

1 **TITLE:**

2 Evaluating the clinical validity of gene-disease associations: an evidence-based
3 framework developed by the Clinical Genome Resource

4

5 **AUTHORS:**

6 Natasha T. Strande*¹, Erin Rooney Riggs*², Adam H. Buchanan³, Ozge Ceyhan-
7 Birsoy⁶⁻⁹, Marina DiStefano⁶, Selina S. Dwight⁴, Jenny Goldstein¹, Rajarshi Ghosh⁵,
8 Bryce A. Seifert¹, Tam P. Sneddon⁴, Matt W. Wright⁴, Laura V. Milko¹, J. Michael
9 Cherry⁴, Monica A. Giovanni³, Michael F. Murray³, Julianne M. O'Daniel¹, Erin M.
10 Ramos¹⁰, Avni B. Santani¹¹⁻¹², Alan F. Scott¹³, Sharon E. Plon⁵, Heidi L. Rehm⁶⁻⁹,
11 Christa L. Martin^{#2,3}, Jonathan S. Berg^{#1}

12

13 **AFFILIATIONS:**

14 ¹ Department of Genetics, School of Medicine, University of North Carolina at Chapel
15 Hill, Chapel Hill, North Carolina, USA.

16 ² Autism & Developmental Medicine Institute, Geisinger Health System, Danville,
17 Pennsylvania, USA.

18 ³ Genomic Medicine Institute, Geisinger Health System, Danville, Pennsylvania, USA.

19 ⁴ Department of Genetics, Stanford University, Stanford, California 94305, USA.

20 ⁵ Department of Pediatrics, Baylor College of Medicine, Houston, Texas, USA.

21 ⁶ Laboratory for Molecular Medicine, Partners Personalized Medicine, Boston,
22 Massachusetts, USA.

23 ⁷ The Broad Institute of Harvard and MIT, Cambridge, Massachusetts, USA.

1 ⁸ Harvard Medical School, Boston, Massachusetts, USA.

2 ⁹ Department of Pathology, Brigham & Women's Hospital, Boston, Massachusetts.

3 ¹⁰ National Human Genome Research Institute, National Institutes of Health, Bethesda,
4 Maryland, USA.

5 ¹¹ Department of Pathology and Laboratory Medicine, Perelman School of Medicine,
6 University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA.

7 ¹² Division of Genomic Diagnostics, Children's Hospital of Philadelphia, Philadelphia,
8 Pennsylvania 19104, USA.

9 ¹³ McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of
10 Medicine, Baltimore, MD 21287, USA.

11

12 **JOINT AUTHORSHIP:** *, # Authors contributed equally

13 **CORRESPONDING AUTHOR:** Jonathan S. Berg (jonathan_berg@med.unc.edu)

1 **ABSTRACT**

2 With advances in genomic sequencing technology, the number of reported gene-
3 disease relationships has rapidly expanded. However, the evidence supporting these
4 claims varies widely, confounding accurate evaluation of genomic variation in a clinical
5 setting. Despite the critical need to differentiate clinically valid relationships from less
6 well-substantiated relationships, currently no standard guidelines for such evaluation
7 exist. Thus the NIH-funded Clinical Genome Resource (ClinGen) has developed a
8 framework to define and evaluate the clinical validity of gene-disease pairs across a
9 variety of Mendelian disorders. Relevant genetic and experimental evidence supporting
10 or contradicting a gene-disease relationship is evaluated semi-quantitatively and
11 assigned a preliminary classification: “Definitive”, “Strong”, “Moderate”, “Limited”, “No
12 Reported Evidence” or “Conflicting Evidence.” Classifications are reviewed and
13 confirmed or adjusted based on clinical expertise of appropriate disease experts. This
14 evidence-based, systematic method to assess the strength of gene-disease
15 relationships will facilitate more knowledgeable utilization of genomic variants in clinical
16 and research settings.

17

1 **INTRODUCTION**

2
3 The human genome comprises approximately 20,000 protein-coding genes¹, of which
4 about 3,000 have been reported in association with at least one Mendelian disease².
5 Roughly half² of these gene-disease relationships have been identified over the last
6 decade, as technological advances have made it possible to use sequence information
7 from small families or even single individuals to discover new candidate gene-disease
8 relationships^{3, 4}. However, there is substantial variability in the level of evidence
9 supporting these claims, and a systematic method for curating and assessing evidence
10 is needed.

