Genetic Database Recognition Decision Summary
for
OncoKB

Submitter:	Memorial Sloan Kettering Cancer Center (MSK)
Genetic Database Name:	OncoKB (www.oncokb.org)
Submission Number:	Q191007

1. Summary of FDA Review to Support Recognition

OncoKB is a precision oncology knowledge base developed at Memorial Sloan Kettering (MSK) that collects and stores information on somatic cancer gene alterations. Alterations included in OncoKB are DNA-based, nonsynonymous mutations, rearrangements, insertions and deletions in cancer. This document uses "alterations", "mutations" and "variants" interchangeably.

OncoKB qualifies as a database per FDA's guidance document, "Use of Public Human Genetic Variant Databases to Support Clinical Validity for Genetic and Genomic Based In Vitro Diagnostics."[1] MSK submitted information to support the recognition of the "FDA-Recognized Content" portion of the OncoKB database which lists tumor type-specific somatic alterations and the corresponding FDA level of evidence.[2] This evaluation was based upon whether OncoKB demonstrated conformance with the recommendations described in the FDA's guidance document.

The information submitted included detailed descriptions and standard operating protocols (SOPs) of the oversight and governance procedures for creating, maintaining, and expanding the database and its content within the scope described below, as well as transparency, security, and privacy of such information. FDA evaluated whether these procedures provide reasonable assurance that the variant assertions are accurate and could be used as a source of valid scientific evidence in support of clinical validity of somatic genotyping tests in regulatory submissions. FDA also evaluated the procedures for upkeep and protections for maintenance of the database. Based upon the information evaluated, the FDA determined that OncoKB conforms to the recommendations described in the guidance supporting the recognition of the OncoKB FDA-Recognized Content portion of the database. FDA's review of the information provided is described herein.

Therefore, FDA recognizes the "FDA-Recognized Content" tab within the OncoKB database. This recognition is expected to provide test developers the opportunity to leverage the OncoKB database to support the FDA's regulatory review of a submission for a tumor profiling test[3] and other similar somatic genotyping tests seeking authorization.

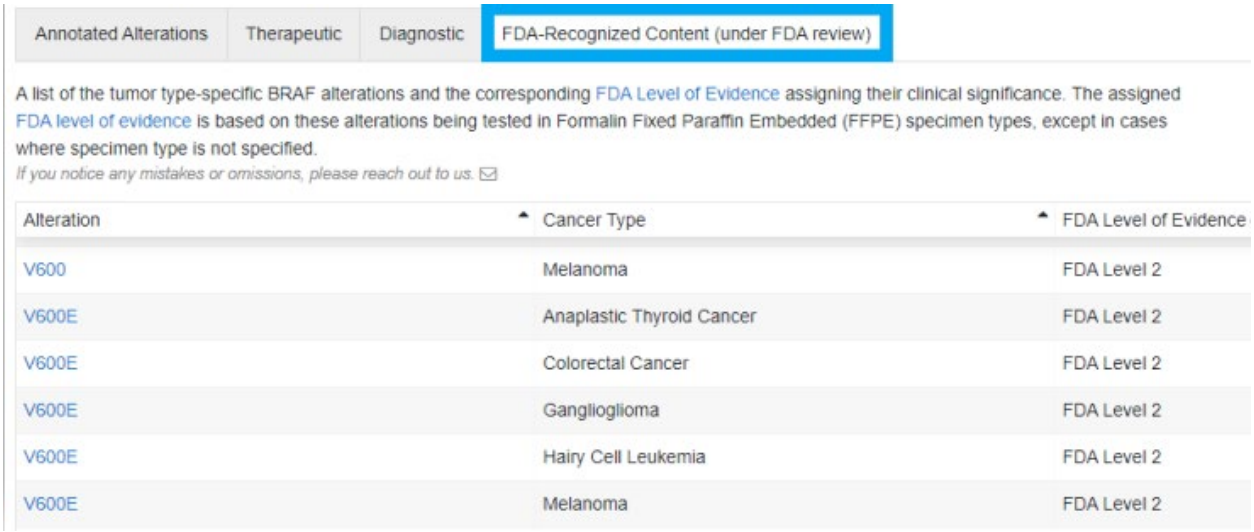
[1] FDA Guidance for Stakeholders and Food and Drug administration Staff available at: https://www.fda.gov/media/99200/download

[2] CDRH'S Approach to Tumor Profiling Next Generation Sequencing Tests available at: https://www.fda.gov/media/109050/download

[3] 21 CFR 866.6080 Next generation sequencing based tumor profiling test; https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/cfrsearch.cfm?fr=866.6080

2. Scope of Recognition

This recognition is for the "FDA-Recognized Content" tab (also referred to as the "FDA Tab") within the OncoKB database which provides listings of the tumor-specific somatic alterations and associated FDA level of evidence when the alteration is detected in a specific cancer type: FDA Level 2 refers to "cancer mutations with evidence of clinical significance" and FDA Level 3 refers to "cancer mutations with potential clinical significance". An example of the "FDA-Recognized Content" tab is shown below:



Alteration	Cancer Type	FDA Level of Evidence
V600	Melanoma	FDA Level 2
V600E	Anaplastic Thyroid Cancer	FDA Level 2
V600E	Colorectal Cancer	FDA Level 2
V600E	Ganglioglioma	FDA Level 2
V600E	Hairy Cell Leukemia	FDA Level 2
V600E	Melanoma	FDA Level 2

To communicate the scope of FDA recognition on the OncoKB website, when an OncoKB user exits the FDA-recognized portion of the website, a pop-up disclaimer appears that 1) temporarily greys out the website, 2) states "You are now leaving the FDA-recognized portion of this page" and 3) requests acknowledgement of this statement by the user via clicking an "OK" button before the user can continue to use the OncoKB website. An image to illustrate how the pop-up is presented on the website is shown below:

2.1. OncoKB Curation Standard Operating Procedure

In support of the recognition of the "FDA-Recognized Content" Tab, MSK provided their master document summarizing the processes and protocols for management of the OncoKB database. This document (Version 2.0, dated March 2021) is available at www.oncokb.org. The SOP is a comprehensive document which summarizes the full scope of the processes for the database and therefore not all content is within the scope of the recognition. The recognition of the OncoKB database is "partial" in that the recognition is of the information provided in support of the "FDA-Recognized Content" Tab. For example, references to protocols and SOPs used to provide information related to biomarker-tumor specific therapeutic options are outside the scope of the recognition. Review of the processes for decisions such as biological

effect/oncogenicity were reviewed in support of decisions related to inclusion or exclusion as a Level 3 variant.

3. OncoKB Oversight and Governance

OncoKB is a precision oncology knowledge base developed and maintained by the Knowledge Systems group in the Marie Josée and Henry R. Kravis Center for Molecular Oncology at the Memorial Sloan Kettering Cancer Center (MSK). OncoKB is publicly available at an interactive website: www.oncokb.org. The OncoKB database collects and stores information on somatic cancer gene alterations including variant assertions and the evidence supporting those assertions that have been fully evaluated by OncoKB staff. Based on the curated evidence, all alterations in OncoKB are assigned an OncoKB Level of Evidence that can be mapped to an FDA Level of Evidence, presented in the FDA Tab within the OncoKB database.

Oversight and Governance of OncoKB is under the purview of the Lead Scientist and the Clinical Genomics Annotation Committee (CGAC) that is comprised of selected members of the scientific and clinical leadership at MSK. A variety of data sources are reviewed by the OncoKB staff which includes public cancer variant databases, alterations identified as statistically significantly recurrent, disease specific treatment guidelines, proceedings of major scientific and clinical conferences, and scientific literature. Variant information is entered and reviewed into the OncoKB curation platform by OncoKB curators and the Scientific Content Management Team (SCMT) respectively. OncoKB staff creates and maintains oversight and governance procedures for the OncoKB staff of individuals with the scientific and clinical expertise to evaluate gene function and disease manifestations, as well as curators who are trained in evaluating evidence sources that support a variant assertion. The OncoKB staff implements robust variant curation and assessment for a single gene or set of genes associated with a single disease/condition or a set of related diseases/conditions in accordance with the procedures described by OncoKB.
The OncoKB staff are fully trained and qualified for their respective positions which consist of the following:

3.1. OncoKB Staff Roles and Responsibilities

The OncoKB staff is a diverse group of scientists, physicians, and engineers. OncoKB has established the minimum qualifications criteria required for their staff members, including educational background, professional training and required skills. Specifics of the experience was provided. An overview of the OncoKB Staff and committees and their roles is described below.

- OncoKB Lead Scientist: Ph.D-level scientist with expertise in translational cancer biology, and clinical cancer genomics with 5 years minimal training. The Lead Scientist creates and maintains general oversight and governance procedures for the OncoKB staff including the development, approval, and coordination of all variant assessment activities. The Lead Scientist also liaises between the variant curation processes and their oversight and governance by CGAC.

- Clinical Genomics Annotation Committee (CGAC): A Clinical Genomics Annotation Committee (CGAC) member is an MD or MD/PhD who is an attending physician at MSK and who is considered an expert in their field and disease specialty. CGAC provides oversight and governance of OncoKB while setting and maintaining standards for the database, especially the assignment of the OncoKB Levels of Evidence to specific alterations. CGAC is responsible for establishing standards and oversight of all processes in the scope of OncoKB. CGAC provides expertise in cancer variant interpretation and the assignment of the OncoKB Levels of Evidence to specific alterations. CGAC is composed of "Core" members and "Extended" members.
    - Extended members are selected physicians and scientists who represent the broader MSK clinical leadership across departments and services, including service chiefs, physicians with clinical expertise in their fields, and scientists with specific gene or pathway expertise.
    - Core CGAC members guide OncoKB development, are at the forefront of clinical management and research, and have translational cancer biology expertise in their respective major disease entities. Core members, in addition to responding to requests regarding clinical consensus, also maintain an active and responsive dialogue with the Lead Scientist, providing insight or updates regarding genomic biomarker-based clinical data.

