



# A Review of Precision Oncology Knowledgebases for Determining the Clinical Actionability of Genetic Variants

Xuanyi Li<sup>1</sup> and Jeremy L. Warner<sup>2,3\*</sup>

<sup>1</sup> Vanderbilt University School of Medicine, Nashville, TN, United States, <sup>2</sup> Department of Medicine, Vanderbilt University, Nashville, TN, United States, <sup>3</sup> Department of Biomedical Informatics, Vanderbilt University, Nashville, TN, United States

## OPEN ACCESS

### Edited by:

Howard Raymond Petty,  
University of Michigan, United States

### Reviewed by:

Joel Lefferts,  
Dartmouth College, United States

Alexis Carter,  
Children's Healthcare of Atlanta,  
United States

Jennifer Laudadio,  
University of Arkansas for Medical  
Sciences, United States

### \*Correspondence:

Jeremy L. Warner  
jeremy.warner@vumc.org

### Specialty section:

This article was submitted to  
Molecular Medicine,  
a section of the journal  
Frontiers in Cell and Developmental  
Biology

**Received:** 01 October 2019

**Accepted:** 20 January 2020

**Published:** 11 February 2020

### Citation:

Li X and Warner JL (2020) A  
Review of Precision Oncology  
Knowledgebases for Determining  
the Clinical Actionability of Genetic  
Variants. *Front. Cell Dev. Biol.* 8:48.  
doi: 10.3389/fcell.2020.00048

The increased availability of tumor genetic testing and targeted cancer therapies contributes to the advancement of precision medicine in the field of oncology. Precision oncology knowledgebases provide a way of organizing clinically relevant genetic information in a way that is easily accessible for both oncologists and patients, facilitating the genetic-based clinical decision making. Many organizations and companies have built precision oncology knowledgebases, intended for multiple users. In general, these knowledgebases offer information on cancer-related genetic variants as well as their associated diagnostic, prognostic, and therapeutic implications, but they often differ in their information curations, designs, and user experiences. It is advisable that oncologists use multiple knowledgebases during their practice to have them complement each other. In the future, convergence toward common standards and formats is needed to ensure that the comprehensive knowledge across all sources can be unified to bring the oncology community closer to the achievement of the goal of precision oncology.

**Keywords:** precision oncology, knowledgebase, variant interpretation, actionability, targeted therapy, genotype-selective clinical trial, tumor genetic testing

## INTRODUCTION

Precision oncology is defined as utilizing tumor molecular profiles to identify diagnostic, prognostic, and therapeutic implications pertaining to the specific tested cancer (Schwartzberg et al., 2017). Fundamentally, precision oncology is based upon the idea that tumor biomarkers are predictive of disease phenotype, clinical outcomes, and therapy responses. The first-ever cancer precision therapy under this definition was imatinib for chronic myeloid leukemia (Druker et al., 2001). Since then, a growing number of therapies targeted at specific genetic alterations have been introduced. Targeted therapies have led to improved outcomes in cancer patients with targetable mutations (Schwaederle et al., 2015, 2016). They have been shown to benefit patients with hard-to-treat cancer without increasing healthcare costs (Massard et al., 2017; Haslem et al., 2017).

Tumor molecular profiles consist of DNA, RNA, and protein alterations, as well as epigenetic changes. Currently, the most widely used molecular profiling method is tumor DNA next-generation sequencing (NGS) (Tran et al., 2013; Van Allen et al., 2014). The introduction of NGS into clinical oncology has provided oncologists with a large amount of genomic information,

which can be incorporated into clinical decision making. As shown in **Figure 1**, the tasks of ordering, measuring, interpreting, and acting upon NGS information are shared between laboratory professionals and pathologists and oncologists. Not all genomic information from NGS is clinically relevant. After genetic variants are identified through NGS, several additional steps are required and are typically performed by laboratory professionals and pathologists. “Noises” such as germline polymorphisms (Tuchman et al., 1997; Calado et al., 2008; Muller et al., 2010), false-positive artifacts generated by NGS technology, and clinically insignificant synonymous variants are filtered out from the report. Moreover, the clinical significance and actionability of the remaining variants will be assessed by laboratory professionals and pathologists in order to identify potential treatment targets (Gao et al., 2019). This complicated process causes difficulties in the implementation of precision cancer medicine in the clinic, which is additionally hampered by incomplete incorporation into clinical health information systems (Warner et al., 2016). Clinical evidence on tumor molecular biomarkers can be scattered and chaotic. With the introduction of large numbers of targeted therapy drugs and genotype-selected clinical trials, it is challenging for oncologists to fully explore all relevant treatment options.

For this purpose, there are several precision oncology knowledgebases developed at least in part for providing clinical decision support for oncologists in interpreting genomic data and identifying therapy targets (Bates et al., 2003). They can help oncologists determine therapeutic options and clinical trial availability. With few exceptions, the bulk of the content of these knowledgebases is based upon curation of published literature. These knowledgebases were created with slightly different goals in mind and differ in data curation approaches, information organization, and anticipated user experiences. The goal of this focused review is to introduce the knowledgebases specific to the tasks of interpretation and clinical decision making highlighted in **Figure 1**.