11 Despite this variability, clinical laboratories may include genes with preliminary evidence
12 of a gene-disease relationship on disease-targeted panels, or in results returned from
13 exome/genome sequencing. Some of the gene-disease relationships are either unable
14 to be confirmed for many years or are ultimately proven wrong⁵. Evaluating the clinical
15 impact of variants identified in genes with an unclear role in disease is exceedingly
16 difficult, and could lead to an incorrect diagnosis for the patient, preventing further
17 evaluations and/or resulting in errant management of the patient and family. This
18 scenario highlights the need for a standardized method to evaluate the evidence
19 implicating a gene in disease and thereby determine the clinical validity³ of a gene-
20 disease relationship.

21 The NIH-funded Clinical Genome Resource (ClinGen)⁶ is creating an open-access
22 resource to better define clinically relevant genes and variants based on standardized,
23 transparent evidence assessment for use in precision medicine and research. Our

1 group has developed a method that 1) qualitatively defines gene-disease clinical validity
2 using a classification scheme based on the strength of evidence supporting the
3 relationship, and 2) provides a standardized semi-quantitative approach to evaluate
4 available evidence and arrive at such a classification. Currently, this framework is
5 optimized for genes associated with monogenic disorders following autosomal
6 dominant, autosomal recessive, or X-linked inheritance. Future iterations will expand the
7 framework to consider other modes of inheritance, such as mitochondrial, and diseases
8 with more complex genomic etiologies, including oligogenic or multifactorial conditions.
9 Our approach is neither intended to define multifactorial disease risk, nor to be a
10 substitute for well-established statistical thresholds used for genome-wide association
11 studies^{7; 8}.

12 This novel framework classifies gene-disease relationships by the quantity and quality
13 of the evidence supporting such a relationship. It builds on efforts to catalog gene-
14 disease associations, such as the Online Mendelian Inheritance in Man (OMIM)¹ and
15 OrphaNet, by systematically organizing the supporting and refuting evidence, and
16 categorizing the strength of evidence supporting these relationships. The resulting
17 clinical validity classifications are valuable to both clinicians and clinical laboratories.
18 First, they provide insight into the strength of clinical associations for clinicians
19 interpreting genetic test results for patient care. Second, they serve to guide clinical
20 genetic testing laboratories as they develop disease-specific clinical genetic testing
21 panels or interpret genome-scale sequencing tests. By including only those genes with
22 established clinical validity, the possibility of returning ambiguous, incorrect, or
23 uninformative results is reduced, improving the quality of interpretation of genomic data.

1 **QUALITATIVE DESCRIPTION: CLINICAL VALIDITY CLASSIFICATIONS**

2 The ClinGen Gene Curation Working Group (GCWG) is comprised of medical
3 geneticists, clinical laboratory diagnosticians, genetic counselors, and biocurators with
4 broad experience in both clinical and laboratory genetics. Over the course of three
5 years, this group convened bi-monthly to develop the described framework for
6 assessing gene-disease clinical validity through expert opinion and working group
7 consensus (additional details in Supplemental Methods). We first defined six classes to
8 qualitatively describe the strength of evidence supporting a gene-disease association
9 (Figure 1). The amount and type of evidence required for each clinical validity
10 classification builds upon that of the previous classification level. Evidence used within
11 this framework to assign a classification to a gene-disease pair is divided into two main
12 types: genetic evidence and experimental evidence (described below). As evidence is
13 likely to change over time, any given classification is only representative of the level of
14 evidence at the time of curation.

15 The classification “No Reported Evidence” is used for genes that have not yet been
16 asserted to have a causal relationship with a human monogenic disorder, but may have
17 some experimental data (e.g., model system data) suggesting a potential role for that
18 gene in disease. The “Limited” classification requires at least one variant, asserted to be
19 disease-causing, to have plausible genetic evidence to support the association with
20 human disease with or without gene-level experimental data. “Moderate” classification
21 encompasses additional clinical evidence (eg. multiple unrelated probands harboring
22 variants with potential roles in disease) and supporting experimental evidence, all of
23 which may be provided by multiple studies or a single robust study. Replication of the

1 gene-disease association in subsequent independent publications and additional
2 substantial genetic and experimental data are critical factors for the “Strong”
3 classification. Finally, the hallmark of a “Definitive” gene-disease association is that, in
4 addition to the accumulation of convincing genetic and experimental evidence, the
5 relationship has been replicated, and ample time has passed since the initial publication
6 (in general, greater than three years) for any conflicting evidence to emerge. It is
7 important to highlight that these classifications do not reflect the effect size or relative
8 risk attributable to variants in a particular gene, but instead the strength of the evidence.
9 For example, a definitive gene-disease association does not imply that a pathogenic
10 variant in that gene confers 100% penetrance of the phenotype. This metric is not
11 intended to assess the penetrance or risk to develop a disease outcome.