- The Scientific Content Management Team (SCMT): Two Ph.D-level scientists with translational cancer biology expertise that oversee day-to-day curation and management of OncoKB content and provide guidance and management of the OncoKB Curators regarding appropriate curation, editorial and scientific content review.

- Lead Scientist, Knowledge Systems: Ph.D-level scientist with 5 years minimal training and expertise in computer science, bioinformatics and cancer genomics that creates and maintains general oversight and governance procedures for the OncoKB software engineers.

- Lead Software Engineer: Executes database governance and data preservation as well as feature development and maintenance of the OncoKB Curation Platform (curation platform).

- Data and Software Liaison: Liaises between the Lead Software Engineer and SCMT. The liaison executes computational data analysis, provides computation assistance to the scientific team and works with the software team to implement systems for data curation to ensure seamless data maintenance, updates and access.

- Curators*:* Curators are pre-doctoral graduate students, postdoctoral fellows and clinical fellows. They assess and curate alterations, their biological effects, and associated treatment implications in cancer in compliance with the procedures described by the OncoKB SOP. OncoKB curators are specifically trained in evaluating evidence from various sources, entering appropriate information into the curation platform, variant classification, and the process to map variants into FDA levels.

3.2. Staff qualification and training

OncoKB has established procedures for OncoKB staff initial training, continued education and documentation of training, achievements, deficiencies and competencies which provide a method for OncoKB members to identify individuals or areas of the workflow that may require additional or newly established training.

Through performance reviews required annually, MSK critically assesses and documents the OncoKB staff's training achievements, deficiencies and competencies. Following each evaluation, the reviewer provides each employee with documentation of the assessment outcome, including the employee's strengths, weaknesses and plans for growth and/or improvement. If there is a valid reason to put the employee on probation or terminate his/her position, this decision and a valid reason behind the decision is reviewed and documented.

Procedures for training and ongoing competency assessments for each of the staff and committee positions above was described. All Lead positions are assessed annually, curators, biannually. Curators receive in person training. A description of the extensive training syllabus and protocols were provided and includes the OncoKB curation platform, website content, constructing and understanding graphics, annotations on cBioPortal, performing literature searches, performing searches on external databases and evaluation of various evidence maps and levels of evidence by professional and consensus sources including ASCO-AMP-CAP, ESCAT by ESMO, and FDA product labels. OncoKB curation elements covered in the review include identifying a gene/variant of interest, curation of the variant specific effects and clinical significance. Curators in conjunction with SMCT members also receive training on gene and variant curation for functions outside the scope of the recognition (e.g., biological effect, oncogenic effect and clinical significance). At the end of the initial training the SCMT provides the curator in training (CIT) with two testing worksheets that assess the training process (The Curation Protocol Training Worksheet and The Curation Protocol Proficiency Test). The CIT must complete this test within 1 week. A score of 80% or above on the testing is recommended to initiate a trial curation period. The CIT receives an OncoKB curation assignment to complete within 2 weeks and the curation will be reviewed by a member of the SCMT before being entered into the OncoKB curation platform.

Multiple layers of review occur for the curation of OncoKB variants. OncoKB curators will have variable levels of variant interpretation experience. All staff training achievements, deficiencies and competencies are monitored and documented. The Lead Scientist and SCMT members are responsible for coordinating and monitoring training and proficiency of curators in procuring the appropriate data, assessing the data in the context of variant interpretation, and entering the data with sufficient detail into the OncoKB curation platform. New curators and/or those curators deemed by the Lead Scientist and SCMT members to require additional training are paired with an SCMT member to receive one-on-one training via curation exercises and in person-training sessions.

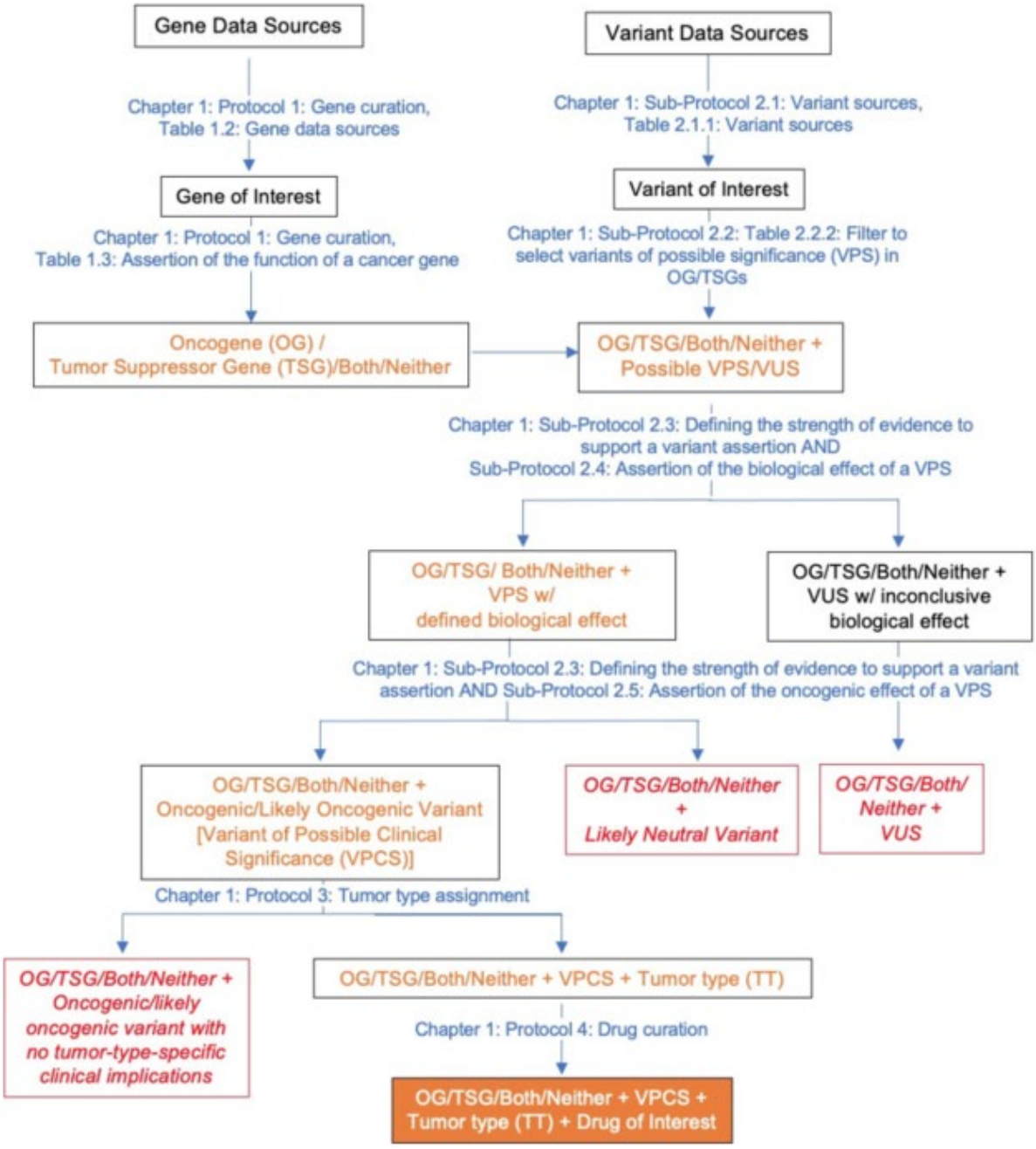
<u>Potential for Conflicts of Interest (COI)</u>
All staff are evaluated for potential COI. Financial conflicts of interest for all OncoKB personnel including CGAC are disclosed publicly on the OncoKB website, www.oncokb.org/team and reported in publications or in conferences as appropriate. In the

event of a conflict of interest arising for a specific CGAC member with regards to a Level of Evidence assignment, he or she is asked to recuse themselves from the consensus request. In the event that consensus cannot be immediately reached, the Lead Scientist is responsible for mediating between conflicting advice to resolve any discrepancy. The Lead Scientist can request the input from the External Advisory Board to resolve conflicting advice from CGAC. Should consensus still not be reached, the proposed change in the Level of Evidence is rejected.

