# Comparison of Annotation Services for Next-Generation Sequencing in a Large-Scale Precision Oncology Program

Evangelia Katsoulakis, MD[1]; Jill E. Duffy, MBA, PA-C[2]; Bradley Hintze, PhD[2,3]; Neil L. Spector, MD[2,3,4]; and Michael J. Kelley, MD[2,3,4]

**PURPOSE** Next-generation sequencing (NGS) multigene panel testing has become widespread, including the Veterans Affairs (VA), through the VA National Precision Oncology Program (NPOP). The interpretation of genomic alterations remains a bottleneck for realizing precision medicine. We sought to examine the concordance for pathogenicity determination and clinical actionability of annotation services in NPOP.

**METHODS** Unique gene variants were generated from NGS gene panel results using two sequencing services. For each unique gene variant, annotations were provided through N-of-One (NoO), IBM Watson for Genomics (WfG), and OncoKB. Annotations for pathogenicity (all three sources) and actionability (WfG and OncoKB) were examined for concordance. Cohen's kappa statistic was calculated to measure agreement between annotation services.

**RESULTS** Among 1,227 NGS results obtained between 2015 and 2017, 1,388 unique variants were identified in 117 genes. The genes with the largest number of variants included *TP53* (270), *STK11* (92), and *CDKN2A* (81). The most common cancer type was lung adenocarcinoma (440), followed by colon adenocarcinoma (113). For pathogenic and likely pathogenic variants, there was 30% agreement between WfG and NoO (kappa, –0.26), 76% agreement between WfG and OncoKB (kappa, 0.22), and 42% agreement between NoO and OncoKB (kappa, –0.07). For level 1 drug actionability of gene variant–diagnosis combinations, there was moderate agreement between WfG and OncoKB (96.9%; kappa, 0.44), with 27 combinations identified as level 1 by both services, 58 by WfG alone, and 6 variants by OncoKB alone.

**CONCLUSION** There is substantial variability in pathogenicity assessment of NGS variants in solid tumors by annotation services. In addition, there was only moderate agreement in level 1 therapeutic actionability recommendations between WfG and OncoKB. Improvement in the precision of NGS multigene panel annotation is needed.

*JCO Precis Oncol 4:212-221. © 2020 by American Society of Clinical Oncology*

## INTRODUCTION

The framework to implement genomics-driven therapeutic oncology has been rapidly established in leading cancer centers worldwide. An exponential growth in characterized cancer genomes has ensued.[1] The results of next-generation sequencing (NGS) assays are intricate and provide a multitude of somatic mutations in hundreds of genes. Mutation rates vary significantly between different cancers.[2] Many somatic genetic variants are characterized as variants of unknown significance, and the implication of multiple mutations within a single tumor is poorly defined.[3] Moreover, tumors possess intratumoral genetic heterogeneity, adding another layer of complexity to NGS results.[4]

At many institutions, including the Veterans Affairs (VA) healthcare system, molecular tumor boards (MTBs)

have been instituted to interpret the results of NGS and develop treatment recommendations. In some instances the whole genome is sequenced, and in other cases selected subsets of genes are sequenced with smaller panels. Identified alterations that are of particular clinical interest include driver mutations or druggable mutations.[1] For use in clinical decision making, identified gene variants must be annotated or interpreted in terms of the likelihood of their tumorigenicity and drug actionability.[5] There are many resources to assist with data curation, including OncoKB (Memorial Sloan Kettering Cancer Center, New York, NY),[6] My Cancer Genome (Vanderbilt-Ingram Cancer Center, Nashville, TN),[7] Precision Medicine Knowledge Base (Weill Cornell Medicine Englander Institute for Precision Medicine, New York, NY),[8] Personalized Cancer Therapy (MD Anderson Cancer

**CONTEXT**

**Key Objective**

Algorithms to interpret the pathogenicity and clinical actionability of next-generation sequencing (NGS)–detected sequence variants have been implemented; however, little is known about their relative performance in clinical practice. The clinical interpretation of genomic alterations remains a major bottleneck for realizing precision medicine. We sought to examine the concordance for pathogenicity and clinical actionability of three annotation services available to the VA Precision Oncology Program (N-of-One, OncoKB, and Watson for Genomics).

**Knowledge Generated**

We demonstrate that NGS annotation services have wide-ranging agreements in pathogenicity (30%-76%). Moreover, there was moderate agreement (96.9%) in level 1 drug actionability.

**Relevance**

We anticipate these findings will encourage improvement in the precision of NGS multigene panel annotation. We provide detailed information regarding NGS panels, gene variant pathogenicity, annotation ontology, and level 1 drug actionability by annotation service. This study has significant implications for precision oncology clinical trials and molecular tumor boards.

Center, Houston, TX),[9] CIViC (Washington University in St Louis School of Medicine, St Louis, MO),[10] Jackson Laboratory Clinical Knowledgebase (Jackson Laboratory, Bar Harbor, ME),[11] Cancer Genome Interpreter (Barcelona, Spain),[5] Cancer Driver Log (omicX, Le-Petit-Quevilly, France),[12] and N-of-One (NoO; Concord, MA).[13] The identification of mutations that are pathogenic and actionable is generally performed by multidisciplinary teams who reference genomic databases, published literature, and clinical trials. The development of cognitive computing, such as Watson for Genomics (WfG), has also garnered interest in precision oncology. WfG[14] is a cloud-based service that uses computerized models to simulate human thought to analyze large volumes of genome data and generate evidence-based guidelines.