## PUBLICLY AVAILABLE PRECISION ONCOLOGY KNOWLEDGEBASES

In response to the increasing demand for assistance with utilizing NGS information, several institutions have developed dedicated services e.g., the precision oncology decision support team at MD Anderson (Johnson et al., 2017) and the precision cancer medicine team at the Dana-Farber Cancer Institute (2019). Similarly, an increasing number of institutions are implementing molecular tumor boards. For oncologists who do not have access to such support, they may elect to use the publicly available precision oncology knowledgebases (**Table 1**), which allow users to access at least part of their curated content without charge. In alphabetical order:

### Cancer Genome Interpreter (CGI)

Cancer Genome Interpreter (CGI) is a knowledgebase designed for the streamlining and automatization of the entire process

of interpretation of cancer genomes (Tamborero et al., 2018). It incorporates several different databases for the annotation of alterations, the identification of driver mutations, the determination of variant actionability, and exploration of biomarker interactions. CGI is notable for its attempt to assess the tumorigenic potential of variants of unknown significance (VUS). VUS represents one of the challenges of interpretation of genomic information and significantly contributes to uncertainty in clinical decision making (Richter et al., 2013). CGI assesses for the tumorigenic potential of VUS via OncodriveMUT. OncodriveMUT employs a rule-based approach that combines features of the VUS, including the action of the gene, the consequence of the mutation, its position within the transcript, its prevalence within the human population, and whether the mutation occurs in a domain of the protein that is depleted of germline variants. CGI is hosted at the Barcelona Biomedical Genomics Lab.

### Clinical Interpretations of Variants in Cancer (CIViC)

Clinical Interpretations of Variants in Cancer (CIViC) is a crowdsourced knowledgebase for the clinical implications of cancer genome variants (Griffith et al., 2017). CIViC contains 6308 curated clinical evidence records for 2278 variants affecting 396 genes. Each evidence record is associated with a specific gene variant and contains information related to therapy, prognostics, diagnostics, or predisposition for cancer.

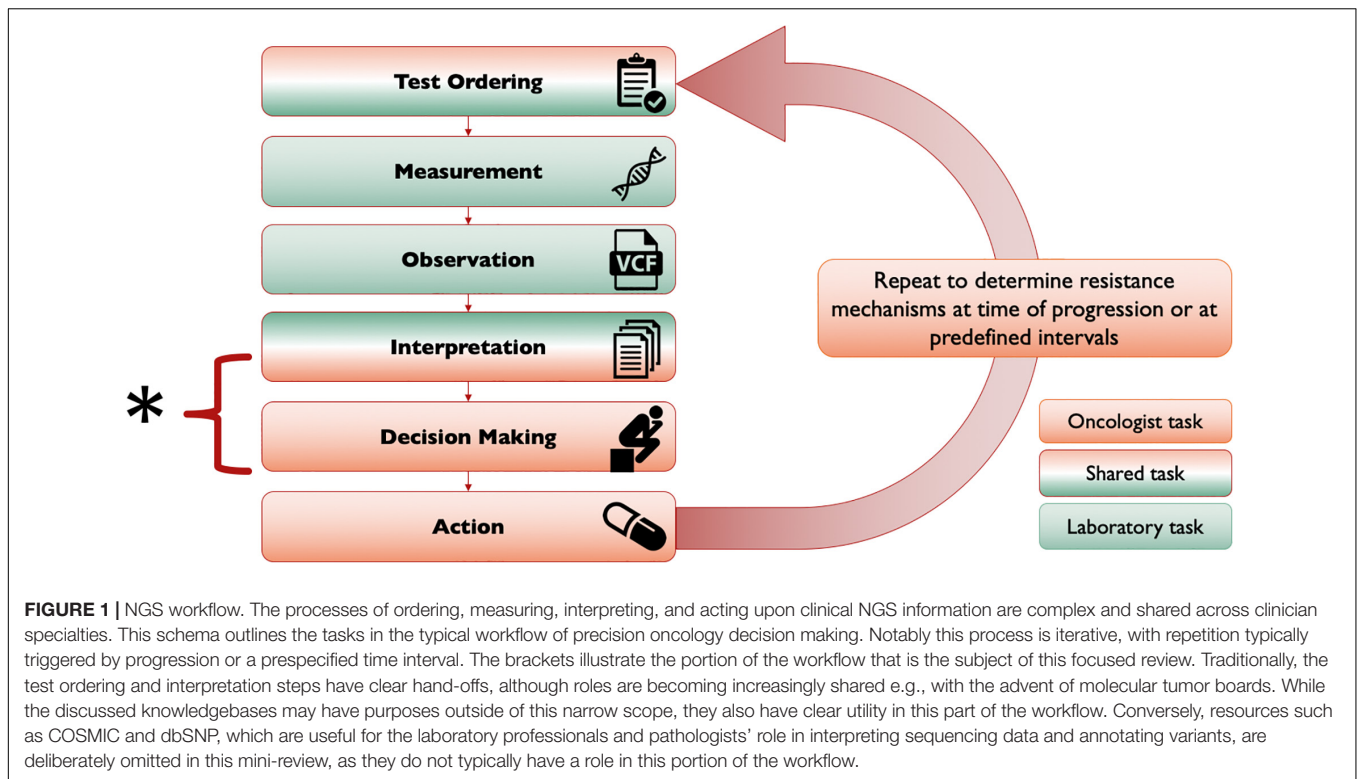
Clinical Interpretations of Variants in Cancer reports an evidence level ranging from established clinical utility (level A) to inferential (level E) for each evidence record it reports (Griffith et al., 2017). Also, CIViC's open model can potentially facilitate the standardization of clinical variant interpretation but can also make it challenging to quality control the curated evidence. CIViC is hosted at the Washington University in St. Louis School of Medicine.