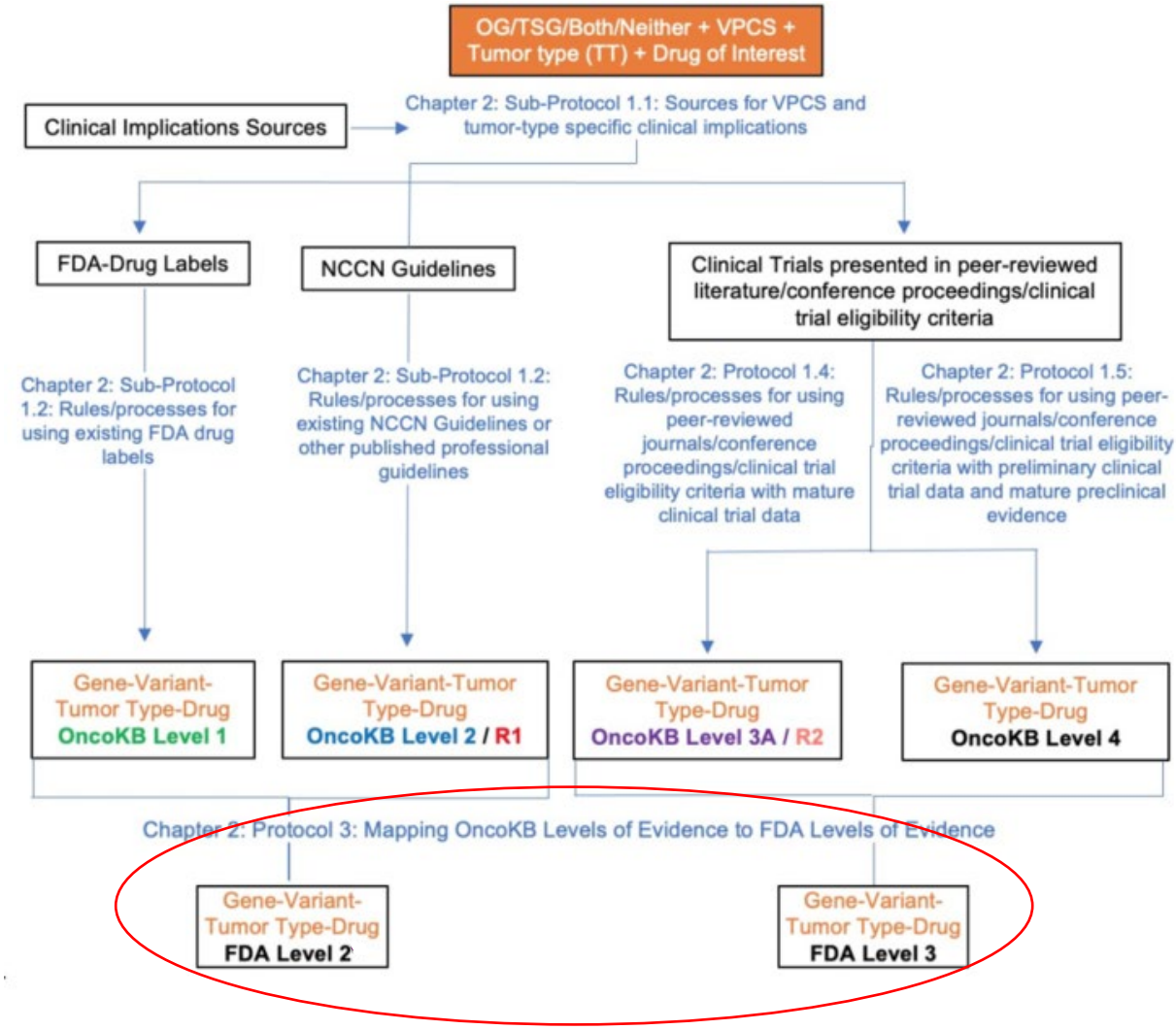
- External Advisory Board (EAB): To help mitigate issues of conflicts of interest (COI), OncoKB has convened an External Advisory Board (EAB) which currently consists of four leaders in the clinical oncology and genomics community external to MSK. As part of the OncoKB EAB, these members have agreed to meet once a year via WebEx to review summarized OncoKB content, comment on any notable process or content changes based on the FDA-approval and clinical trial landscape, assess productivity of the OncoKB team, and advise on improvements to the OncoKB infrastructure, process, or content as necessary. Furthermore, they help mitigate and resolve any COI issues that may arise among members of CGAC.

3.3. Database Curation

Variant information is entered into the OncoKB curation platform, a custom web-based application that allows manual curation and review of variant information. All information entered into the curation platform are structured in a hierarchy of gene, alteration, tumor type and clinical implications. In support of the recognition of the database, FDA reviewed the underlying protocols for how MSK generated the clinical decisions that are used to define the FDA Level 2 or FDA Level 3 mutations to determine that a consistent and well-validated process is established. FDA recognition, however, is not extended to the biological or clinical information outside of the FDA Tab. A flowchart that provides an overview of the OncoKB curation process from gene and variant data sources to FDA level associations is shown below:

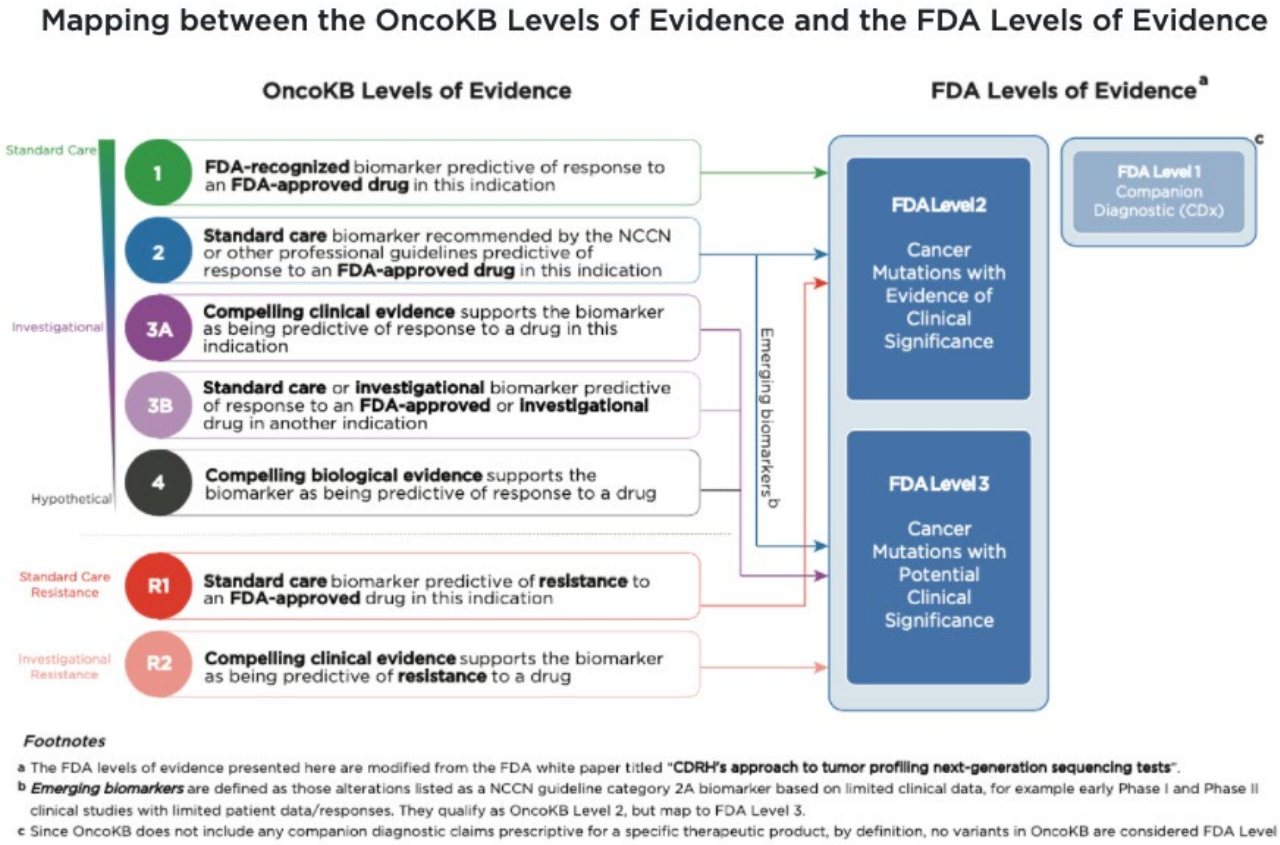


Continued on next page



The process to assign FDA levels of evidence was described in the SOP, Chapter 2, Protocol 3: Mapping OncoKB levels of evidence to FDA levels of evidence as well as integrated to the general variant duration workflow in the SOP, Chapter III: Workflow summaries.

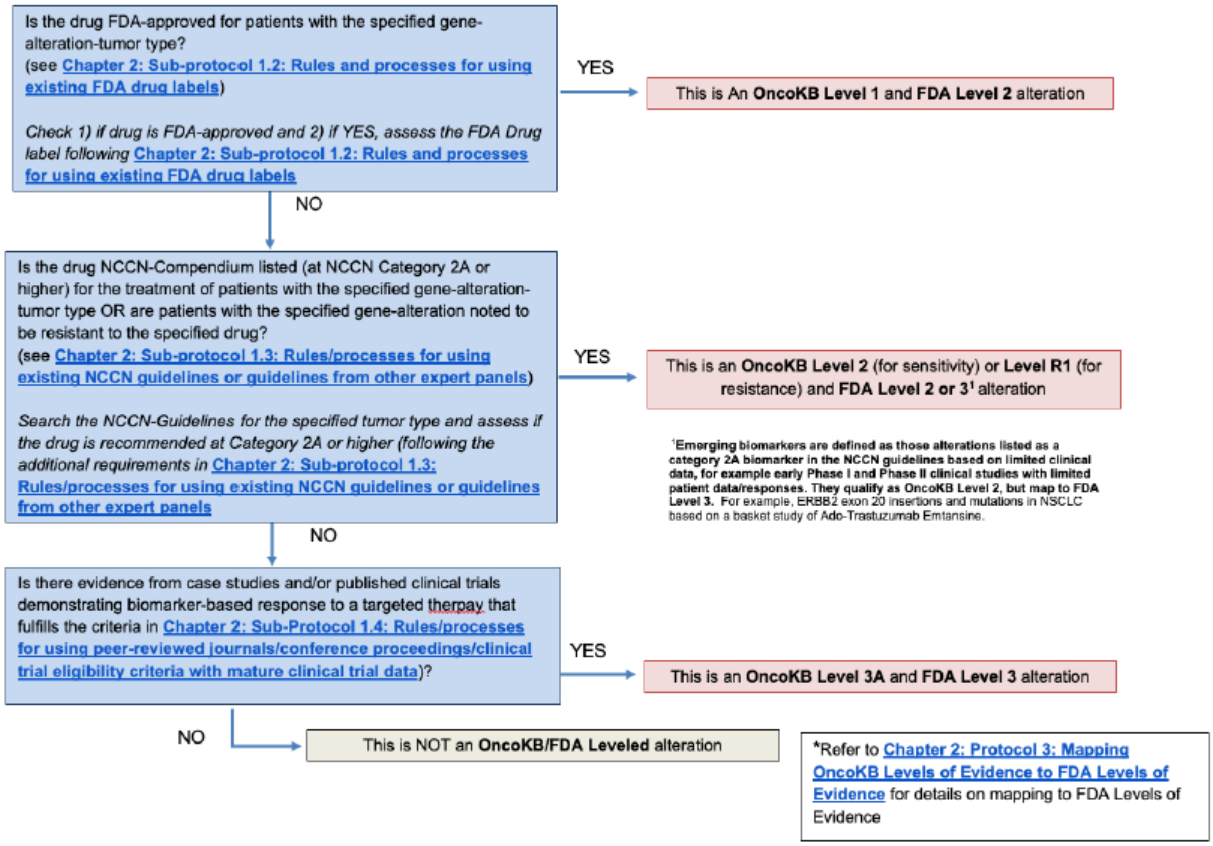
FDA currently has three levels of recognition of the clinical significance of tumor biomarkers for NGS tests for which the Agency has approved somatic variant detection in patients diagnosed with solid neoplasms as described in the FDA fact sheet titled "CDRH's Approach to Tumor Profiling Next Generation Sequencing Tests". Once variants have being assigned an OncoKB level of evidence, the curators follow a specific protocol where qualifying variants are mapped into FDA levels of evidence, and the OncoKB's SOP describe the processes that indicate how every specific OncoKB level of evidence maps to an FDA level of evidence. Since OncoKB is not associated with a companion diagnostic test, OncoKB variants are mapped to FDA levels 2 and 3, and variants in OncoKB are not mapped to FDA Level 1. An example of the mapping between OncoKB Levels of Evidence and FDA Levels of Evidence is shown in the image below.