The ultimate objective of gene sequencing is to improve patient outcomes by matching patients to specific treatments that target mutations driving the growth of individual tumors. The landscape of genetic tests has evolved from single mutation to small hotspot mutation panels, large gene panels, and whole exome and whole genome platforms. The interpretation of the clinical significance of these genomic mutations has become a formidable task because of the number of genes tested and limited understanding of normal genetic variation as well as pathogenic gene-to-gene interaction. Although professional societies have recently published consensus guidelines[15] for the use and interpretation of NGS, they have not been extensively used in practice. Moreover, traditional clinical trial design approaches of molecularly unselected tumor types may no longer be appropriate because of the molecular heterogeneity of tumors from each primary site (eg, breast). This has led to the development of basket and umbrella trial designs as well as precision genomic oncology trials at the point of care. An urgency exists to examine the various annotation services and clinical support tools available to ensure quality of NGS in oncology.

The VA National Precision Oncology Program (NPOP)[16] initially used two commercial vendors for NGS testing, Personal Genome Diagnostics (PGDx, Baltimore, MD) and Personalis (Menlo Park, CA), that used the NoO commercial annotation service. In addition, WfG was available to the VA through a gift from IBM. Until NPOP accrues substantial outcomes data, treatment recommendations are largely based on biomarker panels, annotations, and relevant patient characteristics.[17] NPOP also references open data sources, including OncoKB (an evidence source, not an annotation provider), when interpreting pathogenicity and actionability of gene variants. For some variants and cases, annotation is assisted by a molecular pathologist. The primary objective of this study was to assess the concordance for pathogenicity determination and clinical actionability for the annotation services available to NPOP.

## METHODS

Sequencing data from all patients who underwent NGS testing under NPOP since its genesis in 2015 through 2017 were analyzed. The work described here was conducted to assess the quality of annotation services available to the NPOP, a clinical operational program that is not research and does not require institutional review board review; permission to publish was obtained from the appropriate VA authority. Details of the NPOP have been published previously.[16] The goal is to use NGS testing to facilitate patient access to US Food and Drug Administration (FDA)–approved targeted therapies and immune checkpoint inhibitors and increase participation in clinical trials (Data Supplement). Clinical trials have been the recommended option in > 50% of cases for which there was no FDA-approved therapy. NGS results were generated through two contracted vendors: PGDx[18] (CancerSELECT 125, Plasma-SELECT 64, CancerSELECT 88, and CancerSELECT 203) and Personalis (ACE Cancer Plus 181)[19] (Data Supplement).

We matched gene variants for the two contracted vendors, PGDx and Personalis, to compare vendor recommendations as well as commercial annotation services for the same unique gene variant. NGS-detected variants were annotated by the sequencing vendor using the commercial annotation service NoO. We examined NoO as implemented by the vendors. NPOP staff re-annotated DNA sequence results using WfG and OncoKB. IBM donated use of WfG to the VA; OncoKB is created and maintained by Memorial Sloan Kettering Cancer Center (MSKCC) and available online.[6]

After completion of sequencing, results were uploaded by an NPOP data scientist to WfG (including tumor type, a list of sequence variants as a variant call file, presence of gene fusions, and gene copy number variation). Results were routinely uploaded into WfG as part of routine test interpretation workflow. Batch analysis was performed approximately weekly using a custom-built informatics workflow and a packaging tool provided by IBM. WfG assigned a pathogenicity label to each variant or fusion. Drug matching for pathogenic variants is also part of the WfG analysis pipeline. Variants that are pathogenic or likely pathogenic are matched with FDA-approved drugs and actively recruiting clinical trials using levels of evidence (level 1, 2A, 2B, 3A, 3B, 4, or R1; Tables 1 and 2).[3] For actionability, WfG uses a subset of National Cancer Institute (NCI) thesaurus terms for diagnosis coding.[20] WfG was performed on all samples as part of NPOP workflow, and OncoKB analysis was performed on an ad hoc basis in response to NPOP consults and/or MTB cases.

Using the same unique gene variants annotated by NoO and WfG, we queried all variants using the OncoKB curated database for the current study (Data Supplement). The public database[21] includes information on the clinical actionability for each somatic gene variant organized by indication and four-tier levels of evidence (Tables 1 and 2) which are ultimately incorporated into cBioPortal[22] for Cancer Genomics to aid physicians and cancer researchers.[6] Many genes involved in tumorigenesis are not targetable with currently available drugs. More than 90% of alterations in OncoKB have biologic effects and are classified as oncogenic but not actionable. For actionability, tumor ontology is considered. OncoKB contains 43 tumor types with biomarker-drug associations and uses an open-source ontology, OncoTree,[23] which was developed at MSKCC (699 tumor types). The Clinical Genomics Annotation Committee (CGAC) reviews the OncoKB alteration across 22 disease management teams. Curation reviews occur every 3 months, and CGAC recommendations and feedback are updated in real time.[6]

For this study, levels of evidence 2A and 2B were consolidated to level 2 and levels 3A and 3B to level 3 for both WfG and OncoKB. For actionability annotation comparisons, we mapped ontology from WfG to OncoKB to ensure that gene variants were properly annotated with their respective tumor type. In addition, for actionability annotation comparisons, we excluded microsatellite instability (MSI), as the MSI status was not entered into WfG and we do not have MSI information from WfG.

