papers of special note have been highlighted as either of interest (•) or of considerable interest (••) to readers.

1. Gullapalli RR, Desai KV, Santana-Santos L, et al. Next generation sequencing in clinical medicine: challenges and lessons for pathology and biomedical informatics. J Pathol Inform. 2012;3 DOI:10.4103/2153-3539.103013
2. Xu J Next-generation sequencing. Poole, UK: Caister Academic Press; 2014.
3. Hambuch TM, Mayfield J, Ajay S, et al. Chapter 2 - Clinical Genome Sequencing In: Kulkarni S, Pfeifer J, editors. Clinical Genomics. Boston: Academic Press; 2015 p. 21–35. D0I:10.1016/B978-0-12-404748-8.00002-2.
4. Shen T, Pajaro-Van de Stadt SH, Yeat NC, et al. Clinical applications of next generation sequencing in cancer: from panels, to exomes, to genomes. Front Genet. 2015;6 D0I:10.3389/fgene.2015.00215.
5. Di Resta C, Galbiati S, Carrera P, et al. Next-generation sequencing approach for the diagnosis of human diseases: open challenges and new opportunities. Ejifcc. 2018;29(1):4–14. [PubMed: 29765282]
6. Koboldt DC, Steinberg KM, Larson DE, et al. The next-generation sequencing revolution and its impact on genomics. Cell. 2013;155 (1):27–38 [PubMed: 24074859]
7. Chowkwanyun M. Big data, large-scale text analysis, and public health research. Am J Public Health. 2019;109(Suppl 2):5126–5127.
8. Khotskaya YB, Mills GB, Mills Shaw KR. Next-generation sequencing and result interpretation in clinical oncology: challenges of perso¬nalized cancer therapy. Annu Rev Med. 2017;68(1):113–125. [PubMed: 27813876]
9. Thiffault I, Lantos J. The challenge of analyzing the results of next-generation sequencing in children. Pediatrics. 2016;137(Supplement 0):53–57.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

10. Kochański A, Demkow U. Chapter 17 - next generation sequencing —ethical and social issues In: Demkow U, Poski R, editors. Clinical applications for next-generation sequencing. Boston: Academic Press; 2016 p. 301–307. D0I:10.1016/B978-0-12-801739-5.00017-9.

11. Knoppers BM, Nguyen MT, Sénécal K, et al. Next-generation sequencing and the return of results. Cold Spring Harb Perspect Med. 2016;6:10.

12. Kulynych J, Greely HT. Clinical genomics, big data, and electronic medical records: reconciling patient rights with research when privacy and science collide. J Law Biosci. 2017;4(1):94–132. [PubMed: 28852559]

13. Clinical Laboratory Improvement Act of 1967 (81 Stat. S36, Public Law 90–174, Sec. 5).

14. Kalia SS, Adelman K, Bale SJ, et al. Recommendations for reporting of secondary findings in clinical exome and genome sequencing, 2016 update (ACMG SF v2.0): a policy statement of the American college of medical genetics and genomics. Genet Med. 2017;19(2):249–255. [PubMed: 27854360]

15. Gargis AS, Kalman L, Lubin IM. Assuring the quality of next-gen-eration sequencing in clinical microbiology and public health laboratories. J Clin Microbiol. 2016;54(12):2857–2865. [PubMed: 27510831]

16. Shevchenko Y, Bale S. Clinical versus research sequencing. Cold Spring Harb Perspect Med. 2016;6:11.

17. Naser C, Alpert SA Protecting the privacy of medical records: An ethical analysis (White Paper). 1999 Lexington, MA: National Coaliton for Patient Rights.

18. Allen AL. Confidentiality: an expectation in health care In: Ravitsky V, Fiester A, Caplan AL, editors. The penn center guide to bioethics. New York, NY: Springer Publishing Company; 2009 p. 127–135.

19. Bennett CJ. Regulating privacy: data protection and public policy in europe and the United States. Ithaca, NY: Cornell University Press; 1992.

20 •. Alpert SA. Protecting medical privacy: challenges in the age of genetic information. J Social Issues. 2003;59(2):301–322.

This article examines how the convergence of the computerization of health records and increasingly detailed genetic information impact privacy considerations.

21 •. Alpert SA. Health care information: confidentiality, access, and good practice In: Goodman KW, editor. Ethics, computing, and medicine: informatics and the transformation of health care. New York: Cambridge University Press; 1998 p. 75–101.