The OncoKB requirements for variant evaluation are described in OncoKB's Standard Operation Procedure (SOP) which is made publicly available on the OncoKB website, www.oncokb.org.

FDA reviewed the oversight procedures and mapping to FDA levels of evidence to evaluate whether the protocols adequately produce high-confidence assertions and if the information presented in the "FDA-Recognized Content" Tab properly presents the somatic variant information according to the scope of this recognition. SOPs used to provide information related to biomarker-tumor specific therapeutic options are outside the scope of the recognition.

Flow chart to Determine the OncoKB and FDA Level of Evidence for a Specified Variant of Potential Clinical Significance



3.3.1. Variant Assertion and Curation

Protocols for end-to-end curation were provided and included rules and procedures for curating genes, variants, tumor type and other content that are currently not part of the recognition (e.g., Biological function, oncogenic effect, therapies, and clinical implications). Description of the handling and procedures for database entry were also provided. OncoKB uses OncoTree (http://oncotree.mskcc.org) to manage the precise vocabulary of tumor types.

All information was summarized in terms of input and output. The protocol for variant curation specifies the data sources and methods used to determine if a specified gene-variant is a Variant of Possible Significance (VPS). Prior to executing the variant assertion protocol, the protocol for gene curation is executed. Gene curation entailed identification as oncogene (OG), tumor suppressor gene (TSG) or both or neither. Gene curation for ranking priority (high, moderate or low) are based on the specific sources (e.g., External databases, published information, clinical trials, feedback from users). The INPUT of this variant assertion protocol must be a gene defined as an OG, TSG, Both or Neither. Assertions for gene curation follow pre-specified rules and is shown in the Table below.

Evidence	ASSERTIONS		
	Oncogene (OG)	**Tumor Suppressor (TSG)**	**Both**
I. Weinberg, p.G:20, 2014 Vogelstein et al., 2013	***RULE OG-1*** Any of the following features as demonstrated by the scientific literature in ≥1 study. (1) A cancer-inducing gene when activated by mutation OR (2) A gene that can transform cells by increasing the selective growth advantage of the cell in which it resides as demonstrated by the scientific literature in ≥1 study.	***RULE TSG-1*** Any of the following features as demonstrated by the scientific literature in ≥1 study. (1) A gene whose partial or complete inactivation by mutation, occurring in either the germline or the genome of a somatic cell, leads to an increased likelihood of cancer development by increasing the selective growth advantage of the cell in which it resides OR (2) A gene that is responsible for constraining cell proliferation OR (3) A gatekeeper, a gene that operates to hinder cell multiplication or to further cell differentiation or cell death and in this way prevents the appearance of populations of neoplastic cells 4) Mutated through protein-truncating alterations throughout their length	***RULE TSGOG-1*** Meets at least one of the criteria for both OG and TSG
II. Davoli et al., 2013	***RULE OG-2*** A gene that, in tumor samples, has i) higher functional impact as defined by the PolyPhen2 Hum-Var prediction model and higher amplification frequency in comparison to those observed in neutral genes, AND ii) lower loss-of-function mutations, splicing mutations and frequency of deletions and increased frequency of amplification compared to tumor suppressors	***RULE TSG-2*** A gene that, in tumor samples, has i) higher frequencies of loss-of-function and splicing mutations, higher functional impact, and higher frequency of deletions compared to those found in neutral genes, AND ii) higher frequencies of loss-of-function and splicing mutations, higher deletion frequency and lower amplification frequency compared to those found in oncogenes	***RULE TSGOG-2*** Meets OG AND TSG criteria

A high-level overview of the subprotocols for variant curation is shown in the table below. A list of Variant data sources was also provided as well as the frequency of assessment of the data sources by the OncoKB team. As noted above, data is further reviewed by OncoKB staff (i.e., never imported directly).

Step	INPUT	INPUT to OUTPUT Process Location		OUTPUT
		Protocols (from Chapter 1)	**Table (if applicable; from Chapter 1)**	
1	Variant data sources	Sub-Protocol 2.1: Variant sources	Table 2.1.1 Variant data sources	Variant of Interest
2	Gene defined as OG/TSG/Both/Neither (from Chapter 1: Protocol 1: Gene curation) AND Variant of Interest	Sub-Protocol 2.2: Defining Variant Type	Table 2.2.1 Definitions of variant types and their molecular consequences AND Table 2.2.2 Filter to select Variants of Possible Significance (VPS) in OG/TSGs	Candidate Variant of Possible Significance (VPS)/Variant of Uncertain Significance (VUS)
3	Gene defined as OG/TSG/Both/Neither AND Candidate VPS/VUS	Sub-Protocol 2.3: Defining the type and strength of evidence to support a variant assertion	Table 2.3.1 Types of experimental evidence to support VPS biological or oncogenic assertion	Gene defined as OG/TSG/Both/Neither AND Candidate VPS/VUS with defined biological effect OR Candidate VUS with Inconclusive biological effect
			Table 2.3.2 Definition of the strength of functional (experimental) evidence	
		Sub-Protocol 2.4: Assertion of the biological effect of a VPS	NA	
4	Gene defined as OG/TSG/Both/Neither AND Candidate VPS/VUS with defined biological effect	Sub-Protocol 2.3: Defining the type and strength of evidence to support a variant assertion	Table 2.3.1 Types of experimental evidence to support VPS biological or oncogenic assertion Table 2.3.2 Definition of the strength of functional (experimental) evidence	Oncogenic Variant with defined biological effect = Variant of Possible Clinical Significance (VPCS) OR Likely Neutral Variant with defined biological effect == Likely Neutral Variant1 OR Variant with Inconclusive biological and oncogenic effect == VUS
		Sub-Protocol 2.5: Assertion of the oncogenic effect of a VPS	NA	

Rules and Processes for the type and strength of evidence to support a variant assertion were provided in detail for functional evidence, in silico evidence and preclinical evidence. This evidence is not part of the recognition and therefore is not described in this decision summary. Examples for all sources of evidence and their curation were provided.

3.3.2. Process for assignment of FDA levels of evidence:

3.3.2.1. FDA drug labels:

The process to assign FDA levels of evidence to a variant is described by mapping OncoKB levels of evidence to FDA levels of evidence and integrating this process into the general variant curation workflow. Genetic alterations specified in the FDA drug label or other professional guidelines that may qualify as Variant(s) of Potential Clinical Significance (VPCS) i.e., FDA Level 2 (OncoKB Level 1 or 2) variants, were described. When evaluating the potential FDA Level 2 (OncoKB Level 1 or R1), protocols were provided describing the decision trees for evaluating and interpreting the different sections of the FDA drug label. Level 2 evidence includes Section 1 Indications and Usage, Section 2.1 Patient Selection, Section 12.1 Mechanism of Action and Section 14: Clinical Studies. The different sections of the FDA drug label, the priority/weight assigned to the information in each section, the specific information that is assessed and the rules for determining the FDA Level 2 association is shown in the table below.

FDA drug label section	**Priority/ weight when defining an FDA Level 2 (OncoKB Level 1 or R1) VPCS[1]**	**Information in the FDA drug label that is assessed by OncoKB**	**Rules for determining if the INPUT gene-VPCS- tumor type-drug qualifies as an FDA Level 2 (OncoKB Level 1 or R1) association[2] (per Chapter 2: Sub-protocol 1.2: Rules and processes for using existing FDA drug labels)**	
			Criteria that must be met from the FDA drug label sections	**The FDA Level 2 (OncoKB Level 1 or R1) association**
Section 1: Indications and Usage	High	● Gene ● Alteration ● Tumor Type ● Drug ● Does the section specify "as detected by an FDA-approved test"	If the INPUT VPCS is specifically listed in *Section 1: Indications and Usage* of the FDA drug label AND Patient selection is NOT determined by an FDA-approved test (CDx) (per *Section 2.1: Patient Selection* of the FDA drug label)	