## RESULTS

Among 1,227 NGS results, 1,388 unique variants were observed in 117 genes. The entire set of unique gene variants is shown in the Data Supplement. The genes with the largest number of variants included were: *TP53* (270), *STK11* (92), *CDKN2A* (81), *ATM* (67), *PTEN* (52), *NF1* (46), and *BRCA2* (45). The most common cancer was lung adenocarcinoma in 35.86% (440). Other cancer types included colon adenocarcinoma 9.21% (113) and lung squamous cell carcinoma 9.05% (111). The complete list of cancer types is shown in Figure 1.

Of the 1,388 unique gene variants, 1,082 were identified only by PGDx (a larger proportion of the samples were sequenced by PGDx), and 480 were identified only by Personalis, whereas 174 gene variants were generated by both vendors in different specimens. NoO classification should be similar for both vendors. The unique gene variants generated by each vendor are shown in the Data Supplement. The genes included in the gene panels are listed in the Data Supplement. We also list the well-characterized genes of each panel as well as the genes for which copy analysis is performed. For panel distribution by NGS samples, see the Data Supplement. All unique gene variants were annotated by NoO, WfG, and OncoKB.

**TABLE 1.** Levels of Evidence Used by Watson for Genomics

| Level | Watson for Genomics |
|---|---|
| Level 1 | FDA-approved drug in this cancer type and biomarker |
| Level 2A | Standard-of-care biomarker predictive of response to an FDA-approved drug in this indication |
| Level 2B | Standard-of-care biomarker predictive of response to an FDA-approved drug in a different indication |
| Level 3A | Compelling clinical evidence supports the biomarker as being predictive of response to a drug in this indication |
| Level 3B | Compelling clinical evidence supports the biomarker as being predictive of response to a drug in a different indication |
| Level 4 | Compelling biologic evidence supports the biomarker as being predictive of response to a drug |
| R1 | Standard-of-care biomarker predictive of resistance to an FDA-approved drug in this indication |

Abbreviation: FDA, US Food and Drug Administration.

**TABLE 2.** Levels of Evidence Used by OncoKB

| Level | OncoKB |
|-------|--------|
| Level 1 | FDA-recognized biomarker predictive of response to an FDA-approved drug in this indication |
| Level 2A | Standard care biomarker predictive of response to an FDA-approved drug in this indication |
| Level 2B | Standard care biomarker predictive of response to an FDA-approved drug in another indication but not standard care for this indication |
| Level 3A | Compelling clinical evidence supports the biomarker as being predictive of response to a drug in this indication, but neither biomarker nor drug is standard care |
| Level 3B | Compelling clinical evidence supports the biomarker as being predictive of response to a drug in another indication, but neither biomarker nor drug is standard of care |
| Level 4 | Compelling biologic evidence supports the biomarker as being predictive of response to a drug, but neither biomarker nor drug is standard care |
| Level R1 | Standard care biomarker predictive of resistance to an FDA-approved drug in this indication |

Abbreviation: FDA, US Food and Drug Administration.

Out of the 1,388 unique gene variants, 337 (24.2%) were variants of unknown significance by OncoKB. Out of the 1,388 unique gene variants, 270 (19.4%) were variants of unknown significance by WfG.

For pathogenicity annotation, (ie, pathogenic and likely pathogenic variants *v* all other variants), there was fair agreement between WfG and OncoKB (76%; kappa, 0.22) and no agreement between WfG and NoO (30%; kappa, –0.26) as well as NoO and OncoKB (42%; kappa, –0.07; Table 3; Fig 2; Appendix Fig A1).

There were 91 unique gene variant–diagnosis combinations identified as having level 1 drug actionability recommendations identified by WfG, OncoKB, or both (not available from NoO). As part of actionability annotation, we mapped diagnosis ontology from WfG to OncoKB for each observed diagnosis (Data Supplement).

There was moderate agreement between WfG and OncoKB (96.9%; kappa, 0.445), with 58 variants identified only by WfG as level 1 and 6 variants identified only by OncoKB as



**FIG 1.** Next-generation sequencing (NGS) tumor types. Among 1,227 NGS results, 1,388 unique gene variants were identified. The NGS solid tumor type distribution is indicated. H&N, head and neck; NSCLC, non–small-cell lung cancer.

level 1 (Table 4). The complete set of level 1 drug recommendations for WfG and OncoKB are shown in the Data Supplement. When both annotation services had level 1 drug actionability, the recommended drugs were overwhelmingly identical. An example of a unique gene variant with drug actionability concordance is the exon 19 deletion in lung adenocarcinoma, *EGFR-L747_S752del*, with level 1 evidence for use of erlotinib, afatinib, or gefitinib using both WfG and OncoKB annotation services. Response rates of tumors with *EGFR* exon 19 deletions at *L747* have been reported as high as 83.3%.[24]

