2-1-2020

Evaluation of Commercial Next-Generation Sequencing Bioinformatics Software Solutions

Rama R. Gullapalli
Departments of Pathology and Chemical and Biological Engineering, University of New Mexico, Albuquerque, New Mexico. Electronic address: rgullapalli@salud.unm.edu

Recommended Citation

Gullapalli RR. Evaluation of Commercial Next-Generation Sequencing Bioinformatics Software Solutions. J Mol Diagn. 2020 Feb;22(2):147-158. doi: 10.1016/j.jmoldx.2019.09.007. Epub 2019 Nov 18. PMID: 31751676; PMCID: PMC7031678.

The Journal of Molecular Diagnostics, Vol. 22, No. 2, February 2020

the Journal of Molecular Diagnostics
jmd.amjpathol.org

Evaluation of Commercial Next-Generation Sequencing Bioinformatics Software Solutions

Rama R. Gullapalli

From the Departments of Pathology and Chemical and Biological Engineering, University of New Mexico, Albuquerque, New Mexico

Accepted for publication September 23, 2019.

Address correspondence to Rama R. Gullapalli, M.D., Ph.D., Departments of Pathology and Chemical and Biological Engineering, University of New Mexico, Room 333A, MSC08-4640, Albuquerque, NM 87131. E-mail: rgullapalli@salud.unm.edu.

Next-generation sequencing (NGS) diagnostics continue to expand rapidly in clinical medicine. An ever-expanding menu of molecular biomarkers is deemed important for diagnostic, prognostic, and therapeutic assessment in patients. The increasing role of NGS in the clinic is driven mainly by the falling costs of sequencing. However, the data-intensive nature of NGS makes bioinformatic analysis a major challenge to many clinical laboratories. Critically needed NGS bioinformatics personnel are hard to recruit and retain in small- to mid-size clinical laboratories. Also, NGS software often lacks the scalability necessary for expanded clinical laboratory testing volumes. Commercial software solutions aim to bridge the bioinformatics barrier via turnkey informatics solutions tailored specifically for the clinical workplace. Yet, there has been no systematic assessment of these software solutions thus far. This article presents an end-to-end vendor evaluation experience of commercial NGS bioinformatics solutions. Six different commercial vendor solutions were assessed systematically. Key metrics of NGS software evaluation to aid in the robust assessment of software solutions are described. Comprehensive feedback, provided by the TriCore Reference Laboratories molecular pathology team, enabled the final vendor selection. Many key lessons were learned during the software evaluation process, which are described herein. This article aims to provide a detailed road map for small- to mid-size clinical laboratories interested in evaluating commercial bioinformatics solutions available in the marketplace. *(J Mol Diagn 2020, 22: 147—158; https://doi.org/10.1016/j.jmoldx.2019.09.007)*

Next-generation sequencing (NGS) is rapidly gaining importance as a key driver of personalized medicine efforts in clinical practice.[1,2] Clinical laboratories are the linchpin for the implementation of fast evolving NGS technologies into routine clinical practice.[1,3,4] Research-based NGS studies have identified numerous genomic, transcriptional, and epigenetic biomarkers with potential clinical applications.[1,5] In pathology specifically, molecular biomarker assessment is becoming increasingly crucial for the process of diagnosis, prognosis, and therapeutic assessment of oncology patients.[2] With the availability of an increasing number of US Food and Drug Administration—approved, targeted cancer therapies, there is increasing demand from oncologists for multiplexed detection of cancer biomarkers as opposed to the previous single biomarker paradigms.[6] The ability to detect multiple biomarkers on a single NGS platform in a multiplexed manner (ie, many patients on a single run) is a key advantage of NGS technology.[1]

Although NGS technologies are undoubtedly powerful and have the potential to alter current clinical practice dramatically, there are many hurdles to their routine clinical implementation. The key challenge of clinical NGS is also the main strength of this technique (namely, the amount of sequencing data generated).[3,7,8] NGS is a data-intensive technique that generates a large amount of data at each step of the process: i) the raw data signal calls on the instrument, ii) FASTQ data files for the bioinformatics analysis, and iii) annotation data describe the generated sequencing variants.[2,9] Depending on the type of the NGS assay, the number of sequencing variants generated range

Supported by a New Mexico—Idea Networks for Biomedical Research Excellence Institutional Development Award from the National Institute of General Medical Sciences of the National Institutes of Health under grant P20GM103451.

Disclosures: None declared.

https://doi.org/10.1016/j.jmoldx.2019.09.007

from tens to thousands to millions (eg, panel sequencing versus whole-exome versus whole-genome sequencing, respectively).[1,2,9] In clinical oncology, the number of data variants generated is high because of the inherent genomic instability associated with biological carcinogenesis. With increasing interest in newer parameters of tumor assessment (eg, tumor mutational burden, copy number changes, and epigenetic markers), issues of NGS data management are complicated even further.[1,2,9]

Clinical NGS assays are evolving quickly and growing exponentially in scope.[1,10] Yet, routine availability of NGS assays and trained clinical personnel capable of interpreting high-throughput clinical data are highly uneven. NGS assay expertise is a niche subspecialization in a clinical laboratory, available mostly at major academic centers and large commercial clinical laboratories. The implementation of NGS assays requires significant financial resources and personnel for a routine clinical laboratory. Even with the availability of financial resources to purchase clinical NGS platforms, trained personnel with expertise in bioinformatics, clinical informatics, and data analytics are hard to come by.[1,2] The lack of trained personnel capable of performing NGS bioinformatics analysis continues to be a major impediment to the widespread implementation of NGS testing.[1,2] Currently, doctoral-level bioinformaticists work alongside molecular pathologists to enable clinical NGS assay implementation. The lack of a formal curriculum in genomics and bioinformatics in pathology residency training is a significant hurdle to the widespread adoption of NGS. A vast majority of the current practitioners of NGS analytics in a clinical laboratory are self-educated, often through a process of trial and error. Yet, because of the ever-increasing complexity of NGS testing, it is a challenge for any single individual to gain mastery over all of the aspects of the NGS assay (experimental and bioinformatics).[1,2]

In the absence of adequate resources (financial and personnel), NGS assay options for many mid- to small-level institutions are restricted to send-out tests for biomarker assessment. Yet, standardized guidelines for pretherapy diagnostic assessment increasingly mandate the use of molecular biomarkers to guide therapies in newly diagnosed cancer patients.[9,11,12] It is no longer an option to ignore a patient's molecular biomarker status before the initiation of cancer treatments. Middle- to small-sized clinical laboratories increasingly need to explore NGS assay options in their own setups. For a clinical laboratory with a preexisting non-NGS molecular component, the wet/bench portion of NGS technology is reasonably accessible (ie, NGS instrumentation and molecular technologist expertise). However, the bioinformatics portion of clinical NGS testing eludes easy solutions, irrespective of the institutional size. Some institutions have developed in-house solutions to bridge the bioinformatics gap.[13] Yet, this is not easy or feasible for smaller institutions.[2,13] To overcome this shortcoming faced by smaller institutions, commercial cloud-based bioinformatics solutions have come into existence in the marketplace. Commercial vendors claim to provide solutions capable of bridging the lack of bioinformatics expertise in a clinical laboratory setting.