This chapter addresses issues related to confidentiality and privacy for health information.

22 •. Tabor HK, Berkman BE, Hull SC, et al. Genomics really gets personal: how exome and whole genome sequencing challenge the ethical framework of human genetics research. Am J Med Genet A. 2011;155A(12):2916–2924.

This article addresses three areas in which exome sequencing/ whole genome sequencing present ethical challenges: informed consent; harms associated with data sharing; and obligations concerning unanticipated findings.

[PubMed: 22038764]

23. The health insurance portability and accountability act (hipaa). Washington, D.C.: U.S Dept. of Labor, Employee Benefits Security Administration; 2004.

24. HHS. Protecting personal health information in research: Understanding the HIPAA Privacy Rule. 2004.

25. Federal Register. Modifications to the HIPAA privacy, security, enforcement, and breach notification rules under the health information technology for economic and clinical health act and the genetic information nondiscrimination act. Other Modifications to the HIPAA Rules. 2013;78(17)5566–5702. www.gpo.gov

26. Genetic Information Nondiscrimination Act of 2007, National Human Genome Research Institute, Update as of May 2, 2007.

27. Meller R Addressing benefits, risks and consent in next generation sequencing studies. J Clin Res Bioeth. 2015;6:6.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

28 •. Milner LC, Garrison NA, Cho MK, et al. Genomics in the clinic: ethical and policy challenges in clinical next-generation sequencing programs at early adopter USA institutions. Per Med. 2015;12 (3):269–282.

This review examines the policies and practices adopted by U. S. institutions that are early adopters of clinical next generation sequencing programs. It includes a synthesis of published literature and interviews in order to examine ethical challenges faced by these programs.

[PubMed: 29771644]

29. Rothstein MA. Health Privacy in the Electronic Age. J Leg Med. 2007;28(4):487–501. [PubMed: 18066975]
30. 21st Century Cures Act, 42 USC § 300jj-52(a)(1)(A) (2016).
31. Health Insurance Portability and Accountability Act of1996 (HIPAA) Pub. L. No. 104–191, 110 Stat. 1936 (1996).
32. HealthlTSecurity. De-identification of data: breaking down HIPAA rules. HealthlTSecurity https://healthitsecurity.com/news/de-identifcation-of-data-breaking-down-hipaa-rules Published April 3, 2015. cited 2018 Dec 5
33. Lowrance WW, Collins FS. Identifiability in genomic research. Science. 2007;317(5838):600–602. [PubMed: 17673640]
34. Lin Z, Owen AB, Altman RB. Genomic research and human subject privacy. Science. 2004;305(5681):183. [PubMed: 15247459]
35. Mello M, Wolf L. The Havasupai Indian Tribe Case — Lessons for Research Involving Stored Biologic Samples | NEJM. New England Journal of Medicine. 2010;363:204–207. [PubMed: 20538622]
36 •. Alpert SA Privacy and the analysis of stored tissues. Research involving human biological materials: ethical issues and policy guidance Volume II Commissioned papers (pp. A-1-A-36). 2000 Rockville, MD: National Bioethics Advisory Commission

This article addresses the concept of medical privacy and the rights of research participants to their own biological materials.

37. Butler JM. The future of forensic DNA analysis. Philos Trans R Soc Lond B Biol Sci. 2015;370:1674.
38. Matheson S, Phenotyping: DNA. Snapshot of a criminal. Cell. 2016;166(5):1061−1064. [PubMed: 27565333]
39. Alvarez-Cubero MJ, Saiz M, Martinez-Garcia B, et al. Next generation sequencing: an application in forensic sciences? Ann Hum Biol. 2017;44(7):581−592. [PubMed: 28948844]
40. Yang Y, Xie B, Yan J. Application of next-generation sequencing technology in forensic science. Genomics Proteomics Bioinformatics. 2014;12(5):190–197. [PubMed: 25462152]
41. Molteni M A new type of DNA testing is entering crime investigations | WIRED. WIRED. [Published 2018 December 26; cited 2019 Jan 7] https://www.wired.com/story/the-future-of-crime-fighting-is-family-tree-forensics/
42. Zhao C How was the golden state killer caught? DNA from relative on genealogy website was key. Newsweek. [Published 2018 April 27; cited 2019 Jan 7]. https://www.newsweek.com/how-was-golden-state-killer-caught-dna-relative-genealogy-website-was-key-903590
43. Gymrek M, McGuire AL, Golan D, et al. Identifying personal genomes by surname inference. Science. 2013;339: 321–324. PubMed: . [PubMed: 23329047]
44. Malin B Re-identification of familial database records. AMIA Annu Symp Proc 2006.
45. Heeney C, Hawkins N, de Vries J, et al. Assessing the privacy risks of data sharing in genomics. PHG. 2011;14(1):17–25.
46. Masca N, Burton PR, Sheehan NA. Participant identification in genetic association studies: improved methods and practical implications. Int J Epidemiol. 2011;40: 1629–1642. PubMed: 22158671. [PubMed: 22158671]
47. Heeney C, Hawkins N, de Vries J, et al. Assessing the privacy risks of data sharing in genomics. Public Health Genomics. 2011;14(1):17–25. [PubMed: 20339285]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