Among the 33 unique gene variant diagnoses identified as level 1 by OncoKB, WfG classification was level 1 in 27; for the 6 cases with discordant classification in WfG, 5 cases were lung adenocarcinoma *EGFR* mutations and a melanoma *BRAF-V600R* mutation. For the *BRAF-V600R* mutation in melanoma, WfG had no level 1 recommendations, whereas OncoKB had 5 drug or drug combinations as level 1. *BRAF-V600R* mutations constitute approximately 3%-7% of all *BRAF* mutations and were not included in the original *BRAF/MEK* inhibitors clinical trials.[25] Although *V600R* is not listed on the *BRAF/MEK* FDA label, National Comprehensive Cancer Network panel guidelines from as early as 2016 consider single-agent *BRAF* inhibitor monotherapy and *BRAF/MEK* inhibitor combination therapy an appropriate treatment of all activating *BRAF* mutations, including *V600R*, *V600D*, and others.[26] For *EGFR* mutations in lung adenocarcinoma (*EGFR-A750P*, *EGFR-E746_T751delinsl*, *EFGR-G719C*, *EGFR-L861Q*, and *EGFR-S7681*), WfG had no recommendations in 3 variants and had level 2A for *EGFR-G719C* and level 3B for both *EGFR-A740P* and *EGFR L861Q* for drugs afatinib, erlotinib, and gefitinib. For *EGFR-L861Q*, an uncommon *EGFR* mutation, the afatinib FDA label expanded to include *EGFR-L861Q* in January of 2018, which was after our annotation analysis. These approvals were based on findings from the phase 2 LUX-Lung 2 trial as well as the phase 3 LUX-Lung3 and LUX-Lung 6 trials that showed an objective response rate of 66% (95% CI, 47% to 81%).[27]
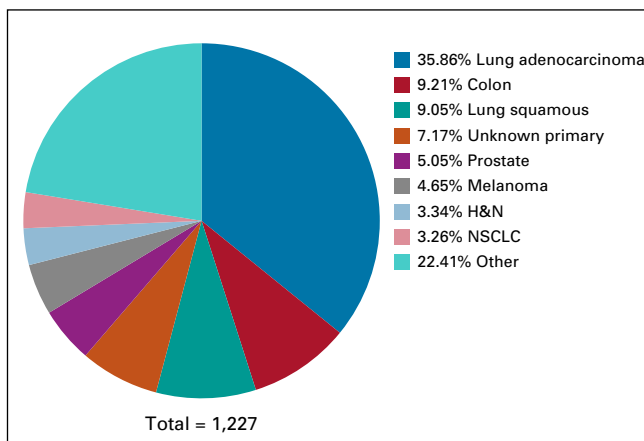
**TABLE 3.** Unique Gene Variants and Their Annotation as Pathogenic or Likely Pathogenic Using Three Annotation Services: N-of-One, WfG, and OncoKB

|  | Pathogenic | Not Pathogenic | Total | Cohen's κ |
|---|---|---|---|---|
| N-of-One |  |  |  |  |
| WfG |  |  |  | −0.26 |
| Pathogenic | 420 | 790 | 420 |  |
| Not pathogenic | 178 | 0 | 968 |  |
| Total | 598 | 790 | 1,388 |  |
| OncoKB |  |  |  |  |
| WfG |  |  |  | 0.22 |
| Pathogenic | 961 | 249 (of these, 246 VUS on OncoKB) | 1,210 |  |
| Not pathogenic | 85 | 93 (of these, 91 VUS on OncoKB) | 178 |  |
| Total | 1,046 | 342 | 1,388 |  |
| N-of-One |  |  |  |  |
| OncoKB |  |  |  | −0.07 |
| Pathogenic | 422 | 176 (of these, 174 VUS on OncoKB) | 598 |  |
| Not pathogenic | 624 | 166 (of these, 163 VUS on OncoKB) | 790 |  |
| Total | 1,046 | 342 | 1,388 |  |

NOTE. There was 30% agreement between WfG and N-of-One, 76% agreement between WfG and OncoKB, and 42% agreement between OncoKB and N-of-One.

Abbreviations: WfG, Watson for Genomics; VUS, variant of unknown significance.

Among the 85 unique gene variant diagnoses identified as level 1 by WfG, OncoKB classification was level 1 in 27, and 58 had discordant classification. Most of the gene variant diagnoses (40/58 [68.9%]) with discordant recommendations were mismatch repair (MMR) genes, such as *MLH1, MSH2, MSH6,* and *PMS2.* The remaining gene variant diagnoses (18/58; 31.0%) with discordant recommendations included genes *CDK4, EGFR, FGFR1, FLT4, KIT, PDGFRA, PIK3CA, POLE, PTEN, TSC1, TSC2, VEGFA,* and *VHL.* For a lung adenocarcinoma case with *MLH1-A42,* which is part of DNA MMR genes and known to have increased MSI, WfG provide a level 1 recommendation for atezolizumab, nivolumab, and pembrolizumab. Although MMR status predicts the clinical benefit of immune checkpoint blockade in multiple tumor types,[28,29] none of these drugs are FDA approved for *MLH1* mutants. Similarly, for other MMR gene variants, WfG provided level 1 recommendations in the absence of an FDA on-label indication.

## DISCUSSION

Precision medicine has promoted the development of biomarker-driven treatment strategies. There has been a surge in the genetic testing market, with a 10% annual increase in new genetic tests and a 20% annual increase in gene-based diagnostic tests.[15,30] NGS generates massive amounts of data, and different biologic questions require the development of specific bioinformatics pipelines, which are frequently platform specific.[30,31] The processing of raw sequence data has a profound effect on patient care and

outcomes, and NGS test validation is necessary.[32] Until recently, there were no established uniform technical standards for reporting tumor-derived NGS gene panel sequencing results.[33] In this analysis, we examine a large number of clinically observed unique gene variants and compare annotations for pathogenicity and actionability through two commercial services, NoO and WfG, as well as an open database, OncoKB.