### Database of Evidence for Precision Oncology

Database of evidence for precision oncology (DEPO) is a curated knowledgebase of clinically relevant genomic variants and their respective drug therapies (Sun et al., 2018). It has a web portal that allows users to query the knowledgebase using standard gene names and optionally, cancer type, variant type or position. Notably, besides incorporating the present variant actionability information, DEPO also employs HotSpot3D (Niu et al., 2016), which analyzes the clustering of mutations on a three-dimensional protein structural model for predicting potential treatment targets. Also, DEPO utilizes the JSmol 3D molecular viewer to visualize the 3D mutational clusters and potentially druggable sites. DEPO is hosted by the Ding lab at Washington University in St. Louis.

### HemOnc.org

HemOnc.org is a clinical knowledgebase for hematology and oncology providers that includes information on cancer drugs



**TABLE 1 |** Summary of publicly available precision oncology knowledgebases.

Knowledgebase	Variant annotation	Drug availability	Trial matching	Literature citation	Website
CGI	Yes	Yes	No	Yes	<a href="http://www.cancergenomeinterpreter.org/">www.cancergenomeinterpreter.org/</a>
CIVIC	Yes	Yes	No	Yes	<a href="https://civicdb.org">https://civicdb.org</a>
DEPO	Yes	Yes	No	Yes	<a href="http://depo-dinglab.ddns.net/">http://depo-dinglab.ddns.net/</a>
HemOnc.org	Yes	Yes	No	Yes	<a href="http://www.hemonc.org/">www.hemonc.org/</a>
JAX CKB <sup>+</sup>	Yes	Yes	Yes*	Yes	<a href="https://ckb.jax.org/">https://ckb.jax.org/</a>
MCG	Yes	Yes	Yes	No	<a href="https://www.mycancergenome.org/">https://www.mycancergenome.org/</a>
OncoKB	Yes	Yes	No	Yes	<a href="https://oncokb.org/">https://oncokb.org/</a>
PCT	Yes	Yes	Yes	Yes	<a href="https://pct.mdanderson.org/">https://pct.mdanderson.org/</a>
PMKB	Yes	Yes	No	Yes	<a href="https://pmkb.weill.cornell.edu/">https://pmkb.weill.cornell.edu/</a>
<b>Drug-associated knowledgebases</b>					
Genomics of drug sensitivity in cancer	Yes	Yes	No	No	<a href="https://www.cancerrxgene.org/">https://www.cancerrxgene.org/</a>
PharmGKB	Yes	Yes	No	No	<a href="https://www.pharmgkb.org/">https://www.pharmgkb.org/</a>
Therapeutic target database	Yes	Yes	No	Yes	<a href="http://db.idrblab.net/ttd/">http://db.idrblab.net/ttd/</a>

\*Available only with subscription to CKB BOOST. <sup>+</sup>Partially publicly available databases.

and regimens (JW is the Deputy Editor of HemOnc.org and co-founder of HemOnc.org LLC). Although HemOnc.org is not specifically a precision oncology knowledgebase, approximately 8% of its content is organized by genomic alterations. For example, non-small cell lung cancer (NSCLC) regimens are organized by EGFR, ALK, BRAF, and ROS1 mutation status; kinase inhibitors are categorized by their target kinase(s) and biomarker-specific FDA labeling is noted. These concepts and relationships are modeled in the derivative publicly available HemOnc vocabulary (Warner et al., 2019). Near-term efforts include structuring the genomic eligibility criteria for genomically defined regimens

at the gene and variant levels. HemOnc.org is hosted by HemOnc.org LLC.

## Jackson Laboratory Clinical Knowledge Base (JAX CKB)

Jackson Laboratory Clinical Knowledge Base (JAX CKB) is a knowledgebase that incorporates integrated data related to cancer-associated genomic variants, therapeutic efficacy, and clinical trials for interpretation of genomic data in cancer (Patterson et al., 2016). The public access version of CKB contains information on 82 commonly known driver genes. On the JAX

CKB website, users can search by gene, gene variants, drug, drug class, indication, and clinical trials. Notably, JAX CKB can build complex molecular profiles that allow for the association of therapeutic efficacy to multiple genetic alterations simultaneously instead of each variant in isolation (Patterson et al., 2019). JAX CKB is hosted at The Jackson Laboratory.

### My Cancer Genome (MCG)

My Cancer Genome (MCG) marks one of the earliest attempts at creating an online knowledge resource describing the clinical actionability of tumor molecular biomarkers (Taylor et al., 2016). The MCG website offers information on targeted therapy and clinical trials for an array of genetic variants that are involved in signaling pathways affected by different types of cancers. At the time of writing, MCG contains 2344 molecular biomarkers, including genetic biomarkers, protein expression markers, chromosomal markers, and markers of genomic instability. For each biomarker, MCG has information on the related disease, available drugs, as well as information on clinical trials that are currently recruiting.

Unlike many other precision oncology knowledgebases, MCG organizes clinical evidence for genomic variants in a disease-centric approach instead of a gene-centric approach (Gao et al., 2019). MCG distinguishes between genetic variants within the same gene (Swanton, 2012), which could potentially discourage the use of targeted agents outside their indicated genetic variants. MCG has a user-friendly web interface and is beneficial for patient education and empowerment (Kusnoor et al., 2016). Additionally, MCG uses a pathway approach to curate genetic variants, which supports pathway-based treatment strategies with vertical and parallel inhibition. A limitation of MCG is that clinical evidence is not prioritized based on a predefined schema for levels of evidence. MCG is hosted at Vanderbilt University Medical Center and has a commercial relationship with GenomOncology LLC (Cleveland, OH, United States).