In this article, evaluative assessments of the different commercial NGS bioinformatics solutions available in the market are described. At the University of New Mexico (UNM)/Tri-Core Reference Laboratories (TRL), routine oncology NGS testing is performed in various tumor categories (solid tumor and hematological). A major driver behind the need for a bioinformatics upgrade was the observation that the existing bioinformatics pipelines did not scale over time (2014 to current) and increasing testing volume (increasing annually). In addition, the lack of historical variant annotation data reduced the efficiency of the medical director's weekly sign-out process in the laboratory. The overall goal of this project was to assess NGS software in a rigorous and robust manner to differentiate the key bioinformatics features of solutions available in the marketplace. Thus, this evaluation experience may provide a useful roadmap to other small- to mid-sized laboratories interested in bringing NGS-based testing (and the bioinformatics software necessary) into their own setups.

Materials and Methods

Six medical directors along with technical personnel in the laboratory drive the strategic long-term molecular diagnostics (MDx) agenda at UNM/TRL. The medical directors are collectively responsible for the planning, strategy, and execution of MDx projects originating in the division. A federated model of NGS bioinformatic analytics was adopted since starting clinical NGS assay services in 2014. A federated model implies that various components of the bioinformatics pipeline are executed by different individuals working in the MDx laboratory. TRL MDx technicians are trained to perform the initial manipulation of the raw NGS data (read mapping, BAM file generation, initial variant call reads, and quality assurance assessment). After the preliminary evaluation, each medical director in the laboratory (and clinical fellows) perform the higher-level variant assessment to issue the final NGS report associated with each patient. Cases with nondefinitive outcomes are reviewed collectively at a consensus meeting as necessary. Individuals with advanced bioinformatics expertise (R.R.G.) serve in a consultative role to drive the broad bioinformatics agenda in the division. The federated NGS bioinformatics model eliminated the need to hire specialized bioinformatics personnel, saving valuable resources for the institution.

Because of the lack of scalability of the federated NGS workflow model over time (2014 to current), a decision was made to obtain additional specialized NGS software in January 2017. The key problems associated with the original bioinformatics pipeline included the inability to review historical NGS sign-out data (by various medical directors), a lack of access to comments of previously signed-out

Table 1 A List of Potential Questions to Consider Before Software Assessments

Software questions	Laboratory workflow questions
Is the software necessary? What are the key payoffs due to the acquisition?	Is the software user friendly at all levels? (pathologists, supervisors, and technologists?)
Windows or a Linux platform?	How does the software improve the overall efficiency of the clinical sign out?
Local on-site install or cloud based?	Is the QC metrics collection improved?
Is the software scalable for different kinds of assays in the future? (eg, DNA sequencing, RNA sequencing, and microbiome data)	Is the transition likely to be disruptive? (for laboratories with existing NGS workflows)
Reputation of the software in the user laboratory professional community?	How does it help with improving the quality of the reporting?
Is the software a sole product offering of the company or one of many?	Can the software be customized to the laboratory NGS workflows?
How many FTEs are necessary for the implementation and long-term use of the software?	Ease of access to historical sign-out data in the laboratory and from other institutions?
Is the software HIPAA compliant?	What are the features sought in the desired molecular annotation database?
Pricing: one-time setup costs or a software-as-a-service pricing model?	Utility of the software at tumor board presentation (eg, NGS data displays)

A list of software and laboratory workflow-related questions considered upfront is shown in two separate columns.
FTE, full-time equivalent; HIPAA, Health Insurance Portability and Accountability Act; NGS, next-generation sequencing; QC, quality control.

clinical cases, and a lack of historical quality control (QC) data across NGS assay runs. These shortcomings caused a reinvention of the wheel for case sign out every week, reducing the overall efficiency of the NGS sign-out process. The broad goal of the intended software purchase was to enable a scalable workflow to accommodate the increasing NGS needs of the MDx division as well as enable efficiencies of the physician sign-out process. The NGS information technology (IT) assessment project obtained an executive formal go-ahead in February 2017. A team of four individuals with expertise in technical project management were selected to drive the software evaluation process. This included the following: i) two medical directors, ii) one TRL MDx technical supervisor, and iii) one administrative specialist with expertise in project management to track the data collection during the review process. Executive management at TRL reviewed the consensus decision of the medical directors, and the project was green lighted in March 2017. At this initial planning stage, two key lessons were learned, which are worthy of mention: An IT project of this evaluative scope requires a key champion to drive forward the overall project, and adequate support staff is critical to track various aspects of the review during the entire software assessment process. It is a well-known fact that resources (human and financial) are often in short supply at small- to mid-level institutions. Yet, it is critically important to allocate adequate human resources upfront (at least one project management specialist) to ensure the successful completion of NGS software evaluation.

A major challenge encountered in the initial phase of the project was determining the evaluation criteria of the NGS software assessment process.[1,2,14] To our knowledge, there are no descriptions of an NGS vendor assessment process previously reported in the literature. It was quickly learned that health care IT evaluation is nontrivial and a complicated process. Health care software assessment must involve not only the technical evaluation of the software features but also the potential impact on the end users (pathologists and technical staff). The initial stages mainly relied on the following resources to determine the assessment criteria of the in-house NGS software: i) official College of American Pathologists' molecular pathology checklists,[15] ii) Next-generation Sequencing: Standardization of Clinical Testing and additional Association for Molecular Pathology guidelines for bioinformatics pipelines,[9,10,16–18] and iii) a book, *Evaluative Methods in Biomedical Informatics*, authored by C.P. Friedman and Jeremy C. Wyatt.[19] All these resources collectively provided an excellent starting point to enable a thorough assessment of commercial NGS software solutions described in this article. After surveying the literature, a list of key preassessment considerations was identified to help the team with the downstream evaluation process. Table 1 describes some of the key preliminary questions considered by the team before initiating the software review process. The major focus in this initial evaluation period related to the nature of the software itself and an assessment of the likely impact on the established NGS workflow within our laboratory (Table 1). Many of these questions are likely to be similar in nature for other laboratories as well. It is critical for the project leadership to thoroughly assess the basic laboratory NGS requirements in discussion with the key stakeholders at the institution (ie, executive management and laboratory personnel).

Results

This section discusses the roadmap and final results of our NGS software assessment experience obtained at UNM/TRL from January 2017 to October 2018. The key

outcomes obtained over the course of the project are summarized in a separate section. Results generated from the NGS software evaluation project are described in five different phases. A workflow schematic detailing the software evaluation process is presented in Figure 1.

Phase 1: Survey—Needs Assessment

After obtaining approval from the executive management of TRL, an internal needs assessment survey was performed within the TRL MDx division. On the basis of our prior NGS workflow experience from 2014 to 2017 and key questions considered in the preassessment phase (Table 1), detailed feedback was obtained from all of the stakeholders involved in the project: medical directors, molecular fellows, technical supervisors, and technologists working within the TRL MDx division. Feedback was obtained as a free-form survey. Laboratory personnel were asked to provide detailed written feedback on the perceived shortcomings and bioinformatics needs within the laboratory. Data were then aggregated by one of the medical directors (R.R.G.), and the overall results were presented at a regular laboratory meeting. The results of this internal survey are summarized in Table 2. The broader concerns of our existing NGS bioinformatics data analysis pipeline were similar in any given category (eg, medical directors versus technicians). This provided us confidence in our overall assessment approach and outcomes. However, specific NGS workflow concerns of the medical directors and technical staff focused on different issues. Although the medical directors focused mainly on issues related to annotation, decision support, and case reporting in the laboratory information system (LIS), the technicians focused on issues related to data entry and raw bioinformatic data processing pipelines. The lack of historical report annotation to improve case sign-out efficiency was identified as a key shortcoming of our in-house bioinformatics pipeline. Thus, the initial free-form survey and internal group review helped us to focus on the key challenges associated with the existing NGS bioinformatics workflow at TRL.