48. McGuire AL, Gibbs RA. No longer de-identified. Science. 2006;312 (5772):370–370. [PubMed: 16627725]

49. Ienca M, Eerretti A, Hurst S, et al. Considerations for ethics review of big data health research: A scoping review. PLoS One. 2018;13:10.

50. Grady C, Cummings SR, Rowbotham MC, et al. Informed Consent. N Engl J Med. 2017;376(9):856–867. [PubMed: 28249147]

51. Sharp RR. Downsizing genomic medicine: approaching the ethical complexity of whole-genome sequencing by starting small. Genet Med. 2011;13(3):191–194. [PubMed: 21311340]

52. Rigter T, Henneman L, Kristoffersson U, et al. Reflecting on earlier experiences with unsolicited findings: points to consider for next- generation sequencing and informed consent in diagnostics. Hum Mutat. 2013;34(10)1322–1328. [PubMed: 23784691]

53. Appelbaum PS, Roth LH, Lidz C. The therapeutic misconception: informed consent in psychiatric research. Int J Law Psychiatry. 1982;5(3–4):319–329. [PubMed: 6135666]

54. Kost RG, Poppel SM, Coler BS. Informed consent for next-genera-tion nucleotide sequencing studies: aiding communication between participants and investigators. J Clin Transl Sci. 2017;1 (2):115–120. [PubMed: 28649453]

55. Bruinooge SS. American society of clinical oncology policy statement update: genetic testing for cancer susceptibility. J Clin Oncol. 2003;21 (12)2397406.

56. Landry LG, Ali N, Williams DR, et al. Lack of diversity in genomic databases hs a barrier to translating precision medicine research into practice. Health Affairs. 2018;37(5):780–785. [PubMed: 29733732]

57 •. Char DS, Lee SS-J, Magnus D, et al. Anticipating uncertainty and irrevocable decisions: provider perspectives on implementing whole-genome sequencing in critically ill children with heart disease. Genet Med. 2018;20(11):1455–1461.

Using qualitative interviews of clinicians this study addressed clinician perspectives regarding the impact of whole genome sequencing results on treatment and care decisions.

[PubMed: 29493583]

58 •. Char D Pediatric acute care decision implications of genetically discoverable mental disorders. Am J Bioeth. 2017;17(4)32–33.

This study of clinicians at a pediatric cardiac center uses qualitative interviews to examine how clinicians perceive the implications of genetically discoverable mental disorders on treatment decisions.

59. Char D, Cho M, Magnus D. Whole genome sequencing in critically ill children. Lancet Respir Med. 2015;3(4):264–266. [PubMed: 25704991]

60. Mathieu G, Groisman IJ, Godard B. Next generation sequencing in psychiatric research: what study participants need to know about research findings. Int J Neuropsychopharmacol. 2013;16(9):2119–2127. [PubMed: 23725748]

61 •. Berkman BE, Hull SC, Eckstein L. The unintended implications of blurring the line between research and clinical care in a genomic age. Per Med. 2014;11(3):285–295.

This article discusses the blurring of the line between research and clinical care in next generation sequencing. It discusses potential implications such as expanding obligations imposed on researchers and also challenges to ethical frameworks for research.

[PubMed: 25506378]

62. Bertier G, Cambon-Thomsen A, Joly Y. Is it research or is it clinical? Revisiting an old frontier through the lens of next-generation sequencing technologies. Eur J Med Genet. 2018;61(10):634–641. [PubMed: 29704685]

63. Kho AN, Pacheco JA, Peissig PL, et al. Electronic medical records for genetic research: results ofthe eMERGE consortium. Sci Transl Med. 2011;3(79):79re71.