FDA drug label section	Priority/ weight when defining an FDA Level 2 (OncoKB Level 1 or R1) VPCS[1]	Information in the FDA drug label that is assessed by OncoKB	Rules for determining if the INPUT gene-VPCS- tumor type-drug qualifies as an FDA Level 2 (OncoKB Level 1 or R1) association[2] (per Chapter 2: Sub-protocol 1.2: Rules and processes for using existing FDA drug labels)	
			Criteria that must be met from the FDA drug label sections	**The FDA Level 2 (OncoKB Level 1 or R1) association**
Section 2.1: Patient Selection	High	● Does the section specify "as detected by an FDA-approved test" ● If YES - proceed to http://www.fda.gov/CompanionDiagnostics	If *Section 2.1: Patient Selection* of the FDA drug label specifies that patient selection must be determined by an FDA-approved test (CDx test) AND the INPUT VPCS is specifically listed in the corresponding CDx test	The INPUT gene-VPCS-tumor type-drug qualifies as an FDA Level 2 (OncoKB Level 1) association
www.FDA.gov/ Companion Diagnostics	High	● Gene ● Alteration(s) ● Tumor Type ● Specimen Type ● For a specified CDx test, the specific sections that require review are: 1. Premarket Approval (PMA) 2. Approval Order 3. Labeling		
Section 14: Clinical Studies	Moderate	● Clinical Trial Details and Metrics:	If patient selection is NOT determined by an FDA-approved test (CDx test) per	

FDA drug label section	Priority/ weight when defining an FDA Level 2 (OncoKB Level 1 or R1) VPCS[1]	Information in the FDA drug label that is assessed by OncoKB	Rules for determining if the INPUT gene-VPCS- tumor type-drug qualifies as an FDA Level 2 (OncoKB Level 1 or R1) association[2] (per Chapter 2: Sub-protocol 1.2: Rules and processes for using existing FDA drug labels)	
			Criteria that must be met from the FDA drug label sections	**The FDA Level 2 (OncoKB Level 1 or R1) association**
		○ Phase ○ Drug ○ Tumor type ○ Total Number of patients ○ Patient cohort stratification ○ Biomarker-based eligibility criteria ○ Primary and Secondary outcomes ○ Efficacy Results (for biomarker-based cohort)	*Section 2.1: Patient Selection* of the FDA drug label AND the INPUT VPCS is included under an umbrella term listed in *Section 1: Indications and Usage* of the FDA drug label AND the INPUT VPCS is specified as being tested in the referenced clinical trial in *Section 14.1: Clinical Studies*	
Section 12.1: Mechanism of Action	High	● Gene ● Alteration ● Mention of clinically acquired resistance mutation	If the INPUT association is being evaluated in the context of resistance AND *Section 12.1: Mechanism of Action* of the FDA drug label specifies the VPCS is a clinically acquired resistance mutation	The INPUT gene-VPCS-tumor type-drug qualifies as an FDA Level 2 (OncoKB Level R1) association

[1] *Section 1: Indications and Usage* and *Section 2.1: Patient Selection* of the FDA drug label should be assessed simultaneously and the variants they reference should be directly compared.

3.3.2.2. Professional Guidelines

Protocols for using external guidelines such as NCCN guidelines or other professional guidelines when determining which variants to designate as Level 2 was also provided. Examples of how to define genetic alterations specified in Section 1: Indications and Usage of the FDA drug label or in the NCCN or other professional guidelines when the terminology in the data source is vague (including when umbrella terms are used) was also described.

3.3.2.3. Emerging biomarkers

Emerging biomarkers are defined as those alterations listed as a category 2A biomarker in the NCCN guidelines based on limited clinical data, for example early Phase I and Phase II clinical studies with limited patient data/responses. They qualify as OncoKB Level 2, but map to FDA Level 3.

3.3.2.4. Peer-reviewed journals/conference proceedings/clinical trial eligibility criteria with mature clinical trial data

Rules/processes for using peer-reviewed journals/conference proceedings/clinical trial eligibility criteria with mature clinical trial data were provided to determine whether mutations qualify as Mutations with Potential Clinical Benefit (Level 3). The protocol describes the process for determining FDA Level 3 (OncoKB Level 3A or R2) associations. The protocol specifically details the approach for evaluating and interpreting peer-reviewed journals, conference proceedings and clinical trial eligibility criteria with mature clinical data. Input includes evaluating location of the mutation relative to the functional domain (e.g., DNA binding domain or kinase domain), the number of patients with the specific tumor type with published evidence of a RECIST-defined clinical response, or trial defined clinical benefit, and robust biological studies about the effect of the mutation in the ability to sensitize the cancer cells to the drug of interest. The level of preliminary clinical data and mature preclinical evidence is considered in the assessment.

3.3.2.5. Biomarker-based clinical studies

The types of studies evaluated by OncoKB members when assessing the strength and validity of clinical evidence and determining whether data presented from clinical trials qualifies for an FDA Level of Evidence was included.

Example of the clinical data that an OncoKB curator or SCMT member must assess and extract when evaluating evidence from peer-reviewed, published biomarker-based clinical studies is shown below. Once collected, the data is summarized and reviewed to determine if the VPCS qualifies for an FDA and OncoKB Level of Evidence.
Each bullet below represents a column in the Table that is filled in by the OncoKB curator or SCMT member.

To comprehensively curate the clinical data from biomarker based clinical studies Fifty-two data points (listed below) are used to document the following information per study:

- Gene
- Alteration
- Tumor type
- Drugs
- OncoKB Level of Evidence
- References
- Other relevant drugs (in the same drug family)
- Number of studies with clinical data
- Reference study (PMID or Abstract)

- PMID or abstract of additional studies with clinical data (non-reference study)
- Notes on additional studies (non-reference study)
- Reference study type e.g. Basket Study
- Reference study drug
- Trial Name/ID e.g. NCT01226316
- Phase
- Disease
- Setting
- Total number of patients (N)
- Number of patients who responded (n)
- Primary endpoint
- Notes on primary endpoint
- Secondary endpoint
- Notes on secondary endpoint e
- PFS (experimental group)
- 95% CI (experimental group)
- PFS (control group)
- 95% CI (control group)
- PFS gain
- PFS HR
- OS (experimental group)
- 95% CI (experimental group)
- OS (control group
- 95% CI (control group)
- OS gain
- OS HR
- ORR
- Clinical benefit rate
- CR
- PR
- SD
- PD
- Not evaluable
- DOR
- If case study, describe response
- Quality of life e
- Toxicity: No. (%) of Grade ≥ 3 Adverse Events e.
- Notes on toxicity
- Number or preclinical studies e
- Preclinical study PMID or abstract
- Preclinical data summary
- Summary of data

3.3.3. Tumor Type Assignment:

A protocol specifying how tumor types are assigned when a variant of possible clinical significance (VPCS) is associated with tumor type-specific clinical implications was reviewed. Curation of tumor types for OncoKB utilize the nomenclature found in OncoTree (http://oncotree.info) to describe tumor types as a subtype of a specific tumor maintype (Kundra et al., JCO Clinical Cancer and Informatics, 2021) as outlined in Chapter 1: Figure 3: OncoTree Homepage and tree structure. OncoTree (http://oncotree.info) is a cancer classification system that was developed and is updated by a cross-institutional committee of oncologists, pathologists, and scientists and is accessible via an open-source web user interface and an application programming interface (API). All cancer types and subclassifications are represented through a taxonomic tree/branch system based on the cell of origin and histologic architecture. This structure of the tree not only allows grouping of tumor types under the tissue of origin but also connecting nodes across branches based on histology

3.3.4. OncoKB Data Sources

OncoKB provided rules and procedures for using various data sources in the curation of variants. Four primary data sources are used to identify and curate cancer variants and their biological and clinical therapeutic implications:

- Public cancer variant databases of alterations identified in tumor sequencing studies,
- Statistically significant and recurrent variants identified based on 24,592 sequenced tumors using methods described in Chang et al., 2018.
- Disease-specific treatment guidelines such as those provided by the National Cancer Compendium Network (NCCN) and proceedings of major scientific and/or clinical conferences such as the American Society of Clinical Oncology (ASCO) and the American Association of Cancer Research (AACR).
- Manual review of general scientific literature accessed through PubMed.

These databases are not used as primary curation sources but are used for variant candidate selection by downloading the comprehensive list of alterations in each database and comparing them to the mutations curated in OncoKB. Post candidacy, each variant is independently curated using the processes specified in and undergo necessary review. reanalysis, and re-review as needed.

Data sources from which information is reviewed and critically assessed when assigning FDA Level of Evidence are shown below.

Data source type that contains evidence for a leveled association	Data source example or clarification		FDA Level of Evidence
FDA Drug Label	Specific sections of the FDA drug label to investigate are: Section 1: Indications and Usage Section 2.1: Patient Selection Section 14: Clinical Section 12.1: Mechanism of Action		2
NCCN Guidelines	www.nccn.org		2 or 3
Peer Reviewed Journals [2]See Chapter 2: Table 1.4.1: Types of biomarker-based studies or analyses evaluated by OncoKB	Cell Cancer Discovery Medicine JAMA Oncology Nature Nature Medicine Nature Reviews Clinical Oncology Journal of Clinical Investigation Lancet Oncology Nature Reviews Cancer Cancer Cell Annals of Oncology Clinical Cancer Research Cancer Research	JAMA New England Journal of Science Science Translational Medicine JCO JCO PO J Thoracic Oncol Target Oncol Lung Cancer BMC Cancer Haematologica Leukemia Hematology	3
Conference Proceedings (Abstracts, Posters or Presentations)	AACR Annual Meeting ASCO Annual Meeting ESMO Annual Meeting ASH Annual Meeting	IASLC WCLC SABCS AACR-EORTC-NIH MTCT	
Clinical Trial Eligibility Criteria	Biomarkers must be specified in patient inclusion or exclusion criteria		

3.3.5. Approval of Assignment of Level of Evidence/Data Validation:

Data validation is required to check all internally, independently reviewed OncoKB curated data. Data curated in the OncoKB curation platform is not publicly available [on cBioPortal for Cancer Genomics (www.cbioportal.org) or the OncoKB public website (www.OncoKB.org)] until it is internally reviewed by a member of the OncoKB staff. Internal, independent review of curated data is performed in the OncoKB curation platform Review Mode. Review Mode details all changes made in a specified Gene Page since the time of the last review. Specific additions/deletions/edits are highlighted to designate the specific text or entries that have been added, deleted or removed since the time of the last review. Review Mode also notes the name of the user who made the data changes and the

date/time of the data entry/removal. A reviewer may not accept his/her own changes in Review Mode and must ask another member of the SCMT or the Lead Scientist to review this data. All curated data is internally reviewed by an OncoKB staff member who did not themselves curate the data.