In efforts to establish guidelines for NGS gene panel interpretation of somatic variants, the Association of Molecular Pathology (AMP) and College of American Pathologists (CAP) issued a joint consensus of classification (Table 5).[32,33] The four-tier system has different nomenclature and classification as compared with WfG and OncoKB (Tables 1 and 2). The WfG and OncoKB levels of evidence are similar and compatible, but not identical, to the AMP/CAP/ASCO evidence levels. The OncoKB evidence levels were mapped to the evidence levels for AMP/CAP/ASCO tier I and II variants (Table 5).[34] We used the WfG and OncoKB classification systems to compare their levels to each other, and we found that variants classified as pathogenic had a wide range of concordance, from 30% to 76% (Fig 2; Table 3; Appendix Fig A1). In routine clinical practice, only gene variants that are pathogenic or likely pathogenic are considered for actionability, so misclassification in pathogenicity could readily affect provider prescribing. For drug actionability, there was moderate-strength agreement 96.9% (kappa, 0.445) for levels 1. In practice, levels of evidence 1 and 2A both frequently result in use of a targeted drug, so
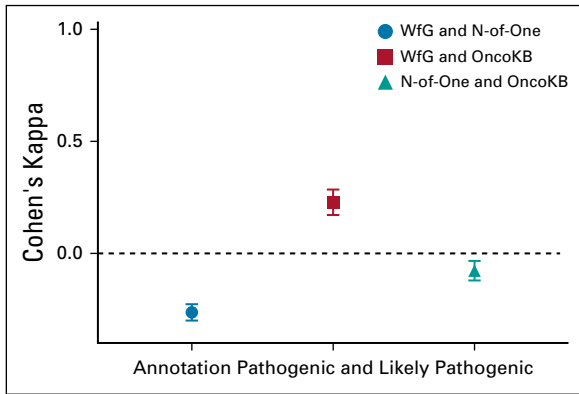
**FIG 2.** Agreement of pathogenicity annotation: Watson for Genomics (WfG) and N-of-One good strength agreement, WfG and OncoKB fair agreement, and N-of-One and OncoKB no agreement. Cohen's kappa coefficient of 1 indicates perfect agreement, and 0 indicates that any agreement is due to chance. The number of samples in each group is shown in Table 3.

concordance in level 2A may be similar to level 1, and their combination may improve concordance among annotation services for actionability. However, we expect greater discordance among levels of evidence 2B to 4.

Although no direct comparisons between NoO, WfG, and OncoKB have been previously performed, WfG has been compared with MTBs.[3] The utility of WfG was compared with MTBs by examining 1,018 cases analyzed by the University of North Carolina's Human-MTB. WfG identified an additional 8 actionable genes (7 of which passed actionability criteria by Human-MTB). Mutations in these 8 genes were identified in 231 and 96 patients out of 703 and 315 patients with already identified actionable and no actionable mutations, respectively.[3] Of the 7 actionable genes identified by WfG, 3 had no clinical trial available, and 4 were made potentially eligible for a biomarker-selected clinical trial. The reason for the added genes by WfG was believed to be the opening of several clinical trials

**TABLE 4.** Level 1 Drug Actionability Comparing IBM WfG and Open Academic Somatic Variant Database OncoKB

| | OncoKB | | | |
| --- | --- | --- | --- | --- |
| | Level 1 Drug Actionability Evidence | Other Levels | Total | Cohen's κ |
| WfG | | | | 0.445 |
| Level 1 drug actionability evidence | 27 | 58 | 85 | |
| Other levels | 6 | 2,007 | 2,013 | |
| Total | 33 | 2,065 | 2,098 | |

NOTE. Level 1 drug actionability was not available from N-of-One. Between WfG and OncoKB, there was moderate agreement: 96.9%; κ, 0.445; SE, 0.0572; 95% CI, 0.333 to 0.5572.

Abbreviation: WfG, Watson for Genomics.

within weeks of the WfG analysis. Most genes identified and reclassified as actionable, however, have yet to demonstrate their utility as predictive biomarkers of response to the recommended therapy. These actionable mutations identified by WfG were found retrospectively in 323 patients, of whom only 47 (4.6%) had active disease requiring further therapy. Moreover, none of the patients had treatment altered based on WfG's recommendation. Cognitive computing may supplement MTBs, and multiple studies have shown concordance of > 90%.[35,36] WfG may provide additional decision support in the current era of rapid generation of information from clinical trials.

Currently, existing evidence does not support population-based universal tumor sequencing.[37] The SHIVA trial (ClinicalTrials.gov identifier: NCT01771458) examined patients with any metastatic solid tumor refractory to standard treatment and randomly assigned 195 to receive treatment on the basis of molecular profile (n = 99) versus standard treatment (n = 95). Median progression-free survival was 2.3 months (95% CI, 1.7 to 3.8 months) in the experimental group versus 2.0 months (95% CI, 1.8 to 2.1 months) in the control group (hazard ratio, 0.88; 95% CI, 0.65 to 1.19; $P = .41$). The study used a variety of tumor types and histologies with single molecular alterations as a predictor for efficacy of targeted agents in heavily pretreated patients, possibly reducing their effectiveness. Nevertheless, there were no differences in progression-free survival or overall survival, and off-label use of targeted agents was discouraged.[38] The benefit of off-label use of molecularly targeted agents is largely debated in oncology; however, most agree that clinical trial enrollment should be encouraged to identify predictive druggable biomarkers.