Phase 2: Vendor Demonstrations/Site Visits

After gathering preliminary feedback from all the stakeholders in the UNM/TRL MDx laboratory, detailed feedback summary was provided to the TRL executive management for further action. A project manager was assigned to initiate formal contact with commercial software vendors in the marketplace. On the basis of the laboratory's initial feedback, a request for information form was generated in collaboration with the IT division at TRL (Supplemental Table S1). The standard TRL IT vendor procurement form was extensively modified to obtain additional data necessary for the NGS software assessment. The final request for information form was a 55-question information assessment form designed to obtain key

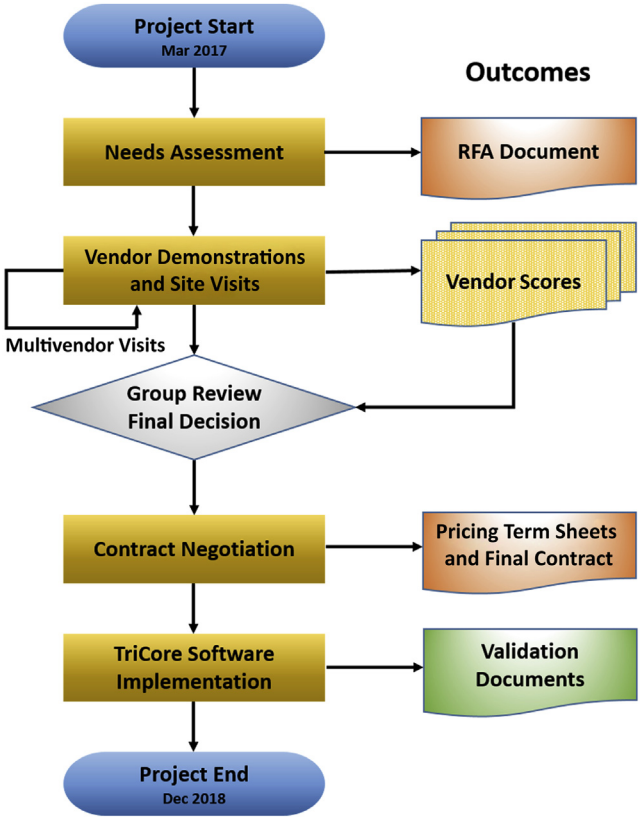

Figure 1 Schematic showing the project workflow of the next-generation sequencing bioinformatics software vendor evaluation process. RAF, request for application.

information necessary to enable the NGS bioinformatics software assessment. Six NGS bioinformatics vendors responded to the UNM/TRL request for information announcement call [GenomOncology, Cleveland, OH; GeneInsight/Mitogen LIMS, Tucson, AZ; Qiagen QCI Interpret, Germantown, MD; PieranDx, St. Louis, MO; Agilent Cartagenia (now Agilent Alissa), Santa Clara, CA; and Bina-Roche, Indianapolis, IN]. One vendor (Bina-Roche) dropped out from the final evaluation process subsequently because of product development issues. After reviewing the request for information answers, individual vendor demonstrations were scheduled over a duration of 6 months. A key point to remember at this stage of the evaluation is the lengthy duration necessary for a comprehensive NGS software assessment. Because of the time involved in the evaluative process (approximately 6 months minimum), it is critical to establish key assessment parameters by the laboratory before scheduling each individual vendor's demonstrations. It was observed that it is often difficult to remember all the key features of a software solution by the laboratory members toward the end of the assessment period. By establishing a strict set of assessment criteria a priori, it is feasible to compare the key software features in an objective manner at the end of the assessment process. It is critical for the core project team members (as

Table 2 Summary of Findings from the Preassessment Survey Conducted in the Laboratory

Summary findings of preassessment survey
Medical directors ($n = 6$)
• Easy retrieval of cross-sectional history of mutations and reports by various medical directors
• Ability to preprocess and flag known false positives a priori and update them in real time
• Streamlined and easy access to various annotation databases (eg, COSMIC and dbSNP)
• Consistent (and flexible) reporting scheme and the ability to free text
• Ability to remotely access and analyze the NGS data
• Intuitive access and incorporation of references into the report with graphics
• Integration with IGV and availability of advanced visualization capabilities
• Ease of use for translational research
• Optimal integration with the existing LIS system within the organization
Technical supervisors ($n = 2$)
• Intuitive, easy-to-use, centralized database of NGS data for all personnel (physicians, supervisors, and technologists)
• Ability to cross-reference data from cases beyond our own
• Eliminating manual transcription of data by technologists
• Easy retrieval, collation, and analysis of QC data time series for regulatory purposes
• Automatically pair (and analyze) the current bioinformatics pipelines without manual intervention
• Bidirectional data flow from database to reports to enable building the database and result reporting
Medical technologists ($n = 6$)
• Automatic initiation of data analysis protocols overnight to save technologist time
• Display accurate coverage metrics to track poorly performing primers on individual and collective clinical samples
• Minimal to no manual intervention in the data preprocessing [ie, no Excel (Microsoft Corp., Redmond, WA) spreadsheets]
• Simultaneous and automatic data analysis from various bioinformatics pipelines
• Easy way to flag for strand bias without requiring manual and visual confirmation
• Faster computer processing speeds
• Identifying suboptimal data and QC metrics automatically without manual intervention

Input was obtained via a free-form survey from individuals in the laboratory. Fourteen individuals provided the data included in this survey [medical directors ($n = 6$), technical supervisors ($n = 2$), and technicians ($n = 6$)]. Final data were collated, and key unique observations are presented in Table 2.

COSMIC, Catalog of Somatic Mutations in Cancer; dbSNP, Single Nucleotide Polymorphism database; IGV, Integrated Genome Viewer; LIS, laboratory information system; NGS, next-generation sequencing; QC, quality control.

opposed to the broader laboratory) to track these preestablished assessment criteria much more closely as each vendor's software demonstration is scheduled and completed.

Phase 3: Defining Assessment Criteria and Final Group Review

A major challenge faced in NGS software evaluation process was the lack of objective measures to compare software solutions. IT evaluation also lacks a consistent terminology to describe the key outcomes of a software assessment process. For a detailed overview of these issues, the interested reviewer is referred to chapters 5 to 8 in the book by Friedman and Wyatt.[19] As a part of the evaluation, the Delphi technique of assessment described in this book was adopted.[19] The Delphi method is a method of obtaining a consensus judgment while evaluating different attributes in an iterative manner until a convergence criterion is met (in our case, the final selection of a software).[19] A major prerequisite is, of course, defining the key assessment criteria a priori in as much detail as possible to form the basis of an objective software evaluation. An additional resource worthy of mention is the term usability, as defined by the National Institute of Standards and Technology. In health care IT, the term usability is defined by the National Institute of Standards and Technology as "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use" (ISO9241, *https://www.nist.gov/programs-projects/health-it-usability*, last accessed December 2019). The interested reader is directed to a key document expanding the definitions and use of usability in health care IT evaluation (NISTIR 7804, *https://nvlpubs.nist.gov/nistpubs/ir/2015/NIST.IR.7804-1.pdf*, last accessed December 2019). Although the National Institute of Standards and Technology document is highly useful to define usability definitions and metrics within a health care IT evaluation process, it is still incumbent on the laboratory to adopt these definitions in the context of the specific needs of the laboratory. On the basis of our prior experience with our in-house NGS bioinformatics workflow and a review of the health care IT literature, nine different metrics were established to form the basis of the NGS software evaluation process. Some of these criteria make use of the National Institute of Standards and Technology usability criteria, as applied to bioinformatics software. The final list of our nine NGS software evaluation criteria and a brief description of each are provided below. All these nine criteria were

used as a benchmark to compare each software solution against the others during the review process.