12 A gene-disease relationship can be determined to have one of the above classifications
13 provided no substantial relevant and valid contradictory evidence exists to call the gene-
14 disease relationship into question. If such evidence emerges, then the relationship is
15 described as “Conflicting Evidence Reported.” Types of contradictory evidence may
16 come from population studies (such as ExAC⁹), attempts to experimentally validate the
17 gene-disease association, or re-analysis of the original family or cohort that was
18 previously studied. Although the role of a specific *variant* in a given disease may be
19 called into question by new evidence, this may not be sufficient to invalidate the role of
20 the *gene* in that disease. Thorough evaluation by experts in the particular disease area
21 is recommended to determine whether the contradictory evidence outweighs the
22 existing supportive evidence to classify a gene into either a “Disputed” or “Refuted”
23 category (see Figure 1 for additional details).

1 **METHODS: SEMI-QUANTITATIVE ASSESSMENT OF EVIDENCE**

2 Assigning a clinical validity classification to a gene-disease pair requires assessment of
3 the evidence supporting the association. We developed a semi-quantitative approach to
4 evaluate both genetic (Figure 2) and experimental evidence (Figure 3) in a standardized
5 manner that promotes consistent collection and weighting of evidence. Defined sub-
6 categories of genetic and experimental evidence are given a suggested default “score.”
7 However, given that evidence of the same general type may vary in its strength
8 (particularly when considering different diseases), the scoring system also allows these
9 scores to be adjusted within a set range of points, with final approval by experts within
10 the particular disease domain. Finally, the maximum number of points allowed for the
11 various types of genetic and experimental evidence is capped to prevent a
12 preponderance of weak evidence from inappropriately inflating the gene-disease
13 classification. Similarly, certain evidence categories are provided higher maximum
14 scores, allowing key pieces of stronger evidence to proportionately influence the
15 classification of a gene-disease pair.

16 **Genetic Evidence**

17 For the purposes of scoring, genetic evidence is divided into two categories: case-level
18 data and case-control data (Figure 2). Studies describing individuals or families with
19 genetic variants are scored as case-level data, while studies using statistical analyses
20 to compare variants in cases and controls are scored as case-control data. When case-
21 level and case-control data are present in a single publication, points can be assigned in
22 each category, but the same piece of evidence should not be counted more than once.

1 For example, an individual case that is also included within a case-control cohort should
2 not be given points in both the “case-level data” and “case-control data” categories. In
3 this scenario, points should be assigned to the most compelling and informative
4 evidence.

5 Assessing case-level data requires consideration of the inheritance pattern and
6 evaluation of the individual variants identified in each case. Within this framework, a
7 case should only be counted towards supporting evidence if the reported variant has
8 some indication of a potential role in disease (e.g., impact on gene function, recurrence
9 in affected individuals, etc.) and does not have evidence that would contradict
10 pathogenicity (e.g., population allele frequency). Unless otherwise noted, the term
11 “qualifying variant” implies that these criteria are met. In addition, points are assigned
12 separately for segregation data to reflect the statistical probability that the locus is
13 implicated in the disease. Figure 2 and Figure S1 provide guidance on the number of
14 points that should be considered for segregation evidence by LOD score; if a LOD score
15 is not provided within the publication being evaluated, an estimated LOD score may be
16 calculated in certain scenarios, as described in the Supplemental Methods.

17 Each study categorized as “case-control data” should be independently assessed to
18 evaluate the quality of the study design (see Figure 2 and Supplemental Methods).
19 Consultation with a clinical domain expert group (such as those affiliated with ClinGen,
20 <https://www.clinicalgenome.org/working-groups/clinical-domain/>) is recommended. For
21 the purposes of this framework, studies are classified based on whether they include
22 single variant analysis or aggregate variant analysis. Single variant analyses are those
23 in which individual variants are evaluated for statistical enrichment in cases compared

1 to controls. More than one variant may be analyzed, but the variants have been
2 independently assessed with appropriate statistical correction for multiple testing.
3 Aggregate variant analyses are those in which the total number of variants is assessed
4 for enrichment in cases compared with controls. This comparison is typically
5 accomplished by sequencing the entire gene in both cases and controls and
6 demonstrating an increased “burden” of variants of one or more types.