Other prospective studies, including EXACT,[39] MOSCATO-01 (ClinicalTrials.gov identifier: NCT01566019),[40,41] and MD Anderson's Phase I Clinic[42,43] have shown promising results for targeted therapy on previously treated solid tumors. Additional studies examining the efficacy of off-label use of targeted therapy on the basis of NGS test results are underway: ASCO's Targeted Agent and Profiling Utilization Registry (TAPUR) study (ClinicalTrials.gov identifier: NCT02693535),[44] NCI-MATCH (Molecular Analysis for Therapy Choice) trial (ClinicalTrials.gov identifier: NCT02465060),[45] and Pediatric MATCH (ClinicalTrials.gov identifier: NCT03155620). Results from NCI-MATCH for patients with *PIK3CA* mutations treated with taselisib show 0% objective response rate,[46] whereas patients with *FGFR* pathway mutations treated with AZD4547 showed a 5% objective response rate.[47] Currently, there are no randomized controlled trial results supporting a universal NGS-based treatment paradigm. Despite this, NGS testing continues to be incorporated into the clinic.[48] In a recent survey of 1,281 US oncologists, 75.6% reported using NGS tests to guide treatment decisions. Of these, 34% used them for patients with advanced refractory disease, and 17.5% used them for decisions on off-label use of

**TABLE 5.** Guidelines for Evidence-Based Variant Categories by the AMP, CAP, and ASCO Compared With Levels of Evidence for OncoKB and WfG

| Tier | AMP/CAP/ASCO | OncoKB/Watson |
|---|---|---|
| Tier I: Variants of strong clinical significance | Level A evidence: FDA-approved therapy, included in professional guidelines | Level 1, 2A, R1 |
| | Level B evidence: Well-powered studies with consensus from experts in the field | Level 3A |
| Tier II: Variants of potential clinical significance | Level C evidence: FDA-approved therapies for different tumor types or investigational therapies | Level 2B, 3B |
| | Multiple small published studies with some consensus | |
| | Level D evidence: Preclinical trials or a few case reports without consensus | Level 4 |
| Tier III: Variants of unknown clinical significance | Not observed at a significant allele frequency in the general or specific subpopulation databases, or pan-cancer or tumor-specific variant databases | |
| | No convincing published evidence of cancer association | |
| Tier IV: Benign or likely benign variants | Observed at significant allele frequency in the general or specific subpopulation databases | |
| | No existing published evidence of cancer association | |

Abbreviations: AMP, Association for Molecular Pathology; CAP, College of American Pathologists; FDA, US Food and Drug Administration; WfG, Watson for Genomics.

FDA-approved drugs.[49] More than 50% reported that NGS tests were difficult to interpret either often or sometimes, which is a separate but related challenge facing personalized medicine.[49] The use of NGS will likely continue to grow, with providers facing uncertainty as to how to integrate its use into the clinic. Studies like ours are critical and further highlight the current shortcomings in precisely interpreting results of NGS gene panels for use in clinical management.

### AFFILIATIONS
[1]Department of Radiation Oncology, James A. Haley Veterans Affairs Healthcare System, Tampa, FL
[2]VA National Oncology Program Office, Durham, NC
[3]Medical Service, Durham VA Healthcare System, Durham, NC
[4]Department of Medicine and Duke Cancer Institute, Duke University Medical Center, Durham, NC

### CORRESPONDING AUTHOR
Michael J. Kelley, MD, 508 Fulton St, Hematology-Oncology 111G, Durham, NC 27705; Twitter: @VAResearch; e-mail: michael.kelley@duke.edu.

### AUTHOR CONTRIBUTIONS
**Conception and design:** Evangelia Katsoulakis, Michael J. Kelley
**Provision of study material or patients:** Bradley Hintze
**Collection and assembly of data:** Evangelia Katsoulakis, Jill E. Duffy, Bradley Hintze
**Data analysis and interpretation:** Evangelia Katsoulakis, Bradley Hintze, Neil L. Spector, Michael J. Kelley
**Manuscript writing:** All authors
**Final approval of manuscript:** All authors

### AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

The following represents disclosure information provided by authors of this manuscript. All relationships are considered compensated unless otherwise noted. Relationships are self-held unless noted. I = Immediate Family Member, Inst = My Institution. Relationships may not relate to the subject matter of this manuscript. For more information about ASCO's conflict of interest policy, please refer to www.asco.org/rwc or ascopubs.org/po/author-center.

Open Payments is a public database containing information reported by companies about payments made to US-licensed physicians (Open Payments).

**Evangelia Katsoulakis**
**Stock and Other Ownership Interests:** AbbVie, Alexion Pharmaceuticals, Biogen
**Research Funding:** Advantagene (Inst)

**Neil L. Spector**
**Stock and Other Ownership Interests:** Eydis Bio, Bessor Pharma
**Research Funding:** Immunolight
**Patents, Royalties, Other Intellectual Property:** I am on a patent related to my work with the company Immunolight; Patent through Eydis Bio

## REFERENCES

1. Hayes DN, Kim WY: The next steps in next-gen sequencing of cancer genomes. J Clin Invest 125:462-468, 2015

2. Kandoth C, McLellan MD, Vandin F, et al: Mutational landscape and significance across 12 major cancer types. Nature 502:333-339, 2013

3. Patel NM, Michelini VV, Snell JM, et al: Enhancing next-generation sequencing-guided cancer care through cognitive computing. Oncologist 23:179-185, 2018

4. Morris LG, Riaz N, Desrichard A, et al: Pan-cancer analysis of intratumor heterogeneity as a prognostic determinant of survival. Oncotarget 7:10051-10063, 2016

5. Tamborero D, Rubio-Perez C, Deu-Pons J, et al: Cancer Genome Interpreter annotates the biological and clinical relevance of tumor alterations. Genome Med 10:25, 2018