### Personalized Cancer Therapy (PCT)

Personalized Cancer Therapy (PCT) is a knowledgebase for clinical evidence of tumor genomic variants (Kurnit et al., 2017; Dumbrava and Meric-Bernstam, 2018). At the time of writing, PCT provides publicly accessible information for 32 actionable genes. For each gene, PCT includes general gene information, variants in that gene, frequencies and outcomes of these variants, therapeutic implications of variants, FDA-approved drugs or investigational therapeutics in clinical trials, as well as information on related genotype-selected and -relevant clinical trials.

Unlike MCG, PCT is organized in a gene-centric and patient tumor type-agnostic fashion (Gao et al., 2019). This organizational approach is designed to facilitate the process of identifying available targeted therapies during a patient encounter. Notably, a substantial portion of this knowledgebase is not available to the public and is instead used only for in-house precision oncology decision support. PCT is hosted at the MD Anderson Cancer Center Sheikh Khalifa Bin Zayed Al Nahyan Institute for Personalized Cancer Therapy.

### Precision Medicine Knowledge Base (PMKB)

Precision Medicine Knowledge Base (PMKB) is an online knowledgebase for structured clinical-grade cancer mutation interpretations (Huang et al., 2017). At the time of writing, PMKB contains 2246 variants with 1767 graded interpretations. Similar to CIViC, PMKB rates variant interpretations by a numeric tier indicating the clinically actionability with tier 1 as strong evidence of clinical utility, tier 2 as potential clinical relevance, and tier 3 as undetermined clinical significance (Huang et al., 2017). The majority of variant interpretations in PMKB are rated as tier 2. Also, PMKB supports crowdsourcing of content by contributors in specific subspecialties. PMKB is hosted at Weill Cornell Medicine Englander Institute for Precision Medicine.

### Precision Oncology Knowledge Base (OncoKB)

Precision Oncology Knowledge Base (OncoKB) is a precision oncology knowledgebase of tumor markers and their related FDA-approved therapies and investigational agents that are under evaluation in clinical trials (Chakravarty et al., 2017). OncoKB offers information for about 5000 genetic variants in 642 cancer-associated genes from 45 tumor types. OncoKB is also tightly integrated into the local environment of the Memorial Sloan Kettering Cancer Center and can take information directly from the MSK-IMPACT testing to produce a nearly automated report (Schram et al., 2017; Zehir et al., 2017).

OncoKB highlights adverse outcomes of off-label use of targeted therapies in specific mutational contexts (Gao et al., 2019). Off-label use of cancer drugs is prevalent, especially in patients with rare cancers, for which randomized clinical trials may not be feasible (Conti et al., 2013). However, off-label medication use can be costly and can lead to adverse effects and compromise clinical trials (Krzyzanowska, 2013). A limitation of OncoKB is that the website does not contain information on genotype-selective clinical trials. OncoKB is hosted at the Memorial Sloan Kettering Cancer Center.

### Drug-Associated Knowledgebases

Several drug-associated knowledgebases also provide information on variant drugability. The therapeutic target database (TTD) provides information about the therapeutic targets and corresponding approved and experimental drugs (Liu et al., 2011). Its data can be explored for target and drug searches and the development of *in silico* target discovery tools. The Genomics of Drug Sensitivity in Cancer (GDSC) contains information on drug sensitivity in cancer cells and molecular markers of drug response (Yang et al., 2013). Notably, GDSC contains large scale genomic and cell-line anticancer drug sensitivity datasets that could facilitate the discovery of new therapeutic biomarkers. The Pharmacogenomics Knowledge Base (PharmGKB) is a public repository of information relevant to pharmacogenetics and pharmacogenomics (Klein and Altman, 2004; Whirl-Carrillo et al., 2012). It provides clinically relevant information, including dosing guidelines, annotated drug



labels, and potentially actionable gene-drug associations and genotype-phenotype relationships.

## COMMERCIALLY AVAILABLE KNOWLEDGEBASES

Besides those above publicly available precision oncology knowledgebases, there also exist several commercially available knowledgebases for interpreting cancer molecular information (Table 2). Several companies offer variant interpretation functionality that is bundled with their sequencing panel workflow (Zeng et al., 2019). Their reports may vary on whether variant annotations or literature citations are offered, but most are consistently reporting available targeted therapies and clinical trial matching.

There are also commercially available stand-alone variant interpretation tools. Their curated efforts are somewhat similar to the publicly available knowledgebases, but less is known about their data curation process. A recent study has been performed to compare some of these commercially available knowledgebases, which shows that they agree upon clinically relevant variants but have a relatively low overall concordance of evidence levels (Sakai et al., 2019).

## PREDICTION-BASED APPROACHES

Several recent efforts have moved away from literature curation and toward the automated prediction of whether an observed somatic mutation is a driver of functional significance. These efforts are based on a broad array of bioinformatics pipelines and have succeeded in identifying functional mutations at scale (Bailey et al., 2018; Bertrand et al., 2018). However, there are studies showing that prediction models may have low specificity (Flanagan et al., 2010; Ernst et al., 2018). With the exception of CGI described above, they are out of scope

for this review, which is primarily concerned with literature-based knowledge.