7 **Experimental Evidence**

8 The experimental data scoring system is presented in Figure 3. The gene-level
9 experimental data used in this framework to assess a gene-disease association are
10 consistent with those proposed by MacArthur and colleagues to implicate a gene in
11 disease¹⁰. The following experimental evidence types are used: biochemical function,
12 experimental protein interactions, expression, functional alteration, phenotypic rescue
13 and model systems (Figure 3 bottom panel). These categories capture the most
14 relevant types of experimental information necessary to determine whether the function
15 of the gene product is at least consistent with the disease with which it is associated, if
16 not causally implicated.

17 **Contradictory Evidence**

18 While curators are encouraged to seek out and document (via qualitative description)
19 conflicting evidence, no specific points are assigned to this category. The types of valid
20 contradictory evidence and their relative weights will be unique to each gene-disease
21 pair, and it would be misleading to attempt to uniformly quantify this type of negative
22 evidence against the reported positive evidence. If there is substantial conflicting

1 evidence, manual review and expert input is required to evaluate the strength of the
2 contradictory evidence, determine whether it outweighs any available supporting
3 evidence, and, if so, decide whether the gene-disease association should be classified
4 as “Disputed” or “Refuted”.

5 **Summary & Final Matrix**

6 The scores assigned to both genetic and experimental evidence are tallied to generate
7 a total score (ranging from 1-18) that corresponds to a preliminary clinical validity
8 classification (Figure 4). The system provides a transparent method for summarizing
9 and assessing all curated evidence for a gene-disease pair, encouraging consistency
10 between curators. While the summary matrix facilitates a preliminary assessment of the
11 gene-disease relationship, the initial curator or expert reviewer may adjust the
12 classification, supplying a specific rationale for the change. Final classifications are
13 determined in collaboration with disease experts, who review the preliminary
14 classification and supporting evidence and work to come to a consensus with the
15 preliminary curators. In the event that the disease experts and preliminary curators
16 disagree on a final classification, a senior member of the ClinGen Gene Curation
17 Working Group may be brought in to facilitate a final classification, erring towards the
18 more conservative classification if consensus cannot be achieved. It should be noted
19 that experimental data alone cannot justify a clinical validity classification beyond “No
20 Reported Evidence,” and at least one human genetic variant with a plausible causal
21 association must be present to attain “Limited” classification. The difference between
22 “Limited,” “Moderate,” and “Strong” gene-disease classifications is justified by the
23 quality and quantity of evidence; it is expected that valid gene-disease associations will

1 gradually accumulate enough supporting evidence and be replicated over time to attain
2 a “definitive” classification. This framework relies on evidence obtained primarily from
3 published literature; however, if necessary, unpublished information available from
4 publicly accessible resources, such as variant databases^{11; 12}, may be used as long as
5 some supporting evidence is provided.

6

7 **RESULTS: VALIDATION OF METHOD**

8 Using this framework we evaluated 33 gene-disease pairs representing a variety of
9 disease domains and spanning the spectrum of clinical validity classifications (see
10 Figure 5 and Supplemental Methods). To assess the reproducibility of our scoring
11 metric, each gene-disease pair was evaluated by two independent curators; paired
12 curators reached concordant clinical validity classifications in 29 of the 31 (93.5%)
13 gene-disease pairs with available published evidence (Figure 5; associations classified
14 as “No Reported Evidence” were excluded). Each gene-disease pair was subsequently
15 reviewed by clinical domain experts; experts agreed with the preliminary classifications
16 for 87.1% (27/31) of the gene-disease pairs with published evidence (Figure 5). The
17 four discrepancies between the expert and curator classifications were each different by
18 only a single category (e.g. limited versus moderate). Of note, the original classifications
19 for *HNRNPK* and *SMARCA1* were at the border between limited and moderate (6.5
20 points); in each case, the preliminary curators’ lack of specific clinical expertise led to
21 uncertainty regarding the scoring of evidence requiring such knowledge. Consulting with
22 clinical experts in the disease resolved these issues resulting in both genes being
23 upgraded to moderate. In the case of *WRAP53*, the expert was aware of additional

1 published experimental evidence that when included increased the classification from
2 limited to moderate. Upon reviewing the curated evidence for *RAD51D* and breast
3 cancer, the domain expert upgraded the classification from disputed to limited (with the
4 approval of the GCWG) due to the specificity of the experimental evidence and
5 insufficient power of the current studies to rule out a role for *RAD51D* in breast cancer
6 (Figure 5). Details and references for each curation are provided in Supplemental
7 Appendix.