Prior to internal review, all proposed OncoKB/FDA leveled associations must be reviewed and approved by CGAC. CGAC members are responsible for entering into consensus regarding the assignment of an OncoKB level of evidence to a biomarker. Requests for consensus from CGAC occur in the form of emails from the Lead Scientist to all CGAC members and are typically prompted by new FDA-approvals, FDA-breakthrough designations, or newly reported results of major clinical trials from clinical oncology conferences or publications. An example of the components of the consensus email were provided and the description of the rationale for the proposed level of evidence is included. The format of the email includes the recipients, deadline for response, proposed level of evidence, proposed change to the level as applicable, reference links, and evidence/clinical summaries,

A validation protocol that assesses the consistency of variant classification to FDA levels of evidence is provided for the purpose of assessment of consistency of variant classification to OncoKB and FDA levels of evidence. An assessment of the effectiveness of the protocols is also presented. Mapping OncoKB Levels of Evidence to FDA Levels of Evidence ranges from 85.7% to 100%.

Data validation is required to check all internally, independently reviewed OncoKB curated data for errors before release to the OncoKB public website (www.oncokb.org). An automated data validation tool is built into the Tools Page on the OncoKB curation platform. By clicking the 'Data Validation' button, the tool queries all curated data (that has been reviewed per the protocol for Data review) and returns database elements that do not pass the data validation test questions outlined in the protocol. These elements are separated into two sections, or "tabs", in the data validation tool.

3.3.6. Conflicting data and conflicting assertions:

Detailed rules, protocols and workflows with regards to OncoKB's requirements for variant assertion as well as resolving conflicting data and conflicting assertions were provided and reviewed. The baseline curation follows an internal process of independent review of curated data performed by an OncoKB staff member who did not themselves curate the data. If conflicting data or conflicting assertions arise during data curation or review, these are resolved following specific protocols and workflows outlined in OncoKB's SOP where an independent review of curated data is performed, and the decisions are evaluated and discussed to reach consensus. A process to evaluate evidence when there is disagreement between different types (experimental vs. clinical) of evidence as well as how to evaluate and resolve conflicting information presented in different publications was provided. OncoKB's SOP also has mechanisms in place to approach instances where drug labels and companion diagnostic labels may present variant information differently in order to clearly determine what areas of the drug label will be

reviewed, the weight assigned to information in different sections, and to define variants in the FDA drug labels or other professional guidelines when non-specific language is used. The resolution entails prioritizing different data sources and their strength, and review by disease specific experts. In cases where consensus cannot be reached, the alteration is not assigned a level of evidence. In cases where majority consensus is reached, the alteration is accepted into OncoKB with note that assertion is a result of majority and not uniform consensus.

3.3.7. Re-analysis and Re-evaluation

OncoKB data continuously undergoes re-analysis and re-evaluation in order to keep the database and SOP procedures current with updated FDA approvals, professional guidelines and data sources. Variant assertions are re-analyzed and re-evaluated by the OncoKB team in specific review cycles and any new content or inconsistencies are corrected at that time. The process for re-evaluating and re-assigning (if applicable) the biological effect of an existing FDA Level 3 variant to an FDA Level 2 variant in OncoKB when new evidence becomes available was provided. The process for variant re-analysis and re-evaluation is initiated by an OncoKB curator (under the management and direction of a SCMT member) following the Variant curation and Data review protocols. The variant's existing biological effect and the validity and strength of the new information is considered during re-analysis and re-evaluation. If new evidence supports the current functional designation of the Variant of Possible Significance (VPS), the VPS's biological effect remains the same, but the reference and data associated with the new evidence is added to the curation system. References for all new evidence are incorporated into the OncoKB curation system as outlined per the protocol for OncoKB alteration nomenclature, style and formatting and data is added to the mutation effect description as outlined per the protocol Generation and formatting of mutation effect description. The stepwise procedure is shown below and is an example of other stepwise procedures reviewed in support of the curation process:

INPUT: A. **Gene** defined as *Oncogene* or *Tumor Suppressor* or *Both* or *Neither* **+** B. **Variant** must be defined as a *Variants of Possible Clinical Significance (VPCS)* as outlined in **Variant curation** C. **Tumor Type** must correspond to a tumor type in OncoTree as indicated in **Tumor type assignment** D. **Drug:** must be a targeted therapy (refer to **Drug curation**)
1. Identify a **data source** that contains evidence to support changing an existing leveled clinical implication (including FDA and/or OncoKB leveled association) Refer to **Procedure for evaluating data sources that may result in a change in an FDA or OncoKB Level of Evidence (column II)** for an overview of data sources that may prompt a change in the FDA and/or OncoKB Level of Evidence of an existing leveled clinical implication in OncoKB a. *Proceed to Step 2*

2. Note the **pre-existing OncoKB curated data** for the specified clinical implication, including the: 1) gene, variant, tumor-type and drug of interest, 2) current OncoKB Level of Evidence, 3) current FDA Level of Evidence, and 4) current referenced data sources and source types (e.g., FDA drug label for capmatinib)

    a. *Proceed to Step 3*

3. **Critically assess the evidence** in the data source identified in Step 1 by following the process outlined in **Procedure for evaluating data sources that may result in a change in an FDA or OncoKB Level of Evidence.** Should the pre-existing clinical implication be assigned a new FDA and/or OncoKB Level of Evidence?

    a. **YES:** *Proceed to:*

        i. **Rules/processes for using existing FDA drug labels** to assess the data for a potential FDA Level 2 (OncoKB Level 1 or R1) association OR

        ii. **Rules/processes for using existing NCCN guidelines or guidelines from other expert panels** to assess the data for a potential FDA Level 2 (OncoKB Level 2, 3A or R1) association OR

        iii. **Rules/processes for using peer-reviewed journals/conference proceedings/clinical trial eligibility criteria with mature clinical trial data** to assess the data for a potential FDA Level 3 (OncoKB Level 3Aor R2) association OR

        iii. **Rules/processes for using peer-reviewed journals/conference proceedings/clinical trial eligibility criteria with preliminary clinical trial data and mature preclinical evidence** to assess the data for a potential FDALevel 3 (OncoKB Level 4) association

    b. **NO**: No further action (curation) is necessary. Exit the protocol.

4. Follow Chapter 2: Protocol 2: CGAC approval of OncoKB level of evidence assignment to **CGAC review and consensus for the proposed FDA and/or OncoKB Level of Evidence change to obtain** CGAC review and consensus for the proposed FDA and/or OncoKB Level of Evidence change.

The SCMT maintains current variant interpretations and level of evidence assignments by performing the below procedures:

- Addressing all inquiries/and or new evidence submitted by public users and/or members of the MSK community within 72 hours of the inquiry. This may involve assessing new evidence for previously curated variant or levels of evidence as well as novel variants or levels of evidence (not already in OncoKB)
- Incorporating new data from data sources within 12 months of their publication
- Reassessing all variants classified as Variants of Unknown Significance or inconclusive at least every two years

OncoKB staff maintain accuracy and currency of OncoKB curated variants and levels of evidence by performing the below procedures:

- Evaluation of new data sources and evidence to determine whether to add newly curated variants or change the evidence of a previously curated variant
- Re-analysis and re-evaluation of existing evidence to determine whether a change in the level of evidence is warranted

Procedures including training and training documentation for the implementation of major changes in the OncoKB SOP were provided as well as the transparency of such changes. (Additional training modules are required for an established OncoKB curator to qualify as an SCMT member.)

Additionally, feedback regarding updated content or inconsistencies reported from users of OncoKB are addressed within 72 hours of receipt. Procedures to address conflict of interest for specific CGA members with regards to levels of evidence are described in the SOP, Chapter 2, Protocol 2: CGAC approval of OncoKB leveled associations

3.3.8. Conflicts of Interest

To address and resolve potential Conflicts of interest, it is required that any new level assignments or changes to an existing level should be approved unanimously by all CGAC members and there are at minimum 3 affirmative verifications from CGAC. The affirmative verifications from CGAC that must be received for a proposed change to the levels of evidence to be entered into OncoKB are the following:

- From the Director of the Center for Molecular Oncology
- From a Disease Management Team Chief in the indication of the proposed level of evidence change
- A miscellaneous member of CGAC

Members of CGAC who may have Conflicts of interest with respect to the introduction or change of the levels of evidence assigned to a specific variant are allowed to provide advice and information regarding the assertion but are excluded from the 3 CGAC member verification committee.

3.3.9. Mechanism of assertion feedback

Assertion feedback by OncoKB users is an important feature of the knowledge base. There are two web-based mechanisms through which users may provide feedback on OncoKB content: 1)The OncoKB website (2) and the cBioPortal for Cancer Genomics.
Feedback, comments or questions may be sent via email to contact@oncokb.org, which is provided in multiple places within the OncoKB website. Emails sent to contact@oncokb.org are received by the Lead Scientist and all SCMT members and answered within 72 hours.

In cBioPortal, variants in both the patient view and mutations tab are annotated with OncoKB information. Users may either click the OncoKB icon to access the OncoKB webpage to provide feedback or click the Feedback button in the OncoKB dialog box. In the "OncoKB Annotation Feedback" pop-up form (B, i), information about the Gene and Alteration, the email address used to log-into the portal, and web-address of the specific portal instance will be pre-populated. Users may then enter specific feedback and associated references in the Feedback and References fields before submitting the feedback.