## HARMONIZATION

With the large number of public and private knowledgebases now available, they will inevitably have differing content coverage and are likely to have discrepancies and disagreements. In recognition of this increasingly problematic situation, the Global Alliance for Genomics and Health (GA4GH) founded the Variant Interpretation for Cancer Consortium (VICC) as a Driver Project. One of the primary goals of the VICC is knowledgebase integration, through the creation of a meta-knowledgebase. This effort concatenated the results of six knowledgebases: CGI (2.7), CIViC (2.4), JAX CKB (2.6), MolecularMatch (commercial; not described in detail), OncoKB (2.3), and PMKB (2.5). This integration has resulted in 12,856 aggregate interpretations covering 3,437 unique variants in 415 genes, 357 diseases, and 791 drugs (Wagner et al., 2018).

## DISCUSSION

Precision oncology knowledgebases offer potential solutions to many challenges associated with the implementation of an individualized treatment approach in clinical oncology (Andre et al., 2014). At their best, they are user-friendly repositories that organize complex information into easy-to-digest clinical evidence that can be applied to clinical decision making in real-time (Kusnoor et al., 2016; Dumbrava and Meric-Bernstam, 2018). They can also facilitate patient communication, which empowers patients to take an active role in their cancer treatment (Kusnoor et al., 2016).

On the other hand, there are several challenges associated with knowledgebases. Currently, few precision oncology

**TABLE 2** | Summary of commercially available precision oncology knowledgebases.

Knowledgebases	Variant annotation	Drug availability	Trial matching	Literature citation
<b>Laboratory-based knowledgebases</b>				
Caris molecular intelligence	Yes	Yes	Yes	No
FoundationOne	Yes	Yes	Yes	Yes
Guardant360	No	Yes	Yes	No
NeoTYPE	Yes	Yes	Yes	No
Perthera	No	Yes	Yes	No
StrataNGS	No	Yes	Yes	No
Tempus	No	Yes	Yes	Yes
<b>Stand-alone knowledgebases</b>				
BaseSpace knowledge network	Yes	Yes	Yes	Yes
MolecularMatch	Yes	Yes	Yes	Yes
OncoPrint knowledge-based reporter	Yes	Yes	Yes	Yes
QIAGEN clinical insight interpret	Yes	Yes	Yes	Yes
Watson for genomics	Yes	Yes	Yes	Yes

*This is a sample of the more widely known sources; omission is not intentional.*

knowledgebases are integrated into clinical workflows, making it difficult to assess user experience. There is also the issue with accountability and trustworthiness of the assertions made in the knowledgebases, which is important since they may affect clinical decisions. Most knowledgebases have extensive disclaimers and exculpatory clauses. However, it is unknown whether there would be legal liability in the case of malpractice lawsuits brought against a medical decision that directly involved the use of knowledgebases. Although there is no current FDA guidance for the clinical validity of somatic knowledgebases, it is possible that guidance similar to that for germline databases will be issued in the future (U.S. FDA, 2018). Precision oncology knowledgebases do not replace the experience and knowledge of a clinician, nor are they meant to replace a careful read of the primary literature. With each knowledgebase having its own profile of strengths and limitations, it is advisable that clinicians use multiple knowledgebases during their practice to take advantage of complementary information (Hughes et al., 2017). We have explored enabling such an approach with an app called SMART Cancer Navigator (Warner et al., 2018), which collects information from many knowledgebases for presentation to a clinical user, using application programming interfaces.

Cancer treatment is becoming more personalized. Although not addressed here, this includes integration of germline and family history information, which has been typically ignored in somatic NGS interpretations. For certain patient populations, e.g., children, the incidence of finding germline pathogenic variants causing hereditary cancer predisposition syndromes is quite high (Zhang et al., 2015). The major challenge as tumors are increasingly molecularly characterized is that literature support for any given mutation and especially for combinations of mutations and other factors such as germline will decrease. A potential solution is to rely on case reports of precision oncology; however, case reports are notoriously difficult to publish, are generally undervalued in the academic setting, and are unlikely to capture the true

breadth of the clinical spectrum. Instead, there are several efforts intended to gather precision oncology diagnoses and outcomes at scale; many of these efforts are commercial e.g., (Singal et al., 2019). One public effort to create such a shared knowledgebase is the AACR Project GENIE, which is a multi-disciplinary effort that integrates de-identified clinical-grade cancer genomic data with clinical outcome data for tens of thousands of cancer patients treated at multiple institutions worldwide (AACR Project Genie Consortium, 2017). The GENIE database (JW has grant support from GENIE) has been used to train machine-learning classifiers and is in the process of adding more in-depth phenotype information (Nicora et al., 2019). In the future, the international oncology community should continue to converge on common standards for knowledge representation, so that a comprehensive and unified knowledgebase that links tumor molecular profile data with approved therapies and available clinical trials, which would lead to the achievement of the goal of precision oncology, can be achieved.

## AUTHOR CONTRIBUTIONS

XL and JW made substantial contributions to the conception and design of the work, drafted and revised the work, approved the final submitted version, and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

## FUNDING

JW was supported in part by NCI grant CA231840 and the Vanderbilt University Office of Health Sciences Education (OHSE). XL was supported in part by the Ann Melly Summer Scholarship in Oncology.