User Interface

The interface is the main window through which the user-software interactions occur (the interface exists either in the cloud or on a local install). Features for consideration were if the interface was a single or multiwindow access. Single (or minimal) window access was considered ideal for the overall user experience. NGS software solutions that required switching between multiple, disconnected browser windows were perceived to be inconvenient to maintain the pathologists' train of thought during the sign-out process. Simple white interfaces were identified to be preferable compared with colored and busy backgrounds.

Ease of Use and Data Displays

Ease of use was defined as the number of clicks needed to move between different windows containing the various elements of the NGS data. Most of the vendors separate out key elements of the NGS workflow in a tabbed manner (eg, patient NGS data, QC data, and demographic data on separate tab structures in a browser window). Some vendors had a more intuitive approach to access the final variant list and the associated annotation data of each call generated from external databases. An additional consideration was the ability to display the patient NGS read pileup data via the Integrated Genome Viewer (IGV; Broad Institute, *https://software.broadinstitute.org/software/igv*, last accessed May 6, 2019; Massachusetts Institute of Technology open license) display interface.[20,21] The original NGS workflow enabled direct visualization of BAM file pileups in IGV software before the final sign out. The ability to visualize the raw data provided confidence to the sign-out pathologists regarding the veracity of the variant calls made by the bioinformatics software pipeline. Thus, it was believed that providing a similar ability to visualize the raw data in IGV would be a key feature of interest in our software assessment. Although all vendors provide some form of sequencing pileup displays, some solutions were observed to be harder to navigate compared with others (eg, needing multiple clicks, lack of the familiar IGV interface, and/or a bare-bones display with no associated read QC metrics similar to those provided by IGV).

Intuitiveness of the Software

The intuitiveness of the software was defined in terms of the logical organization of the tasks performed by the software during a routine NGS bioinformatics workflow in the laboratory. For the technologists, intuitiveness was mainly focused on the ease of evaluation of the QC metrics of any given run. However, for the pathologists, this was focused on the annotation of the data variants and the ability to access this information easily. The ability to navigate various stages of the bioinformatics data in an intuitive manner was a crucial part of the software assessment. Although this is a subjective criterion, good consensus was obtained by our team on what was considered as an intuitive software in the final review. It is thus critical for software vendor companies to test their products with a spectrum of end users to establish the intuitiveness of a software solution.

Configurability

Bioinformatics workflows are different for each laboratory. In addition, LISs also vary widely between different laboratories. A key requirement is thus the ability to configure the NGS software solution to conform to the local LIS interface specifications. Building LIS interfaces is not trivial and often NGS bioinformatics vendors provide support for report generation in commonly used Microsoft Corp. (Redmond, WA) MS-WORD and/or Adobe PDF formats. The final clinical NGS reports are then separately uploaded into the LIS. Additional configurability requirements may exist for each laboratory. For example, in our laboratory, our original pipeline used two parallel but complementary NGS workflows [Torrent Variant Caller (Thermo Fisher Scientific, Waltham, MA) and NextGene software version 2.4.0.1 (Softgenetics, State College, PA)] to ensure redundancy of the variant calling process. This process provides enhanced confidence to the sign-out pathologist assessing the final variant calls generated for each patient. A key requirement for us was the ability to merge the Variant Call Format outputs from each of these software solutions (IGV and NextGene). The vendors were differentiated on their willingness to customize the solution to fit our local laboratory and IT needs. Smaller vendors were much more likely to work collaboratively to provide customized solutions compared with large corporate vendors. Corporate solutions were found to be often locked in with limited flexibility for end-user customization.

Quality of Variant Annotation

The quality of the database backing up the NGS variant call annotation was a key metric in our assessment. For all vendors assessed, there was an active, manual curation process as a part of their variant annotation database offering. Each vendor's database curation process was reviewed in detail. The total number of variant curation scientists involved in the effort, the update frequency of the databases, and the company's workflow of the variant update process were assessed. All the vendor solutions assessed include interfaces with the common, publicly available genomic annotation databases (eg, Single Nucleotide Polymorphism database, Catalog of Somatic Mutations in Cancer, and Genome Aggregation Database). However, each vendor also had separate licensing agreements with institutions and/or companies, to provide additional, specialized variant curation database access. In some instances, access to these specialized databases was a part of the standard fees of the software, whereas in others, additional fees were required to access these specialized databases. It is thus important for an

end user to have a clear idea of the quality (and costs) associated with the variant annotation data associated with each software solution.

A second important point relates to the origin of the database. During the review, database solutions were found to fall into two distinct categories: databases with historical origins in constitutional genetics data and oncology-focused databases. Understanding this distinction is of importance because the end goals of each category of databases are different. Although variant databases originating from an inherited genetics background are focused mostly on diagnostic paradigms (ie, Is this variant likely to cause a Mendelian inherited genetic disorder or not?), oncology-focused databases address questions of a different nature (eg, Is the variant and its associated variant allele frequency likely to be a driver mutation or a mere passenger variant? Does the detected variant affect the eligibility of patient into a clinical trial?). It is thus important to have a clear idea of the origin of the annotation data made available by the software vendor, depending on the application needs of the end user. For oncology-focused solutions, it is key to have an annotation database developed with a focus specifically on oncology. The quality of the database annotation driving the software solution is perhaps the single most important factor while assessing the utility of an NGS software solution.

User Support

Active vendor support at all phases of the bioinformatics workflow implementation and use is a critical consideration. Preinstallation vendor support for the local IT group is important to ensure a smooth implementation of the software solution. The postinstallation support is equally important to enable proper training and use of the bioinformatics solution across different skill levels within the MDx laboratory. The general observation was that smaller NGS bioinformatics vendors were much more responsive to user support concerns compared with larger corporate entities. The user documentation support provided by each vendor was also reviewed in detail. Detailed user documentation is highly useful to troubleshoot postimplementation NGS workflow issues. The pricing associated with the postimplementation software support was variable across vendors. Although some vendors provided it as a part of the initial pricing, others required additional fees for IT support (eg, hourly rates).

QC Data

QC metrics are an important part of the current College of American Pathologists' NGS guidelines.[15] All vendor solutions provided a varying degree of access to the NGS run QC data. Some software solution providers delivered a data and graphic-rich QC report, whereas other solutions were sparse and textual form of the NGS run QC data. The vendors were assessed mainly on their ability to provide a detailed review of the temporal QC metrics necessary to fulfill the mandated College of American Pathologists' regulatory requirements. All the vendors provided enough QC data (within each run and longitudinally across runs) to satisfy the overall QC assessment requirements.