Submission of feedback by a cBioPortal user will auto-populate in a Google spreadsheet (B, ii). Changes to this Google Sheet will trigger an automatic email sent to the Lead Scientist and SCMT alerting them of user feedback via cBioPortal. User feedback is answered within 72 hours of its receipt. Upon completion of any necessary deliverables as suggested by the feedback (either curation or software related), the appropriate OncoKB staff member fills in the "Complete" column and adds their initials as well as any comments related to the feedback item. The Feedback Page collates all cBioPortal user feedback related to OncoKB assertions and is a log of OncoKB development based on cBioPortal user-feedback.

4. Transparency and Public Accessibility

4.1. OncoKB Access

Data from OncoKB is available four ways. The FDA Recognized Content Tab is available from within the database:

- OncoKB data is publicly available for personal and research purposes through an interactive website at www.oncokb.org. Usage terms of OncoKB are specified at https://www.oncokb.org/terms.
- The curated data is also available programmatically through the OncoKB application program interface (API). The different ways to access OncoKB data are documented at www.oncokb.org/DataAccess .
- The cBioPortal for Cancer Genomics (https://www.cbioportal.org) uses the OncoKB API for annotating cancer variants in its database.
- OncoKB data is used to annotate the patient reports of the results from MSK-IMPACT, a targeted tumor sequencing test available to MSK patients.

Additionally, a version controlled OncoKB SOP v2 describing all processes and protocols involved in the maintenance of OncoKB, is publicly available on the web.

All versions of the OncoKB SOP are publicly available on the "About" page of the OncoKB website (oncokb.org/about). The SOP presents information regarding all the rules, protocols and procedures established for the OncoKB database, including data sources, variant assertion and mapping to FDA levels of evidence. All final variant assertions and their respective FDA level of evidence are made publicly available through the OncoKB website.

All changes to OncoKB in a given data release are specifically documented on the OncoKB News page (oncokb.org/news). Each News item and the corresponding data release is dated, and version controlled. Access to previous versions of OncoKB are also provided via GitHub.

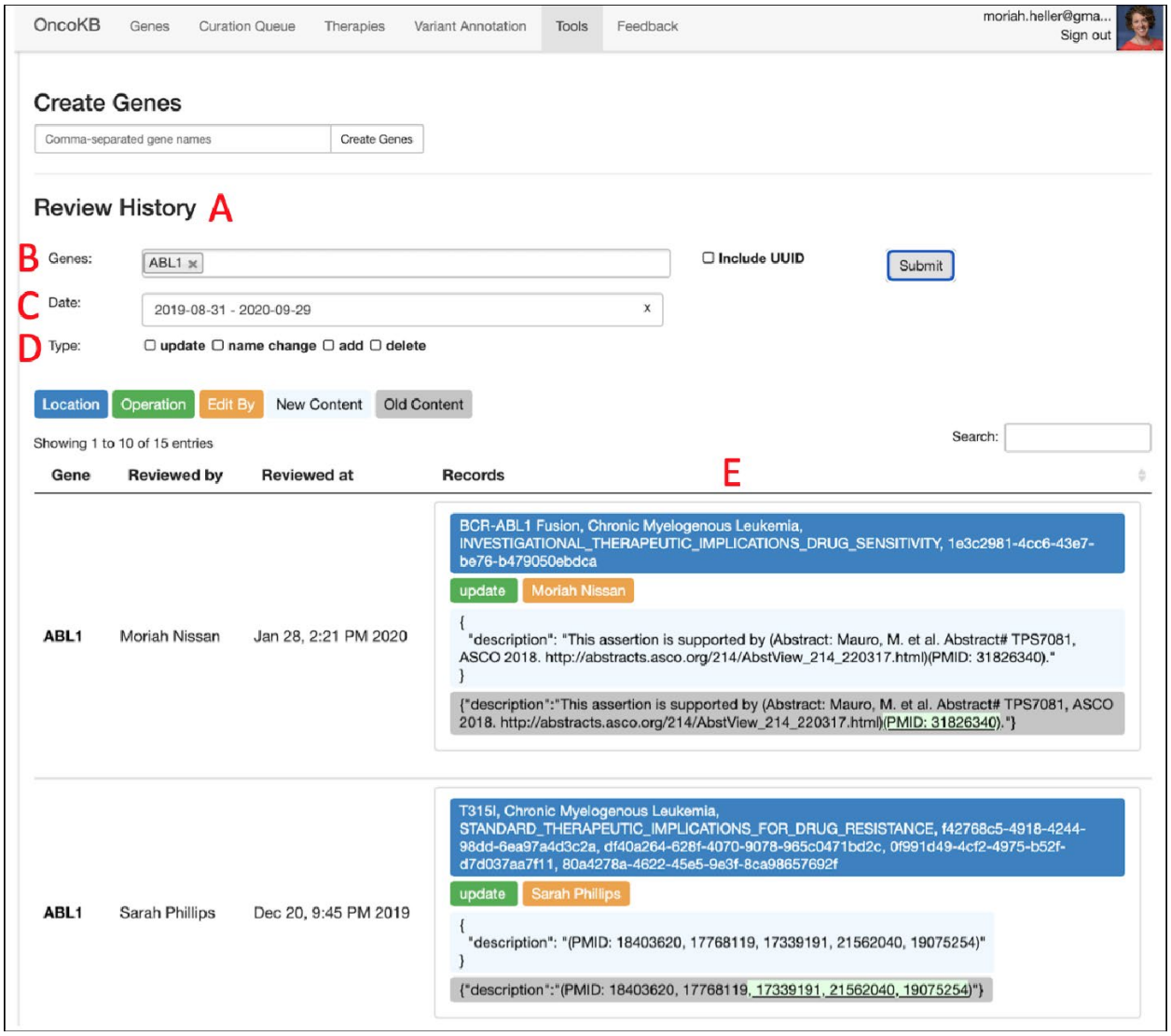
An in-depth protocol about OncoKB's news release is described in OncoKB's Standard Operation Procedure (SOP) which is made publicly available on the OncoKB website, www.oncokb.org.

4.2. Standard Operating Procedure (SOP) Version Control

All SOPs that govern processes for how FDA levels of evidence are developed and approved by OncoKB as well as all information regarding variant curation, validation, staff qualification and training, are presented in OncoKB's Master SOP.

OncoKB operations are reviewed on a weekly basis. The OncoKB SOP is modified as needed and fully reviewed on an annual basis. OncoKB has protocols in place to introduce significant changes to the SOP and to evaluate if such changes merit the re-evaluation of variants that have been previously curated.

Within the Tools page is Review History. All reviewed changes to an indicated gene (those listed in: Data additions, deletions and edits highlighted in Review Mode in the OncoKB curation platform) within a designated date range can be visualized by selecting the dates in the dropdown; alternatively, only changes of a certain type (e.g updates, name change, etc) can be selected using the type checkboxes. Example results retrieved from this query are shown in in the figure below. Review History highlights the difference from the pre-reviewed version as well as the user who initiated the change, the SCMT member who reviewed and accepted the change, and the date the change was reviewed



4.3. Data Preservation and Security

All data connected to FDA levels of evidence within the "FDA-recognized Content" tab on the OncoKB database, are publicly available via the OncoKB website at www.oncokb.org. Data finalized for OncoKB inclusion is stored in a MySQL database, ensuring proper data links and data integrity. Processes such as unit tests are in place to assess overall database consistency and stability.

OncoKB releases its data on a monthly basis through www.oncokb.org. Between data releases, processes are in place to perform daily backup of the OncoKB information to safeguard against system errors and to allow for reinstatement of OncoKB as necessary. At the time of monthly data releases, software updates are simultaneously released to a repository in GitHub.

The OncoKB database is publicly available through an interactive website that can be accessed through the OncoKB application program interface (API). OncoKB has provided information

about its website security protocols with regards to vulnerability to security breaches that could potentially impact the accessibility and integrity of OncoKB data.

GitHub has been designated as the single location to deposit API, versioning information and supporting SOPs and documentation. OncoKB provides interested parties instruction on how to use OncoKB software working with the database dump. The software is publicly available on Docker (https://hub.docker.com/repository/docker/oncokb/oncokb).

4.4. Information Security and Privacy

OncoKB acknowledges that for FDA-recognition, adequate security measures are recommended to be in place to ensure the protection and privacy of personally identifiable information and protected health information. OncoKB does not store personally identifiable information and/or protected health information, therefore this specific information is not connected to or contained within the OncoKB database.

4.5. Data Formats and Nomenclature

The OncoKB curation platform homepage (http://oncokb.mskcc.org/curate/#!/genes) lists all genes in the curation system. The Genes homepage is displayed upon entering the OncoKB curation interface and is the main homepage of the curation interface. This page lists all genes (linking each listed gene to its own Gene Curation Page) in the OncoKB curation system, along with sortable columns containing the following information for each gene:
1. Last modified: Timestamp indicating when the Gene Curation Page was last modified
2. Last modified by: Name of the last user to edit the page
3. Needs to be reviewed: Indicates if there is new content in the Gene Curation Page that needs to be reviewed by the SCMT.
4. Search Box (Allows the user to search for their gene of interest, the last modified user of interest, or the last modified date of interest

OncoKB uses the following commonly accepted data and nomenclature formats:

- For gene names, Human Genome Organization (HUGO) gene symbols from Human Gene Nomenclature Committee (HGNC, https://www.genenames.org) are being used and updated periodically when new HUGO releases are made available.
- For variant names, Human Genome Variation Society (HGVS) nomenclature, endorsed by the American College of Medical Genetics (ACMG), is being used.
- For tumor types, OncoTree (http://oncotree.mskcc.org), adopted as a standard by AACR Project GENIE, is being used.
- or drug names and identities, NCI Thesaurus (NCIt) terminology is being used.

General rules for how input and format variant level data in the OncoKB Curation platform are shown below.