## REFERENCES

- AACR Project Genie Consortium, (2017). AACR project GENIE: powering precision medicine through an international consortium. *Cancer Discov.* 7, 818–831. doi: 10.1158/2159-8290.CD-17-0151
- Andre, F., Mardis, E., Salm, M., Soria, J. C., Siu, L. L., and Swanton, C. (2014). Prioritizing targets for precision cancer medicine. *Ann. Oncol.* 25, 2295–2303. doi: 10.1093/annonc/mdu478
- Bailey, M. H., Tokheim, C., Porta-Pardo, E., Sengupta, S., Bertrand, D., Weerasinghe, A., et al. (2018). Comprehensive characterization of cancer driver genes and mutations. *Cell* 173, 371–385. doi: 10.1016/j.cell.2018.02.060
- Bates, D. W., Kuperman, G. J., Wang, S., Gandhi, T., Kittler, A., Volk, L., et al. (2003). Ten commandments for effective clinical decision support: making the practice of evidence-based medicine a reality. *J. Am. Med. Assoc.* 289, 523–530. doi: 10.1197/jam.2003.11.523
- Bertrand, D., Drissler, S., Chia, B. K., Koh, J. Y., Li, C., Suphavilai, C., et al. (2018). ConsensusDriver improves upon individual algorithms for predicting driver alterations in different cancer types and individual patients. *Cancer Res.* 78, 290–301. doi: 10.1158/0008-5472.CCR-17-1345
- Calado, J., Sznajder, Y., Metzger, D., Rita, A., Hogan, M. C., Kattamis, A., et al. (2008). Twenty-One additional cases of familial renal glucosuria: absence of genetic heterogeneity, high prevalence of private mutations and further evidence of volume depletion. *Nephrol. Dial. Transplant.* 23, 3874–3879. doi: 10.1093/ndt/gfn386
- Chakravarty, D., Gao, J., Phillips, S. M., Kundra, R., Zhang, H., Wang, J., et al. (2017). OncoKB: a precision oncology knowledge base. *JCO Precis. Oncol.* 2017, 1–16. doi: 10.1200/PO.17.00011
- Conti, R. M., Bernstein, A. C., Villafior, V. M., Schilsky, R. L., Rosenthal, M. B., and Bach, P. B. (2013). "Prevalence of off-label use and spending in 2010 among patent-protected chemotherapies in a population-based cohort of medical oncologists. *J. Clin. Oncol.* 31, 1134–1139. doi: 10.1200/JCO.2012.42.7252
- Dana-Farber Cancer Institute, (2019). *The Center for Cancer Precision Medicine Team*. Available at: <https://www.dana-farber.org/research/departments-centers-and-labs/integrative-research-centers/center-for-cancer-precision-medicine/our-team/> [Accessed September 29, 2019].
- Druker, B. J., Talpaz, M., Resta, D. J., Peng, B., Buchdunger, E., Ford, J. M., et al. (2001). Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in chronic myeloid leukemia. *N. Engl. J. Med.* 344, 1031–1037. doi: 10.1056/NEJM200104053441401
- Dumbrava, E. I., and Meric-Bernstam, F. (2018). Personalized cancer therapy-leveraging a knowledge base for clinical decision-making. *Cold Spring Harb. Mol. Case Stud.* 4:a001578. doi: 10.1101/mcs.a001578