64 •. Saulsberry K, Terry SF. The need to build trust: a perspective on disparities in genetic testing. Genet Test Mol Biomarkers. 2013;17 (9):647–648.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

This article discusses the need to involve minority groups in genetic testing.

[PubMed: 24000888]

65. Hitch K, Joseph G, Guiltinan J, et al. Lynch syndrome patients views of and preferences for return of results following whole exome sequencing. J Genet Couns. 2014;23(4):539–551. [PubMed: 24449059]

66 •. McGuire AL, Caulfield T, Cho MK. Research ethics and the challenge of whole-genome sequencing. Nat Rev Genet. 2008;9(2):152–156.

This article discusses three ethical issues relevant to whole-genome research: return of results obligations to relatives and the future use of samples and data. Recommendations are given for the management of research involving whole genome sequencing.

[PubMed: 18087293]

67. Yu JH, Crouch J, Jamal SM, et al. Attitudes of African Americans toward return of results from exome and whole genome sequencing.

68. Knoppers BM, Ma n HZ, Sénécal K. Return of genetic testing results in the era of whole-genome sequencing. Nat Rev Genet. 2015;16:553–559. [PubMed: 26239711]

69. Turnwald BP, Goyer JP, Boles DZ, et al. Learning ones genetic risk changes physiology independent of actual genetic risk. Nat Human Behav. 12 2018 D0I:10.1038/s41562-018-0483-4.

70 •. Lantos JD, Artman M, Kingsmore SF. Ethical considerations associated with clinical use of next-generation sequencing in children. J Pediatr. 2011;159(6):879–80.e1

This article examines the ethical issues related to the clinical use of next generation sequencing for pediatric patients.

71. Roy S, Coldren C, Karunamurthy A, et al. Standards and guidelines for validating next-generation sequencing bioinformatics pipelines. J Mol Diagn. 2018;20(1):4–27. [PubMed: 29154853]

72. Deans ZC, Costa JL, Cree I, et al. Integration of next-generation sequencing in clinical diagnostic molecular pathology laboratories for analysis of solid tumours; an expert opinion on behalf of IQN Path ASBL. Virchows Arch. 2017;470(1):5–20. [PubMed: 27678269]

73. Solomon S. Chapter 24 - ethical challenges to next-generation sequencing In: Kulkarni S, Pfeifer J, editors. Clinical Genomics. Boston: Academic Press; 2015 p. 403–434. D0I10.1016/B978-0-12-404748-8.00024-1.

74. Pyeritz R The coming explosion in genetic testing-is there a duty to recontact? - PubMed - NCBI. N Engl J Med. 2011;365:1367–1369. [PubMed: 21995382]

75. Kohane IS, Masys DR, Altman RB. The incidentalome: a threat to genomic medicine. JAMA. 2006;296(2):212–215. [PubMed: 16835427]

76. Wolf SM, Lawrenz FP, Nelson CA, et al. Managing incidental findings in human subjects research. J Law Med Ethics. 2008;36(2):211–219.

77. Bijlsma RM, Bredenoord AL, Gadellaa-Hooijdonk CG, et al. Unsolicited findings of next-generation sequencing for tumor analysis with in a Dutch consortium: clinical daily practice recons idered. Eur J Hum Genet. 2016;24(10):1496–1500. [PubMed: 27071717]

78. Green RC, Berg JS, Grody WW, et al. ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. Genet Med. 2013;15(7)565–574. [PubMed: 23788249]

79. Cho MK. Understanding incidental findings in the context of genetics and genomics. J Law Med Ethics. 2008;36(2):280–285, 212. [PubMed: 18547195]

80. ESHG.Guidelines for diagnostic next-generation sequencing.Eur J Hum Genet. 2016;24:2–5. [PubMed: 26508566]

81. Angrist M You never call, you never write: why return of ‘omic results to research participants is both a good idea and a moral imperative. Per Med. 2011;8(6):651–657. [PubMed: 22199990]

82 •. Bollinger J, Scott J, Dvoskin R, et al. Public preferences regarding the return of individual genetic research results: findings from a qualitative focus group study. Genet Med. 2012;14(4)451–457.

This study used focus groups to examine public understandings of the return of genetic results.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

[PubMed: 22402755]

83 •. Murphy J, Scott J, Kaufman D, et al. Public expectations for return of results from large-cohort genetic research. Am J Bioeth. 2008;8 (11)36–43.

This study used focus groups to examine the public s preferences regarding return of results in large-scale genetic research.