6. Chakravarty D, Gao J, Phillips SM, et al: OncoKB: A precision oncology knowledge base. JCO Precis Oncol 10.1200/PO.17.00011

7. Taylor AD, Micheel CM, Anderson IA, et al: The path(way) less traveled: A pathway-oriented approach to providing information about precision cancer medicine on My Cancer Genome. Transl Oncol 9:163-165, 2016

8. Huang L, Fernandes H, Zia H, et al: The cancer precision medicine knowledge base for structured clinical-grade mutations and interpretations. J Am Med Inform Assoc 24:513-519, 2017

9. Dumbrava EI, Meric-Bernstam F: Personalized cancer therapy—leveraging a knowledge base for clinical decision-making. Cold Spring Harb Mol Case Stud. 4:a001578, 2018

10. Griffith M, Spies NC, Krysiak K, et al: CIViC is a community knowledgebase for expert crowdsourcing the clinical interpretation of variants in cancer. Nat Genet 49:170-174, 2017

11. Patterson SE, Liu R, Statz CM, et al: The clinical trial landscape in oncology and connectivity of somatic mutational profiles to targeted therapies. Hum Genomics 10:4, 2016

12. Damodaran S, Miya J, Kautto E, et al: Cancer Driver Log (CanDL): Catalog of potentially actionable cancer mutations. J Mol Diagn 17:554-559, 2015

13. N-of-One: A Qiagen Company. Clinical interpretation solutions. https://n-of-one.com/

14. Rhrissorrakrai K, Koyama T, Parida L: Watson for Genomics: Moving personalized medicine forward. Trends Cancer 2:392-395, 2016

15. Robson ME, Bradbury AR, Arun B, et al: American Society of Clinical Oncology policy statement update: Genetic and genomic testing for cancer susceptibility. J Clin Oncol 33:3660-3667, 2015

16. Poonnen PJ, Duffy JE, Hintze B, et al: Genomic analysis of metastatic solid tumors in veterans: Findings from the VHA National Precision Oncology Program. JCO Precis Oncol 10.1200/PO.19.00075

17. Fiore LD, Brophy MT, Turek S, et al: The VA point-of-care precision oncology program: Balancing access with rapid learning in molecular cancer medicine. Biomark Cancer 8:9-16, 2016

18. Personal Genome Diagnostics (PGDx): CLIA certified. https://www.personalgenome.com/

19. Personalis, Inc: A leading precision medicine company focused on advanced NGS based services for cancer and clinical trials and translational research. https://www.personalis.com/

20. Fragoso G, de Coronado S, Haber M, et al: Overview and utilization of the NCI thesaurus. Comp Funct Genomics 5:648-654, 2004

21. OncoKB: Precision oncology knowledge base. https://www.oncokb.org/

22. Cerami E, Gao J, Dogrusoz U, et al: The cBio cancer genomics portal: An open platform for exploring multidimensional cancer genomics data. Cancer Discov 2:401-404, 2012

23. OncoTree: Displaying: oncotree_latest_stable. http://oncotree.mskcc.org/#/home

24. Su J, Zhong W, Zhang X, et al: Molecular characteristics and clinical outcomes of *EGFR* exon 19 indel subtypes to EGFR TKIs in NSCLC patients. Oncotarget 8:111246-111257, 2017

25. Klein O, Clements A, Menzies AM, et al: BRAF inhibitor activity in V600R metastatic melanoma. Eur J Cancer 49:1073-1079, 2013

26. Coit DG, Thompson JA, Algazi A, et al: NCCN guidelines insights: Melanoma, version 3.2016. J Natl Compr Canc Netw 14:945-958, 2016

27. Wu TH, Hsiue EH, Lee JH, et al: New data on clinical decisions in NSCLC patients with uncommon EGFR mutations. Expert Rev Respir Med 11:51-55, 2017

28. Le DT, Uram JN, Wang H, et al: PD-1 blockade in tumors with mismatch-repair deficiency. N Engl J Med 372:2509-2520, 2015

29. Le DT, Durham JN, Smith KN, et al: Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. Science 357:409-413, 2017

30. Gargis AS, Kalman L, Bick DP, et al: Good laboratory practice for clinical next-generation sequencing informatics pipelines. Nat Biotechnol 33:689-693, 2015

31. Roy S, Coldren C, Karunamurthy A, et al: 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 20:4-27, 2018

32. Jennings LJ, Arcila ME, Corless C, et al: 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 19:341-365, 2017

33. Li MM, Datto M, Duncavage EJ, et al: 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 19:4-23, 2017

34. Wagner AH, Walsh B, Mayfield G, et al: A harmonized meta-knowledgebase of clinical interpretations of cancer genomic variants. Nat Genet (in press)

35. Somashekhar SP, Sepúlveda MJ, Puglielli S, et al: Watson for Oncology and breast cancer treatment recommendations: Agreement with an expert multi-disciplinary tumor board. Ann Oncol 29:418-423, 2018

36. Kim YY, Oh SJ, Chun YS, et al: Gene expression assay and Watson for Oncology for optimization of treatment in ER-positive, HER2-negative breast cancer. PLoS One 13:e0200100, 2018

37. West HJ: No solid evidence, only hollow argument for universal tumor sequencing: Show me the data. JAMA Oncol 2:717-718, 2016

38. Le Tourneau C, Delord JP, Gonçalves A, et al: Molecularly targeted therapy based on tumour molecular profiling versus conventional therapy for advanced cancer (SHIVA): A multicentre, open-label, proof-of-concept, randomised, controlled phase 2 trial. Lancet Oncol 16:1324-1334, 2015