- Ernst, C., Hahnen, E., Engel, C., Nothnagel, M., Weber, J., Schmutzler, R. K., et al. (2018). Performance of in silico prediction tools for the classification of rare BRCA1/2 missense variants in clinical diagnostics. *BMC Med. Genom.* 11:35. doi: 10.1186/s12920-018-0353-y
- Flanagan, S. E., Patch, A., and Ellard, S. (2010). Using SIFT and PolyPhen to predict loss-of-function and gain-of-function mutations. *Genet. Test. Mol. Biomarkers* 14, 533–537. doi: 10.1089/gtmb.2010.0036
- Gao, P., Zhang, R., and Li, J. (2019). Comprehensive elaboration of database resources utilized in next-generation sequencing-based tumor somatic mutation detection. *Biochim. Biophys. Acta Rev. Cancer* 1872, 122–137. doi: 10.1016/j.bbcan.2019.06.004
- Griffith, M., Spies, N. C., Krysiak, K., McMichael, J. F., Coffman, A. C., Danos, A. M., et al. (2017). CIViC Is a community knowledgebase for expert crowdsourcing the clinical interpretation of variants in cancer. *Nat. Genet.* 49, 170–174. doi: 10.1038/ng.3774
- Haslem, D. S., Burke Van Norman, S., Fulde, G., Knighton, A. J., Belnap, T., Butler, A. M., et al. (2017). A retrospective analysis of precision medicine outcomes in patients with advanced cancer reveals improved progression-free survival without increased health care costs. *J. Oncol. Pract.* 13, e108–e119. doi: 10.1200/JOP.2016.011486
- Huang, L., Fernandes, H., Zia, H., Tavassoli, P., Rennert, H., Pisapia, D., et al. (2017). The cancer precision medicine knowledge base for structured clinical-grade mutations and interpretations. *J. Am. Med. Inform. Assoc.* 24, 513–519. doi: 10.1093/jamia/ocw148
- Hughes, K. S., Ambinder, E. P., Hess, G. P., Yu, P. P., Bernstam, E. V., Routbort, M. J., et al. (2017). Identifying health information technology needs of oncologists to facilitate the adoption of genomic medicine: recommendations from the 2016 American Society of Clinical Oncology Omics and Precision Oncology workshop. *J. Clin. Oncol.* 35, 3153–3159. doi: 10.1200/JCO.2017.74.1744
- Johnson, A., Khotskaya, Y. B., Brusco, L., Zeng, J., Holla, V., Bailey, A. M., et al. (2017). Clinical use of precision oncology decision support. *JCO Precis. Oncol.* 1, 1–12. doi: 10.1200/PO.17.00036
- Klein, T. E., and Altman, R. B. (2004). “PharmGKB: the pharmacogenetics and pharmacogenomics knowledge base. *Pharmacogenomics J.* 4, 1–1. doi: 10.1038/sj.tpj.6500230
- Krzyzanowska, M. K. (2013). Off-Label use of cancer drugs: a benchmark is established. *J. Clin. Oncol.* 31, 1125–1127. doi: 10.1200/JCO.2012.46.9460
- Kurnit, K. C., Bailey, A. M., Zeng, J., Johnson, A. M., Abu Shufean, M., Brusco, L., et al. (2017). Personalized cancer therapy: a publicly available precision oncology resource. *Cancer Res.* 77, e123–e126. doi: 10.1158/0008-5472.can-17-0341
- Kusnoor, S. V., Koonce, T. Y., Levy, M. A., Lovly, C. M., Naylor, H. M., Anderson, I. A., et al. (2016). My cancer genome: evaluating an educational model to introduce patients and caregivers to precision medicine information. *AMIA Summits Transl. Sci. Proc.* 2016, 112–121.
- Liu, X., Zhu, F., Ma, X., Tao, L., Zhang, J., Yang, S., et al. (2011). The therapeutic target database: an internet resource for the primary targets of approved, clinical trial and experimental drugs. *Expert Opin. Ther. Targets* 15, 903–912. doi: 10.1517/14728222.2011.586635
- Massard, C., Michiels, S., Fertié, C., Le Deley, M. C., Lacroix, L., Hollebecque, A., et al. (2017). High-throughput genomics and clinical outcome in hard-to-treat advanced cancers: results of the MOSCATO 01 trial. *Cancer Discov.* 7, 586–595. doi: 10.1158/2159-8290.CD-16-1396
- Muller, J., Stoetzel, C., Vincent, M. C., Leitch, C. C., Laurier, V., Danse, J. M., et al. (2010). Identification of 28 novel mutations in the bardet–biedl syndrome genes: the burden of private mutations in an extensively heterogeneous disease. *Hum. Genet.* 127, 583–593. doi: 10.1007/s00439-010-0804-9
- Nicora, G., Marini, S., Limongelli, I., Rizzo, E., Montoli, S., Tricomi, F. F., et al. (2019). “A semi-supervised learning approach for pan-cancer somatic genomic variant classification,” in *Artificial Intelligence in Medicine*, eds D. Riaño, S. Wilk, and A. ten Teije, (Cham: Springer International Publishing), 42–46. doi: 10.1007/978-3-030-21642-9\_7
- Niu, B., Scott, A. D., Sengupta, S., Bailey, M. H., Batra, P., Ning, J., et al. (2016). Protein structure guided discovery of functional mutations across 19 cancer types. *Nat. Genet.* 48, 827–837. doi: 10.1038/ng.3586
- Patterson, S. E., Liu, R., Statz, C. M., Durkin, D., Lakshminarayana, A., and Mockus, S. M. (2016). The clinical trial landscape in oncology and connectivity of somatic mutational profiles to targeted therapies. *Hum. Genom.* 10:4. doi: 10.1186/s40246-016-0061-7
- Patterson, S. E., Statz, C. M., Yin, T., and Mockus, S. M. (2019). Utility of the JAX clinical knowledgebase in capture and assessment of complex genomic cancer data. *NPJ Precis. Oncol.* 3, 1–6. doi: 10.1038/s41698-018-0073-y
- Richter, S., Haroun, I., Graham, T. C., Eisen, A., Kiss, A., and Warner, E. (2013). Variants of unknown significance in BRCA testing: impact on risk perception, worry, prevention and counseling. *Ann. Oncol.* 24(Suppl.\_8), viii69–viii74. doi: 10.1093/annonc/mdt312
- Sakai, K., Takeda, M., Shimizu, S., Takahama, T., Yoshida, T., Watanabe, S., et al. (2019). A comparative study of curated contents by knowledge-based curation system in cancer clinical sequencing. *Sci. Rep.* 9:11340. doi: 10.1038/s41598-019-47673-9
- Schram, A. M., Reales, D., Galle, J., Cambria, R., Durany, R., Feldman, D., et al. (2017). Oncologist use and perception of large panel next-generation tumor sequencing. *Ann. Oncol.* 28, 2298–2304. doi: 10.1093/annonc/mdx294
- Schwaederle, M., Zhao, M., Lee, J. J., Eggermont, A. M., Schilsky, R. L., Mendelsohn, J., et al. (2015). Impact of precision medicine in diverse cancers: a meta-analysis of phase ii clinical trials. *J. Clin. Oncol.* 33, 3817–3825. doi: 10.1200/JCO.2015.61.5997
- Schwaederle, M., Zhao, M., Lee, J. J., Lazar, V., Leyland-Jones, B., Schilsky, R. L., et al. (2016). Association of biomarker-based treatment strategies with response rates and progression-free survival in refractory malignant neoplasms: a meta-analysis. *JAMA Oncol.* 2, 1452–1459. doi: 10.1001/jamaoncol.2016.2129
- Schwartzberg, L., Kim, E. S., Liu, D., and Schrag, D. (2017). Precision oncology: who, how, what, when, and when not? *Am. Soc. Clin. Oncol. Educ. Book* 37, 160–169. doi: 10.1200/EDBK\_174176
- Singal, G., Miller, P. G., Agarwala, V., Li, G., Kaushik, G., Backenroth, D., et al. (2019). Association of patient characteristics and tumor genomics with clinical outcomes among patients with non-small cell lung cancer using a clinicogenomic database. *JAMA* 321, 1391–1399. doi: 10.1001/jama.2019.3241
- Sun, S. Q., Mashl, R. J., Sengupta, S., Scott, A. D., Wang, W., Batra, P., et al. (2018). Database of evidence for precision oncology portal. *Bioinformatics* 34, 4315–4317. doi: 10.1093/bioinformatics/bty531
- Swanton, C. (2012). My cancer genome: a unified genomics and clinical trial portal. *Lancet Oncol.* 13, 668–669. doi: 10.1016/s1470-2045(12)70312-1
- Tamborero, D., Rubio-Perez, C., Deu-Pons, J., Schroeder, M. P., Vivancos, A., Rovira, A., et al. (2018). Cancer genome interpreter annotates the biological and clinical relevance of tumor alterations. *Genome Med.* 10:25. doi: 10.1186/s13073-018-0531-8
- Taylor, A. D., Micheel, C. M., Anderson, I. A., Levy, M. A., and Lovly, C. M. (2016). 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. doi: 10.1016/j.tranon.2016.03.001
- Tran, B., Brown, A. M. K., Bedard, P. L., Winquist, E., Goss, G. D., Hotte, S. J., et al. (2013). Feasibility of real time next generation sequencing of cancer genes linked to drug response: results from a clinical trial. *Int. J. Cancer* 132, 1547–1555. doi: 10.1002/ijc.27817
- Tuchman, M., Morizono, H., Rajagopal, B. S., Plante, R. J., and Allewell, N. M. (1997). Identification of ‘private’ mutations in patients with ornithine transcarbamylase deficiency. *J. Inher. Metab. Dis.* 20, 525–527. doi: 10.1023/A:1005301513465
- U.S. FDA (2018). *Use of Public Human Genetic Variant Databases to Support Clinical Validity for Genetic and Genomic-Based In Vitro Diagnostics: Guidance for Stakeholders and Food and Drug Administration Staff*. Silver Spring, MD: U.S. Food and Drug Administration.
- Van Allen, E. M., Wagle, N., Stojanov, P., Perrin, D. L., Cibulskis, K., Marlow, S., et al. (2014). Whole-exome sequencing and clinical interpretation of formalin-fixed, paraffin-embedded tumor samples to guide precision cancer medicine. *Nat. Med.* 20, 682–688. doi: 10.1038/nm.3559