Scalability for Future Use

As NGS testing volumes increase, it is important to assess if the chosen NGS solution will scale over time. Scalability may be viewed from two perspectives—complexity and volume. With increasing complexity of assays (eg, DNA sequencing, RNA sequencing, epigenetics, and microbiome analysis), the NGS software must be able to work with diverse NGS bioinformatics data sets. For clinical NGS, workflows currently of importance to a practicing molecular pathologist are perhaps the DNA- and RNA-sequencing NGS assays. These NGS assays are important for both constitutional and oncology applications alike. All the NGS bioinformatics solutions that were reviewed have the ability to tackle both of these assay types. In addition, many of these NGS bioinformatics solutions run on cloud web services providers, such as the Amazon Web Services. Thus, scalability of volumes is unlikely to be a major issue if one adopts a cloud-based NGS solution.

User Community

Last, a critically important consideration for us was the robustness of the user community supported by the NGS software vendor. The vendor efforts were assessed to generate an active community of users surrounding their software solutions. It is important for laboratories to obtain existing user input of a software solution before making the final selection decision. No single vendor who enabled an active user community that fully met our expectations was identified (eg, annual conferences or web-based interaction opportunities). It would be useful for commercial vendors to generate data and information sharing opportunities via mailing lists, user websites, and potential annual user meetings to share key ideas and experiences as a community. However, no vendors who were engaged in all these efforts were identified.

All the above nine criteria were formulated on the basis of our NGS workflow experience gained within our laboratory. However, these criteria may vary depending on each laboratory's individual needs and preferences. Thus, it is crucial for each laboratory to come up with its own internally consistent assessment guidelines (before starting the formal assessment process) to guide the NGS vendor selection process. With a set of well-formulated assessment criteria ahead of time, the final decision-making process is vastly simplified.

On completion of the vendor demonstrations, an MDx group meeting was organized with participation from all of the members of the laboratory at UNM/TRL. Key features of each software solution based on the above nine assessment criteria were presented to the MDx group by the project leader (R.R.G.) first. Then, each software's features were assessed in conjunction with a live online demonstration of the vendor's software solution. In addition, the original PDF presentation materials obtained from the vendors were reviewed as well. All attendees of the MDx

group meeting then graded each software solution on the basis of the nine key metrics described above on a scale of 1 to 5. Scores from all members of the laboratory were averaged and presented at a follow-up meeting for the final consensus decision (Table 3). Finally, the top two vendors after the review process were short listed for phase 4 of the assessment.

Phase 4: Pricing Negotiations and Contract Structuring

After the final group assessment, the next step (contract negotiations) was handed back to the TRL executive management team. Depending on the structural organization of the institution, one may not have an executive management team to implement the contract negotiations. This study describes the key lessons of contract negotiation gleaned during our NGS software evaluation process. The commercial space of NGS-based software solutions is relatively crowded, with an increasing number of NGS bioinformatics vendors available in the marketplace. Thus, it is often feasible to negotiate a favorable pricing for NGS software solutions. Companies are often willing to negotiate (extensively) to obtain an institutional toehold. Once a vendor is selected and the NGS bioinformatics workflow is well established in a laboratory, it is rather difficult to change vendors in the future. The NGS workflow familiarity established by the local users over time precludes easy migration to a different platform at a short notice. Thus, it is critical for institutions to negotiate extensively before signing the final contract. Significant variability in the pricing models between individual companies was also noted. Few companies are willing to work with a single, one-time pricing cost. Many of the NGS software companies use a subscription model, which uses a per-click costing structure. That is, for each case signed out, the customer is billed a fixed amount (Supplemental Table S2).

One also needs to be aware of the potential additional costs involved beyond the base subscription costs. These include fees, such as software implementation fees, hosting costs, hardware costs (for local installs), training fees, pipeline customization fees, and customer support fees.

Table 3 Final Cumulative Assessment Scores of All of the Individual Vendors

Vendor	Technologists ($n = 7$)	Supervisors ($n = 2$)	Pathologists ($n = 5$)
A	3.4	3.3	4.1
B	3.4	3.0	3.9
C	4.2	3.6	3.7
D	4.2	4.0	2.2
E	3.4	4.3	4.4

Data were obtained by group feedback from technicians, technical supervisors, and molecular pathologists in the molecular pathology team at TriCore Reference Laboratories. Vendors A and E were selected for final pricing negotiations.

Some companies also use a pricing model based on the number of genes used in the analysis (ie, different pricing for targeted panel gene testing versus whole-exome or whole-genome sequencing). The duration of the contract may be variable as well. Some companies are amenable to annual contracts, whereas others require a 3-year contract. A clear understanding of the cumulative total costs needs to be obtained before signing the final contract. Institutions can and should negotiate to obtain the best possible contract suited for their situation. A second lesson for medical directors is to be aware of a potential future locked-in IT software scenario. Overpriced, legacy IT software situations are often too common in health care. In NGS, one may be better off with having the flexibility to move over to a newer platform with advanced features because of the rapidly evolving nature of NGS bioinformatics workflows. To avoid downstream legacy lock-ins, institutions need to plan upfront to enable data workflow redundancies into their NGS bioinformatics pipelines. Companies may be unlikely to allow NGS data migration (eg, local medical director variant annotations accumulated over time) if the end user decides to discontinue the contract in future. Users must prepare in advance for these kinds of situations in a dynamically evolving field, such as NGS bioinformatics.

Phase 5: Software Implementation

After negotiating the contract, the final phase involves the implementation of the chosen bioinformatic solution within the laboratory. Similar to the initial phase of vendor assessment, it is necessary to have a key champion involved in the local installation and training process. One of the medical directors was nominated to be the key driver of the software implementation process at UNM/TRL. Laboratories may be faced with two different scenarios: implementation of a *de novo* bioinformatics solution or migration from a preexisting bioinformatics workflow IT solution. The experience with the migration of a preexisting process is described as follows.

Preimplementation Phase

First, the key personnel to facilitate the software transition process were identified. This included a medical director overseeing the transition, a technical supervisor of TRL MDx division, a senior technician in the laboratory, technical support personnel from the vendor, and the local IT experts at TRL. To enable the smooth implementation of the software solution, weekly conference calls were scheduled with the vendor to evaluate the progress and troubleshoot any implementation issues. It is key for the vendors to have a clear understanding of the local LIS systems to generate the appropriate interfaces for the NGS report transfer and communication. For TRL, two parallel, but complementary, bioinformatics pipelines were enabled for the raw NGS data analysis, as described earlier. This was done with the intention of ensuring variant call redundancy to avoid

missing key patient variant calls. A similar workflow method was customized with the vendor and implemented to ensure process continuity. Variant Call Format files generated from each of the parallel bioinformatics pipelines were merged using custom written python scripts and tested extensively by the vendor. Alternatively, individual laboratories may also decide to implement the end-to-end bioinformatics solution provided by the vendor. The raw FASTQ file generated by the NGS instrument is uploaded into the cloud and analyzed using the vendor-provided bioinformatics alignment and calling methods in its entirety. This may be an attractive option for laboratories just starting out in this space with minimal bioinformatics resources. Currently, the vendor solution is relied on mainly for variant annotation purposes.