	Style and formatting rules for variant-level data in in OncoKB curation platform		Nesting of biological/therapeutic information
General variant input rules	Multiple mutations may be grouped together (comma separated) for curation of shared clinical implications and/or tumor type summaries. The oncogenic and mutation effect of each of the mutations should be curated separately.		Must have an associated oncogenic effect, mutation effect, and description of evidence based on the available evidence. References (PMIDs and abstracts) must be included in the description of mutation effect. Clinical implications and/or tumor type summaries can also be curated under mutational ranges.
	Mutation ranges, which capture all amino acid substitutions in a specified amino acid range, can be used (e.g., TP53 102_292mis [TP53 DNA binding domain mutations]).		
Alteration codes	a. mis = missense mutation - e.g., 102_292mis [DNA binding domain missense mutations] b. dup = duplication of a specified range - e.g., S501_A502dup . del = in-frame deletion of a specified range - e.g., P551_E554del d. ins = in-frame insertion - e.g., W557_V559delinsC; e.g.T574insTQLPYD e. delins = in-frame alteration - interpreted by the number of amino acid changes. f. nontrunc = any non-truncating mutation - e.g., R449_E514 nontrunc g. fs = frameshift - e.g., N457Mfs*22 h. _splice = splice mutations - e.g., X963_D1010splice or X963_splice i. trunc = truncating mutation - e.g., D286_L292trunc j. 1? = start lost - e.g., M1? * = stop gained - e.g., R2019*		
Brackets and parentheses in the mutation header	Square Brackets [ ] - used in the mutation header to rename a curated alteration.	The OncoKB website will display the alteration as the text in the bracket versus variant name (e.g., “Exon 19 insertion” instead of 729_761ins).	
	Parentheses () - used in the mutation header to leave comments.	Any text in () in the mutation header is for administrative purposes only and can only be viewed within the OncoKB curation interface. Does not affect the output of how a mutation is displayed.	
Missense mutations	naming convention for missense mutations is <ref_allele><position><tumor_allele> (e.g., V600E)		Every missense mutation needs to be separately curated with respect to its oncogenic and mutation effect.

	Style and formatting rules for variant-level data in in OncoKB curation platform	Nesting of biological/therapeutic information
	Positional variants, which capture all amino acid substitutions at a given position, can be used for curation of shared clinical implications and/or tumor type summaries (e.g., KRAS G12, BRAF V600).	Do not include curation of oncogenic effect or mutation effect, as this information should be captured under each allele-specific missense mutation for which there is functional data.
Truncating mutations	"Truncating Mutations" can be curated as a specific alteration within a Gene Page. Truncating mutations in a tumor suppressor gene include the following mutations: nonsense/frameshift/deletion/splice site mutation	Must have an associated oncogenic effect, mutation effect, and description of evidence.
		Oncogenic and mutation effect should be marked as "Likely Oncogenic " and "Likely Loss of Function" respectively.
		Clinical implications and/or tumor type summaries can also be curated under "Truncating Mutations."
		The oncogenic effect, mutation effect and clinical implications associated with "Truncating Mutations" can be limited by defining a range for the truncation (e.g., "CCND1 256_286trunc [C Terminal Truncating Mutations]").
	"Truncating Mutations" include the following based on the Sequence Ontology : a. Stop_lost: A sequence variant where at least one base of the terminator codon (stop) is changed,resulting in an elongated transcript b. Start_lost: A codon variant that changes at least one base of the canonical start codon c. Stop_gained: A sequence variant where at least one base of a codon is changed, resulting in a premature stop codon and leading to a shortened transcript d. TFBS_ablation: A feature ablation where the deleted region includes a transcription factor binding site e. Feature_truncation: A sequence variant that causes the reduction of a genomic feature, with regard to the reference sequence f. Frameshift_variant: A sequence variant which causes a disruption of the translational reading frame,i.e., the number of nucleotides inserted or deleted is not a multiple of three g. Transcript_ablation: A feature ablation whereby the deleted region includes a transcript feature h. Splice_donor_variant: A splice variant that changes the 2-base	

	Style and formatting rules for variant-level data in in OncoKB curation platform	Nesting of biological/therapeutic information
	region at the 5' end of an intron i. Splice_region_variant: A sequence variant in which a change has occurred within the region of thesplice site, either within 1-3 bases of the exon or 3-8 bases of the intron j. Stop_retained_variant: A sequence variant where at least one base in the terminator codon ischanged, but the terminator remains k. Splice_acceptor_variant: A splice variant that changes the 2-base region at the 3' end of an intron l. Incomplete_terminal_codon_variant: A sequence variant where at least one base of the final codon of an incompletely annotated transcript is changed.	
Fusions	"Fusions" can be curated as a specific gene alteration within a Gene Page, and include any fusion that involves the specified gene	Must have an associated oncogenic effect, mutation effect, and description of evidence.
		Oncogenic and mutation effect should be marked as "Likely Oncogenic " and "Likely Gain of Function" respectively.
		Clinical implications and/or tumor type summaries can also be curated under "Fusions."
	Specific fusions, in which both fusion partners are specified, can be curated if there is functional evidence in the literature describing their oncogenic and/or mutation effect. These have the format "GeneA-GeneB Fusion" (e.g., BCR-ABL1 Fusion)	Oncogenic effect, mutation effect, and clinical implications of the specific fusion alteration will be prioritized over those of the "Fusions" alteration.
		Specific fusion names two gene partners, the alteration is only curated in one Gene Page - the gene that is the main driver (or hypothesized to be the main driver) of the fusion oncoprotein
Copy number aberrations	"Amplification" and "Deletion" can be curated as specific gene alterations within a Gene Page if appropriate functional data exists	Must have an associated oncogenic effect, mutation effect, and description of evidence.
		Prognostic implications, clinical implications and/or tumor type summaries can also be curated under "Amplification" and "Deletion."
In-frame Deletions or Insertions	In-frame deletions or insertions can be curated as a specific gene alteration within a Gene Page	Each curated alteration must have an associated oncogenic effect, mutation effect, and description of evidence.

	Style and formatting rules for variant-level data in in OncoKB curation platform		Nesting of biological/therapeutic information	
			Clinical implications and/or tumor type summaries can also be curated under an in-frame deletion or insertion.	
	1. “del” = in-frame deletion (e.g., P551_E554del, P191del) 2. “ins” = in-frame insertion (e.g., T574insTQLPYD) 3. “delins” = a specified in-frame alteration. Whether the alteration is an in-frame deletion or in-frame insertion is determined by the specified number of amino acid changes.			
Oncogenic Mutations	can be curated as a specific gene alteration within a Gene Page. is used when there is tumor-specific information that applies to ALL functional (oncogenic/likely oncogenic) alterations within a Gene Page.		The tumor-specific information will automatically get linked to all mutations in the Gene Page that have the ” Yes” or “Likely” boxes checked next to the Oncogenic label.	If a gene has “Amplification” curated as “Oncogenic” or “Likely Oncogenic”, this alteration will NOT be associated with the tumor-type specific information under “Oncogenic Mutations.”
Hard-coded Alteration Names	Alterations that do not follow the above nomenclature are not supported unless they are hard coded.	1. FLT3: internal tandem duplication 2. EGFR: vIII 3. EGFR: Kinase domain duplication 4. EGFR: C-terminal domain		
Citation Type		**Format**	**Example**	
Publication in PubMed		(PMID: ########)	(PMID: 28890946)	
Conference Abstract		(Abstract: Author et al. Abstract# ###, Meeting, Year. URL).	(Abstract: Suehnholz et al. Abstract# 3208, AACR 2020. https://cancerres.aacrjournals.org/content/80/16_Supplement/3208)	

4.6. Metadata

OncoKB operates independently of sequencing results, and therefore does not store variant detection-associated metadata such as variant allele frequency for somatic variants.

4.7. Data Uniqueness

The OncoKB curation platform automatically ensures that individual data points are not represented more than once in OncoKB. A gene-variant is considered as a single data element in OncoKB. Duplicated variants are programmatically not allowed in the OncoKB curation platform. Entry of a duplicate variant in a gene page will trigger a dialogue box notifying the curators or database administrator that the variant already exists in the database. A curator or administrator is required to delete the previous variant entry or change annotation of the existing variant entry; duplicate entries are not allowed in the system.

5. Software and Cybersecurity

OncoKB is publicly available to anyone using a browser without a log-in requirement. In order to access the OncoKB API programmatically or to use the OncoKB content in commercial/hospital settings, users are required to register and get a license. Documentation was provided to indicate how unauthorized use is prevented including maintenance of code, data and execution integrity.
A detailed description of the design of the device to detect, respond and recover from cybersecurity threats was summarized. Information included a description of the process and software for maintaining code integrity, data integrity, and execution integrity. A risk analysis was conducted, and the cybersecurity controls described.

6. Discussion of the Evidence to Support Recognition

The OncoKB oversight and governance procedures, which includes specific protocols for the OncoKB mapping of somatic variants into FDA levels of evidence, demonstrate that OncoKB operates in a manner that provides sufficient information and assurances regarding the source data, evidence review, and variant assertions presented in the "FDA-recognized Content" tab within the OncoKB database. OncoKB provides transparency regarding its data sources and operations, and contains genetic variant information generated by validated methods. The procedures to determine the FDA levels of evidence using OncoKB protocols are sufficiently robust to provide a high degree of confidence that the assertions are accurate and the OncoKB database constitutes valid scientific evidence that can be used to support clinical validity of genetic tests in future premarket submissions. These procedures also collect, store, and report data and conclusions in compliance with all applicable requirements regarding protected health information, patient privacy, research subject protections, and data security.

7. Conclusions

To determine if recognition was appropriate, FDA evaluated the oversight and governance of OncoKB as well as the protocols and procedures to establish the FDA levels of evidence presented in the "FDA-recognized Content" tab within the OncoKB database and the containing annotations for somatic variants consistent with tumor profiling claims. That review is described above.

The policies and procedures established for the OncoKB database provide assurance that assertions displayed on the "FDA-recognized Content" tab within the OncoKB database constitute valid scientific evidence and support recognition of this portion of the database. The FDA concludes that the FDA tab within the OncoKB database could be used to support clinical validity of tumor profiling tests in future premarket submissions.