39. Prager GW, Unseld M, Waneck F, et al: Results of the extended analysis for cancer treatment (EXACT) trial: A prospective translational study evaluating individualized treatment regimens in oncology. Oncotarget 10:942-952, 2019

40. Verlingue L, Malka D, Allorant A, et al: Precision medicine for patients with advanced biliary tract cancers: An effective strategy within the prospective MOSCATO-01 trial. Eur J Cancer 87:122-130, 2017 [Erratum: Eur J Cancer 93:156-157, 2018]

41. Massard C, Michiels S, Ferté C, et al: High-throughput genomics and clinical outcome in hard-to-treat advanced cancers: Results of the MOSCATO 01 trial. Cancer Discov 7:586-595, 2017

42. Tsimberidou AM, Iskander NG, Hong DS, et al: Personalized medicine in a phase I clinical trials program: The MD Anderson Cancer Center initiative. Clin Cancer Res 18:6373-6383, 2012

43. Tsimberidou AM, Wen S, Hong DS, et al: Personalized medicine for patients with advanced cancer in the phase I program at MD Anderson: Validation and landmark analyses. Clin Cancer Res 20:4827-4836, 2014

44. Mangat PK, Halabi S, Bruinooge SS, et al: Rationale and Design of the Targeted Agent and Profiling Utilization Registry (TAPUR) Study. JCO Precis Oncol 10.1200/PO.18.00122

45. Moore KN, Mannel RS: Is the NCI MATCH trial a match for gynecologic oncology? Gynecol Oncol. 140:161-166, 2016

46. Krop I, Jegede O, Grilley-Olson J, et al: Results from Molecular Analysis for Therapy Choice (MATCH) arm I: Taselisib for PIK3CA-mutated tumors. J Clin Oncol 36:101, 2018 (15_suppl; abstr 101)

47. Chae YK Vaklavas C, Cheng HH, et al: Molecular analysis for therapy choice (MATCH) arm W: Phase II study of AZD4547 in patients with tumors with aberrations in the FGFR. J Clin Oncol 36:2503, 2018 (15_suppl; abstr 2503)

48. Morash M, Mitchell H, Beltran H, et al: The role of next-generation sequencing in precision medicine: A review of outcomes in oncology. J Pers Med 10.3390/jpm8030030

49. Freedman AN, Klabunde CN, Wiant K, et al: Use of next-generation sequencing tests to guide cancer treatment: Results from a nationally representative survey of oncologists in the United States. JCO Precis Oncol 10.1200/PO.18.00169

# APPENDIX

***Veterans Affairs National Precision Oncology Program.*** National Precision Oncology Program (NPOP) services are available to all Veterans Affairs (VA) facilities and as of June 2018, over 72 different VA facilities submitted tumor samples. Following sequencing, a formal report of identified genomic aberrations is collated, annotated and included in patient records. The turnaround time is 14 days. NPOP provides a molecular oncology consultation service to assist VA clinicians with treatment decisions, and a case-based education-focused molecular tumor board. The users of NPOP are VA general oncologists and pathologists. Specialized oncologists and molecular pathologists oversee the program.

***Panel information.*** Panels were constructed to identify base/missense substitutions, insertions/deletions in protein-encoding regions, copy number variations, selected gene fusions/rearrangements and microsatellite instability (only on CancerSELECT 125 panel). There was overlap between the panels and a minimum of 500X DNA sequence coverage was required for all assays.

***OncoKB pathogenicity and actionability.*** Annotation of variant pathogenicity was executed using OncoKB support and through the API (http://oncokb.org/). Annotation of actionability as defined by the OncoKB Levels of Evidence was executed using OncoKB support and are available through the OncoKB annotator (https://github.com/). Data regarding all variants annotated as actionable per OncoKB at the time of the study are publicly available at https://github.com/.

***Panel breakdown of the NGS samples.*** The panel breakdown of the 1,227 NGS samples was: Personalis ACE CancerPlus (n = 404); Personal Genome Diagnostics (PGDx) CancerSELECT 125 (n = 286); PGDx CancerSELECT 203 (n = 75); PGDx CancerSELECT 88 (n = 176); PGDx CancerSELECT 64 (n = 1); PGDx unknown panel (n = 285).
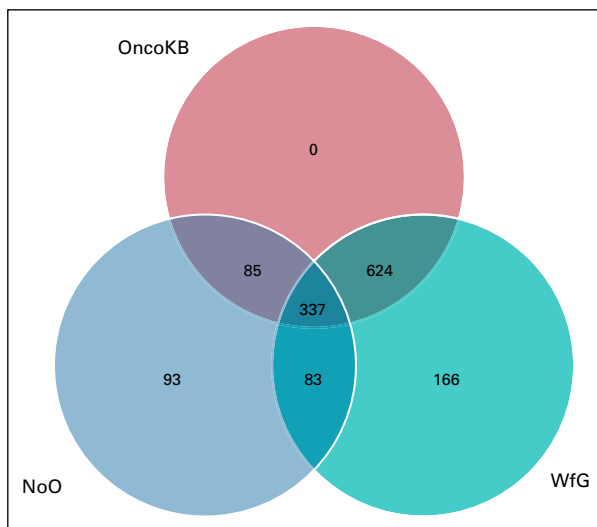


**FIG A1.** Venn diagram showing the relationship between pathogenicity annotation for N-of-One (NoO), Watson for Genomics (WfG), and OncoKB.