Workforce Training and Transition

A training period was established to ensure familiarity with the vendor solution to all the molecular pathologists and the technologists within our division. Fifteen previously signed-out cases were uploaded to the vendor solution. Pathologists and fellows were randomly assigned cases and were signed out in the vendor test server. The final reports and analysis were compared with the original reports to identify any discrepancies of the outcomes. None was observed. An additional 50 cases were reviewed prospectively using the older in-house bioinformatics workflow and the new vendor solution to ensure enhanced familiarity of the new workflow process. The training workflow outcomes were documented as a part of the software validation process. Similar to our previous NGS workflow, the technicians in the laboratory were trained to perform the initial patient identification data entry procedures, followed by Variant Call Format and BAM file uploads into the cloud server after initial training with the new software. The run QC metrics are evaluated by the technicians first and documented as a part of the patient data review process.

Documentation and User Administration

A thorough documentation of the software procedures is important and mandatory for purposes of College of American Pathologists' inspections.[9,15] However, with cloud-based bioinformatics solutions (and even local software installations), it is key to have well-documented procedures in place to enable detailed administrative oversight of user access. Key leaders in the division (or local IT managers) were assigned the responsibility of serving as user administrators for the NGS bioinformatics solution. In most software solutions evaluated, it is possible to provide varying levels of user access to ensure patient confidentiality as well as ensure compliance with standard Health Insurance Portability and Accountability Act (HIPAA) protocols (discussed more in the next section). Varying user-level access in our chosen solution included those of i) accessioners, ii) MDx laboratory technicians, iii) case managers, iv) variant analysts, and v) sign-out pathologists. Laboratory members were assigned the appropriate level of user access by the user administrator with the appropriate documentation. The user administrator is also responsible for enforcing the software password guidelines. It is important to establish proper data hygiene procedures, such as periodic auditing of the access review (eg, quarterly) to ensure only authorized users exist in the system. Legacy users need to be deactivated from the system promptly after their departure from the institution. Audited data must be documented and shared routinely with the local IT compliance divisions.

Privacy Issues of Genomic Solutions in the Cloud

A key consideration in reviewing NGS software solutions is the issue of patient data privacy.[8,22,23] Health care IT systems increasingly face the key decision of locating patient data internally versus in the cloud.[8,22,24] NGS bioinformatic solutions generate a similar dilemma for clinical laboratories. Privacy requirements surrounding genomic testing are a complex and vast issue.[23] The interested reader is highly encouraged to combine this article with the detailed definitions and case studies of genomic privacy described in an excellent review.[3] Some of the basic definitions and the key lessons learned during our NGS software evaluation process are described herein. Protected health information is defined under the Privacy Rule, which is a part of the congressional mandate in the HIPAA of 1996.[3] Under HIPAA, electronic protected health information is defined as individually identifiable health information, generated and transmitted via electronic media. Although cloud-based software-as-a-service solutions are convenient for software vendors to enable regular software updates (and reduce the overall maintenance costs), user laboratories must be fully aware of the detailed regulatory requirements involving the use of cloud software. Laboratories using these solutions must negotiate with vendors upfront to ensure full knowledge of the timing of any future software updates. This is to enable the internal revalidation protocols as needed, depending on the nature of the software updates (major or minor). From a laboratory perspective, the biggest concerns of using cloud-based solutions are potential end-user security issues related to hacked patient data, loss of confidentiality, and potential HIPAA violations.[3,8] The goal of the HIPAA Privacy Rule is to prevent any potential abuses of health information by establishing national standards to protect an individual's medical record and other personal health information. The HIPAA Privacy Rule applies to health plans, health care clearinghouses, and those health care providers that conduct certain health care transactions electronically, which also includes laboratories. Thus, the security of cloud-based solutions is an important consideration for laboratories because of the potential loss of patient privacy in the event of a data breach. It is incumbent on testing laboratories to ensure the safety of the NGS data as a covered entity, as defined by HIPAA.[3,8,22] In the evaluative process of NGS software solutions, issues related to cloud security were robustly considered. It is important for

Table 4 Recommendations Provided by the Cloud Standards Customer Council to Ensure Data Security for Cloud Computing Applications

Ten steps for secure cloud computing
White paper summary recommendations by the Cloud Standards Customer Council version 2.0
1) Ensure effective governance, risk, and compliance processes exist
2) Audit operational and business processes
3) Manage people, roles, and identities
4) Ensure proper protection of data and information
5) Enforce privacy policies
6) Assess the security provisions for cloud applications
7) Ensure cloud networks and connections are secure
8) Evaluate security controls on physical infrastructure and facilities
9) Manage security terms in the cloud service agreement
10) Understand the security requirements of the exit process

The recommendations provided in the white paper are for general cloud computing applications. Yet, many of the Cloud Standards Customer Council recommendations are directly applicable for clinical genomics applications as well. For more in-depth details, please refer to *https://www.omg.org/cloud/deliverables/CSCC-Security-for-Cloud-Computing-10-Steps-to-Ensure-Success.pdf*.

laboratories to assess the degree of compliance of the NGS bioinformatics software with established HIPAA standards. Most of the software vendor solutions evaluated in this article provide cloud-based software-as-a-service NGS solutions only (only one vendor, Genomoncology, provided us with the option for a local install on-site). Thus, it was observed that the end users are unlikely to have many options for localized, on-site NGS bioinformatics vendor solutions. Thus, if one is forced to use a cloud-based software-as-a-service solution, what are the key points a laboratory needs to consider as a part of the software assessment?

The current genomic cloud computing literature was extensively reviewed.[8,22–28] Some key observations based on detailed literature surveys are as follows: For cloud-based software solutions, major vendors, such as the Google health portal, recommend the evaluation of vendor business associate agreements (BAAs). Most of the reputable corporate cloud service providers (eg, Google and Amazon) provide detailed BAAs online for HIPAA-associated protected health information data storage. It is incumbent on the laboratories considering cloud-based NGS solutions to discuss preexisting BAAs with the software vendors ahead of time and incorporate a summary of these discussions into the laboratory documentation to minimize future liability. It would be ideal to incorporate the software vendor's BAA as a part of the laboratory documentation. It is also a good idea to involve the local IT personnel and generate contract agreements that adhere to the requirements of the local institutional IT security policies from the get-go.[3,22] The vendor BAAs should ideally be reviewed annually by the local IT departments. Annual audits must be performed and documented in detail as a part of the routine workflow for cloud-based NGS solutions.[3,22] The encryption standards for data safety used by the NGS software solution providers must be evaluated and documented in the workflow protocols. The vendors must provide detailed information on the security parameters associated with their cloud-based bioinformatics pipelines, as requested by the end-user laboratory. An excellent overview of the 10 steps cloud users can take to ensure effective cloud security is provided in the white paper document published by a user advocacy group, the Cloud Standards Customer Council (Table 4).[29] However, it is important to remember that these 10 steps are only guidelines and, thus, not legally enforceable. One must still rely on the guidelines stated in the HIPAA law as a cornerstone to guide the necessary steps of local health care IT security in conjunction with the institutional IT and legal compliance teams. Issues related to long-term clinical NGS data storage are of critical importance to any laboratory performing diagnostic NGS assays. These issues are not discussed in this article, which mainly focuses on the administrative/evaluative process of selecting commercial NGS software. NGS data storage in the clinical context is an expansive topic that merits perhaps a separate future review. However, laboratories evaluating commercial NGS bioinformatics solutions must include long-term NGS data storage issues as a key part of their evaluative assessments.