84. ACMG Board of Directors. ACMG policy statement: updated recommendations regarding analysis and reporting of secondary findings in clinical genome-scale sequencing. Genet Med. 2015;17(1): 68–69. [PubMed: 25356965]

85. Pitini E, Vito CD, Marzuillo C, et al. How is genetic testing evaluated? A systematic review of the literature. Eur J Hum Genet. 2018;26(5):605. [PubMed: 29422659]

86. Yu K-H, Beam AL, Kohane IS. Artificial intelligence i n healthcare. Nat Biomed Eng. 2018;2(10):719. [PubMed: 31015651]

87. All eyes are on AI. Nat Biomed Eng. 2018;2(3):139. [PubMed: 31015718]

88. Celesti F, Celesti A, Wan J, et al. Why deep learning is changing the way to approach NGS data processing: a review. IEEE Rev Biomed Eng. 2018;11:68–76. [PubMed: 29993643]

89. Libbrecht MW, Noble WS. Machine learning applications in genetics and genomics. Nat Rev Genet. 2015;16(6):321–332. [PubMed: 25948244]

90. Njage PMK, Henri C, Leekitcharoenphon P, et al. Machine learning methods as a tool for predicting risk of illness applying next-gen-eration sequencing data. Risk Anal. 11 2018 D0I:10.1111/risa.13239.

91. Van Den Akker J, Mishne G, Zimmer AD, et al. A machine learning model to determine the accuracy of variant calls in capture-based next generation sequencing. BMC Genomics. 2018;19(1):263. [PubMed: 29665779]

92 •. Char DS, Shah NH, Magnus D. Implementing machine learning in health care —addressing ethical challenges. N Engl J Med. 2018;378(11)981–983.

This article examines the ethical implications of the integration of machine learning into clinical care.

[PubMed: 29539284]

93. Towards trustable machine learning. Nat Biomed Eng. 2018;2 (10)709. [PubMed: 31015650]

94. Knerr S, Wayman D, Bonham VL. Inclusion of racial and ethnic minorities in genetic research: advance the spirit by changing the rules? J Law Med Ethics. 2011;39(3):502–512. [PubMed: 21871045]

95. Knight W Theres a big problem with AI: even its creators can t explain how it works. MIT technology review. [Published 2017 April 11; cited 2018 Dec 14]. Available from: https://www.technologyreview.com/s/604087/the-dark-secret-at-the-heart-of-ai/.

96. Marr B The amazing ways artificial intelligence is transforming genomics and gene editing. [cited 2018 Dec 15]. Available from: http://www.forbes.com/sites/bernardmarr/2018/11/16/the-amazing-ways-artificial-intelligence-is-transforming-genomics-and-gene-editing/63a5ad9d42c1

97. Hansson MG, Lochmuller H, Riess o, et al. The risk of re-identification versus the need to identify individuals in rare disease research. Eur J Hum Genet. 2016;24(11):1553–1558. [PubMed: 27222291]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

rticle highlights

The use of NGS in research and clinical contexts offers tremendous opportunities to advance understandings of health and disease, as well as improve treatments. At the same time, certain features of the research and clinical landscape for NGS present challenges for ethical frameworks for research and clinical care. Efforts to increase research collaboration and data sharing, massive data sets and the use of Big Data approaches for genomic research present challenges for privacy, in particular.

NGS data often moves back and forth between research and clinical contexts. There are differences in the ethical obligations and standards applicable in research and clinical contexts. A large portion of NGS data is generated in research laboratories, which have different standards than those applied to clinical laboratories. The boundary between research and clinical contexts for NGS has therefore blurred. This blurred boundary also has implications for informed consent and the return of results that need to be considered.

The availability of personal information in online data sets, data sharing and advances in data analysis have made it easier to re-identify individuals from their genomic data. NGS is also potentially useful for non-medical applications, such as forensic uses, which presents additional privacy concerns regarding NGS-generated data. Institutions that utilize NGS will need to ensure that there are appropriate security and storage standards for the massive data sets generated by NGS. There will also need to be attentive to protecting people from potential misuse or harmful repercussions from their genomic data.

Machine learning techniques are being applied to NGS. Machine learning applications raise concerns regarding the potential for bias. It will be necessary to educate relevant stakeholders regarding issues such as potential bias in the data and algorithms, the appropriate uses machine learning systems and their limitations. Machine learning tools also have implications for fiduciary obligations in health care that will need to be studied and addressed.