- Wagner, A. H., Walsh, B., Mayfield, G., Tamborero, D., Sonkin, D., Krysiak, K., et al. (2018). A harmonized meta-knowledgebase of clinical interpretations of cancer genomic variants. *BioRxiv* [Preprint] doi: 10.1101/366856
- Warner, J. L., Dymshyts, D., Reich, C. G., Gurley, M. J., Hochheiser, H., Moldwin, Z. H., et al. (2019). HemOnc: a new standard vocabulary for chemotherapy regimen representation in the OMOP common data model. *J. Biomed. Inform.* 96:103239. doi: 10.1016/j.jbi.2019.103239
- Warner, J. L., Jain, S. K., and Levy, M. A. (2016). Integrating cancer genomic data into electronic health records. *Genome Med.* 8:113.
- Warner, J. L., Prasad, I., Bennett, M., Arniella, M., Beeghly-Fadiel, A., Mandl, K. D., et al. (2018). SMART cancer navigator: a framework for implementing ASCO workshop recommendations to enable precision cancer medicine. *JCO Precis. Oncol.* 2018, 1–14. doi: 10.1200/PO.17.00292
- Whirl-Carrillo, M., McDonagh, E. M., Hebert, J. M., Gong, L., Sangkuhl, K., Thorn, C. F., et al. (2012). Pharmacogenomics knowledge for personalized medicine. *Clin. Pharmacol. Therapeutics* 92, 414–417. doi: 10.1038/clpt.2012.96
- Yang, W., Soares, J., Greninger, P., Edelman, E. J., Lightfoot, H., Forbes, S., et al. (2013). Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Res.* 41, D955–D961. doi: 10.1093/nar/gks1111
- Zehir, A., Benayed, R., Shah, R. H., Syed, A., Middha, S., Kim, H. R., et al. (2017). Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. *Nat. Med.* 23, 703–713. doi: 10.1038/nm.4333
- Zeng, J., Johnson, A., Abu Shufean, M., Kahle, M., Yang, D., Woodman, S. E., et al. (2019). Operationalization of next-generation sequencing and decision support for precision oncology. *JCO Clin. Cancer Inform.* 3, 1–12. doi: 10.1200/CCI.19.00089
- Zhang, J., Walsh, M. F., Wu, G., Edmonson, M. N., Gruber, T. A., Easton, J., et al. (2015). Germline mutations in predisposition genes in pediatric cancer. *N. Engl. J. Med.* 373, 2336–2346. doi: 10.1056/NEJMoa1508054

**Conflict of Interest:** JW is the Deputy Editor of HemOnc.org and co-founder of HemOnc.org LLC; these positions are uncompensated and company shares have no monetary value. JW is a member of the IBM Watson Health Oncology and Genomics Advisory Council. JW receives grant funding from the National Institutes of Health, which also funds several of the knowledge bases mentioned; and AACR Project GENIE.

The remaining author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Li and Warner. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.