Beyond routine regulatory measures, common sense local security measures must be a part of all routine laboratory workflows dealing with NGS data.[3,22] These include simple, but effective, steps, such as the following: i) data minimization (uploading only the raw NGS data); ii) deidentification (removing all of the patient-associated identifiers) and the use of nontraceable identifiers to prevent identification matching in the event of data loss; iii) strict access control and user authentication (allowing data access to only a limited number of users with strong passwords); and iv) routine data audits and reviews.[22] It is important to work closely with the local IT department to enable security compliance to ensure the security and safety of the patient NGS data. However, it is worthwhile to remember a zero-risk setup does not exist when it comes to computer security-related issues.

Discussion

NGS methods are increasingly the preferred platform of choice to analyze molecular biomarker changes in patient samples.[6] The main strength of NGS lies in its ability to multitask different forms of molecular assays (eg, DNA analysis, RNA analysis, epigenetics, and microbiome analysis) on the same hardware. With the ever-decreasing costs of the NGS hardware technology, high-throughput sequencing is now within the reach of many small- to mid-sized laboratories here in the United States and around

the globe.[30] In addition, for many oncology-related applications, determining key molecular-level changes is now an integral part of an oncology patient diagnostic and management algorithm.[31] Yet, many significant challenges remain for small- to mid-sized clinical laboratories desiring to implement NGS technologies. Bioinformatics remains a key challenge in the implementation of NGS assays.[1,14] The use of NGS assays is likely to expand at an increasing rate as our understanding of cancer pathophysiology improves at the molecular level. Not only is the volume of NGS assays likely to grow in the near future, but also the complexity associated with NGS informed clinical diagnostic and/or treatment paradigms.

Commercial software solutions in the marketplace aim to alleviate challenges associated with NGS bioinformatics data. This article describes a systematic evaluation of commercial NGS bioinformatics solutions available in the marketplace. Because of the lack of any prior literature focused on this issue, 20 months were spent on the NGS software assessment process and many important lessons were learned. Some of the key lessons gained along the way include the following: i) a clear assessment of the laboratory's NGS requirements is necessary upfront, ii) adequate personnel are critical to perform a thorough and exhaustive evaluation, iii) executive team inputs and feedback are critical to a successful evaluation process, and iv) full buy-in from the members of the laboratory is necessary to enable the evaluation process. The NGS evaluation project took much longer than what was originally anticipated. This was partly due to the lack of an understanding of the complexity involved in the evaluation process.

There are currently multiple NGS software options available to laboratories in the marketplace. With the rapidly increasing importance of this area, traditional LIS vendors have shown an increasing interest in the NGS bioinformatics space (eg, Sunquest). An add-on NGS bioinformatics solution, provided by our existing LIS vendor here at TRL, was also evaluated. Although a monolithic, integrated solution may be attractive to the executive and IT teams (because of the reduced IT vendor redundancy and the ability to negotiate costs effectively), our evaluation of one such add-on NGS software solution did not fully meet our NGS end-user requirements. Users need to be aware of any clunky, ill-formed NGS solutions that are merely convenient from the executive perspective. Such solutions have the potential to end up being burdensome with limited overall utility in the laboratory. Stand-alone companies that designed NGS software from scratch provided better NGS bioinformatics software solutions in our experience. During these evaluations, it was observed that smaller companies were more responsive overall to our specific NGS requirements compared with larger corporate solutions. However, it is important to remember that commercial NGS IT solutions are not a magic bullet to all the data management and clinical sign-out woes within a molecular pathology laboratory. An NGS bioinformatics IT solution does not eliminate the need for thoughtful interpretation and reporting of clinical NGS cases. The added expense of NGS bioinformatics associated with each clinical case is an issue executive management needs to consider in this age of declining reimbursements. Finally, the Pareto principle of software features is worth mentioning herein: in any NGS (or non-NGS) clinical informatics solution, only 20% of a software solution's features are likely to be used on a regular basis by the end users. Therefore, it is important for a laboratory to define the key essential features/requirements of an NGS solution upfront to evaluate the available NGS software solutions as pragmatically as possible.

IT software assessment requires a thorough knowledge of the detailed impact of a software workflow to identify the best possible solution. Yet, to our knowledge, there are no universally accepted templates on how to perform an exhaustive IT software evaluation. This is partly because of the unique nature of individual IT workflows associated with each institution. Often, IT solutions experienced the lack of a rigorous evaluative process from the perspective of the actual end user. Final decisions of software purchases are often made by management with minimal/no inputs from the end users. This is particularly true in health care IT, where corporate-driven software product development often ends up disappointing the true end user: the health care worker/laboratorian. We sought to avoid a similar fate by performing a rigorous assessment of the available IT solutions for NGS bioinformatics. As a result, this article reports the end-to-end experience of the NGS bioinformatics vendor assessment process and the key lessons gained during this process. We hope our experience will benefit institutions considering assessment of commercial NGS software solutions to optimize the existing or future NGS workflows within their own individual laboratories.

Acknowledgments

I thank Dr. Michael Crossey, CEO (TriCore Reference Laboratories), for allowing the generous republication of the TriCore Reference Laboratories request for information template; and Dr. Devon Chabot-Richards, Eric Carbonneau, Gwyneth Olson, Cathy Rawdon, and all the members of the molecular pathology team at TriCore Reference Laboratories for participating in this project.

Author Contribution

R.R.G. conceived and executed the project, analyzed data, and wrote the manuscript.

Supplemental Data

Supplemental material for this article can be found at *http://doi.org/10.1016/j.jmoldx.2019.09.007*.

References

1. Gullapalli RR, Desai KV, Santana-Santos L, Kant JA, Becich MJ: Next generation sequencing in clinical medicine: challenges and lessons for pathology and biomedical informatics. J Pathol Inform 2012, 3:40
2. Roy S, LaFramboise WA, Nikiforov YE, Nikiforova MN, Routbort MJ, Pfeifer J, Nagarajan R, Carter AB, Pantanowitz L: Next-generation sequencing informatics: challenges and strategies for implementation in a clinical environment. Arch Pathol Lab Med 2016, 140:958–975
3. Carter AB: Considerations for genomic data privacy and security when working in the cloud. J Mol Diagn 2019, 21:542–552
4. Luthra R, Chen H, Roy-Chowdhuri S, Singh RR: Next-generation sequencing in clinical molecular diagnostics of cancer: advantages and challenges. Cancers (Basel) 2015, 7:2023–2036
5. Cheng DT, Mitchell TN, Zehir A, Shah RH, Benayed R, Syed A, Chandramohan R, Liu ZY, Won HH, Scott SN, Brannon AR, O'Reilly C, Sadowska J, Casanova J, Yannes A, Hechtman JF, Yao J, Song W, Ross DS, Oultache A, Dogan S, Borsu L, Hameed M, Nafa K, Arcila ME, Ladanyi M, Berger MF: Memorial Sloan Kettering-Integrated Mutation Profiling of Actionable Cancer Targets (MSK-IMPACT): a hybridization capture-based next-generation sequencing clinical assay for solid tumor molecular oncology. J Mol Diagn 2015, 17:251–264
6. Rajeevan MS, Li T, Unger ER: Precision medicine requires precision laboratories. J Mol Diagn 2017, 19:226–229
7. Campbell WS, Carter AB, Cushman-Vokoun AM, Greiner TC, Dash RC, Routbort M, de Baca ME, Campbell JR: A model information management plan for molecular pathology sequence data using standards: from sequencer to electronic health record. J Mol Diagn 2019, 21:408–417
8. Dove ES, Joly Y, Tasse AM; Public Population Project in Genomics and Society (P3G) International Steering Committee; International Cancer Genome Consortium (ICGC) Ethics and Policy Committee, Knoppers BM: Genomic cloud computing: legal and ethical points to consider. Eur J Hum Genet 2015, 23:1271–1278
9. Roy S, Coldren C, Karunamurthy A, Kip NS, Klee EW, Lincoln SE, Leon A, Pullambhatla M, Temple-Smolkin RL, Voelkerding KV, Wang C, Carter AB: Standards and guidelines for validating next-generation sequencing bioinformatics pipelines: a joint recommendation of the Association for Molecular Pathology and the College of American Pathologists. J Mol Diagn 2018, 20:4–27
10. Jennings LJ, Arcila ME, Corless C, Kamel-Reid S, Lubin IM, Pfeifer J, Temple-Smolkin RL, Voelkerding KV, Nikiforova MN: Guidelines for validation of next-generation sequencing-based oncology panels: a joint consensus recommendation of the Association for Molecular Pathology and College of American Pathologists. J Mol Diagn 2017, 19:341–365
11. Li MM, Datto M, Duncavage EJ, Kulkarni S, Lindeman NI, Roy S, Tsimberidou AM, Vnencak-Jones CL, Wolff DJ, Younes A, Nikiforova MN: Standards and guidelines for the interpretation and reporting of sequence variants in cancer: a joint consensus recommendation of the Association for Molecular Pathology, American Society of Clinical Oncology, and College of American Pathologists. J Mol Diagn 2017, 19:4–23
12. Lih CJ, Harrington RD, Sims DJ, Harper KN, Bouk CH, Datta V, Yau J, Singh RR, Routbort MJ, Luthra R, Patel KP, Mantha GS, Krishnamurthy S, Ronski K, Walther Z, Finberg KE, Canosa S, Robinson H, Raymond A, Le LP, McShane LM, Polley EC, Conley BA, Doroshow JH, Iafrate AJ, Sklar JL, Hamilton SR, Williams PM: Analytical validation of the next-generation sequencing assay for a nationwide signal-finding clinical trial: molecular Analysis for Therapy Choice Clinical Trial. J Mol Diagn 2017, 19:313–327
13. Kang W, Kadri S, Puranik R, Wurst MN, Patil SA, Mujacic I, Benhamed S, Niu N, Zhen CJ, Ameti B, Long BC, Galbo F, Montes D, Iracheta C, Gamboa VL, Lopez D, Yourshaw M, Lawrence CA, Aisner DL, Fitzpatrick C, McNerney ME, Wang YL, Andrade J, Volchenboum SL, Furtado LV, Ritterhouse LL, Segal JP: System for informatics in the molecular pathology laboratory: an open-source end-to-end solution for next-generation sequencing clinical data management. J Mol Diagn 2018, 20:522–532
14. Gullapalli RR, Lyons-Weiler M, Petrosko P, Dhir R, Becich MJ, LaFramboise WA: Clinical integration of next-generation sequencing technology. Clin Lab Med 2012, 32:585–599
15. Molecular Pathology Checklist. Chicago, IL: College of American Pathologists (CAP), 2017. pp. 35–51
16. Gargis AS, Kalman L, Berry MW, Bick DP, Dimmock DP, Hambuch T, et al: Assuring the quality of next-generation sequencing in clinical laboratory practice. Nat Biotechnol 2012, 30:1033–1036
17. Gargis AS, Kalman L, Bick DP, da Silva C, Dimmock DP, Funke BH, et al: Good laboratory practice for clinical next-generation sequencing informatics pipelines. Nat Biotechnol 2015, 33:689–693
18. Lubin IM, Aziz N, Babb LJ, Ballinger D, Bisht H, Church DM, Cordes S, Eilbeck K, Hyland F, Kalman L, Landrum M, Lockhart ER, Maglott D, Marth G, Pfeifer JD, Rehm HL, Roy S, Tezak Z, Truty R, Ullman-Cullere M, Voelkerding KV, Worthey EA, Zaranek AW, Zook JM: Principles and recommendations for standardizing the use of the next-generation sequencing variant file in clinical settings. J Mol Diagn 2017, 19:417–426
19. Friedman CP, Wyatt J: Evaluation Methods in Biomedical Informatics. New York, NY: Springer, 2005
20. Robinson JT, Thorvaldsdóttir H, Wenger AM, Zehir A, Mesirov JP: Variant review with the integrative genomics viewer. Cancer Res 2017, 77:e31–e34
21. Thorvaldsdottir H, Robinson JT, Mesirov JP: Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform 2013, 14:178–192
22. Cucoranu IC, Parwani AV, West AJ, Romero-Lauro G, Nauman K, Carter AB, Balis UJ, Tuthill MJ, Pantanowitz L: Privacy and security of patient data in the pathology laboratory. J Pathol Inform 2013, 4:4
23. Erlich Y, Narayanan A: Routes for breaching and protecting genetic privacy. Nat Rev Genet 2014, 15:409–421
24. Fernandez-Aleman JL, Senor IC, Lozoya PA, Toval A: Security and privacy in electronic health records: a systematic literature review. J Biomed Inform 2013, 46:541–562
25. Datta S, Bettinger K, Snyder M: Secure cloud computing for genomic data. Nat Biotechnol 2016, 34:588–591
26. Kaye J: The tension between data sharing and the protection of privacy in genomics research. Annu Rev Genomics Hum Genet 2012, 13: 415–431
27. McEwen JE, Boyer JT, Sun KY: Evolving approaches to the ethical management of genomic data. Trends Genet 2013, 29:375–382
28. Schlosberg A: Data security in genomics: a review of Australian privacy requirements and their relation to cryptography in data storage. J Pathol Inform 2016, 7:6
29. Cloud Standards Customer Council: Security for Cloud Computing Ten Steps to Ensure Success v.3.0; 2017. Available at https://www.omg.org/cloud/deliverables/CSCC-Security-for-Cloud-Computing-10-Steps-to-Ensure-Success.pdf (accessed December 2019)
30. Schrijver I, Farkas DH, Gibson JS, Lyon E; AMP Executive Committee: The evolving role of the laboratory professional in the age of genome sequencing: a vision of the Association for Molecular Pathology. J Mol Diagn 2015, 17:335–338
31. Sepulveda AR, Hamilton SR, Allegra CJ, Grody W, Cushman-Vokoun AM, Funkhouser WK, Kopetz SE, Lieu C, Lindor NM, Minsky BD, Monzon FA, Sargent DJ, Singh VM, Willis J, Clark J, Colasacco C, Rumble RB, Temple-Smolkin R, Ventura CB, Nowak JA: Molecular biomarkers for the evaluation of colorectal cancer: guideline from the American Society for Clinical Pathology, College of American Pathologists, Association for Molecular Pathology, and American Society of Clinical Oncology. J Mol Diagn 2017, 19:187–225