8 9 **DISCUSSION**

10 The evidence-based framework described here qualitatively defines clinical validity
11 classifications for gene-disease associations in monogenic conditions and provides a
12 systematic framework for evaluating key criteria required for these classifications. This
13 method is intentionally flexible to accommodate curation of a wide spectrum of genes
14 and conditions by curators with varying levels of expertise. The semi-quantitative
15 scoring system combined with the qualitative classification scheme guides curators
16 through the preliminary decision-making process, while the expert-level review provides
17 disease-specific experience to weigh in on the final classification.

18 This effort to create a generalized framework may result in some specific challenges
19 due to the heterogeneity of genetic conditions, in both phenotype and prevalence. For
20 example, ultra-rare disorders may have a relatively small number of probands described
21 in the medical literature, thus limiting their potential to achieve a high genetic evidence
22 score within this matrix. This obstacle is mostly circumvented by allowing compelling
23 pieces of genetic evidence to score the maximum number of points (for example, see

1 *CD3E* and severe combined immunodeficiency, detailed in the Supplemental
2 Appendix). When substantial experimental evidence is also available, these conditions
3 can attain a “Strong” or “Definitive” classification.

4 On the opposite end of the spectrum are conditions that occur commonly in the general
5 population, such as cancer, where the predominant etiology is multifactorial rather than
6 monogenic. In the less common Mendelian cancer predisposition syndromes,
7 incomplete penetrance is a typical feature that can lead to confounding factors in family
8 genetic studies such as apparently non-penetrant family members who carry a disease-
9 associated variant and phenocopies among family members without a disease-
10 associated variant. For such conditions, case-control data may provide more compelling
11 evidence to support the gene-disease association (see the curation of *PALB2* and
12 hereditary breast cancer in the Supplemental Appendix as an example).

13 One limitation of any such system is the challenge of balancing thorough literature
14 curation and practical time commitment. This system can accommodate an exhaustive
15 literature review, but in most cases will only require curating the amount of information
16 sufficient to reach the maximum number of points in the matrix. In some scenarios this
17 method may fail to include pertinent information, which could impact the classification
18 (e.g., omission of contradictory evidence). Another potential limitation is the subjective
19 nature of certain evidence types (e.g., experimental), which may lead to variability
20 between different groups assessing evidence. However, due to the transparency of the
21 evidence base, the incorporation of expert review, and the ability to reassess
22 classifications over time, such drawbacks are likely to be self-limiting.

1 ClinGen’s ultimate goal is to enhance the incorporation of genomic information into
2 patient care, an important component of the Precision Medicine Initiative¹³. The
3 implementation of this framework will be supported by an open-access ClinGen curation
4 interface (under development), which will provide a platform for extension to the
5 community. In essence, this framework aims to provide a systematic, transparent
6 method to evaluate a gene-disease relationship in an efficient and consistent manner
7 suitable for a diverse set of users. A detailed standard operating procedure for this
8 framework is available on the ClinGen website. All curated evidence, including clinical
9 validity assessments, will be made readily accessible to clinical laboratories, clinicians,
10 patients, researchers, and others via our website.

11 Carefully evaluated gene-disease clinical validity classifications, as provided by this
12 framework, will be useful to clinical laboratories as they evaluate genes for inclusion on
13 disease-targeted panels, or as they decide how to categorize, prioritize, and return
14 results from exome/genome sequencing. Clinicians may choose to use these types of
15 gene-disease classifications as they interpret laboratory results for their patients; for
16 instance, they may choose not to adjust medical management based on variants in
17 genes of limited clinical validity. Researchers could also utilize this framework to
18 evaluate the clinical validity of their own newly discovered associations and identify
19 promising target genes for future work in order to augment the currently available
20 evidence and attain a “Strong” or “Definitive” classification. In addition, professional
21 societies and regulatory bodies may utilize these clinical validity assessments when
22 making recommendations or guidelines for clinical genetic testing. Ultimately, our

- 1 systematic, evidence-based method for evaluating gene-disease associations will
- 2 provide a strong foundation for genomic medicine.

3

4

1 **DESCRIPTION OF SUPPLEMENTAL DATA:**

2 The Supplemental file includes methods, footnotes for Figure 2, one figure, an appendix
3 with curated evidence for each example presented in Figure 5 and a list of references.

4 **ACKNOWLEDGEMENTS:**

5 This work was supported by grants from the National Human Genome Research
6 Institute (NHGRI), through the following three grants: U41 HG006834-01A1, U01
7 HG007437-01, U01 HG007436-01, as well as from the National Cancer Institute (NCI)
8 through the following contract: HHSN261200800001E. ClinGen also receives funding
9 through the Eunice Kennedy Shriver National Institute of Child Health and Human
10 Development (NICHD) and ClinVar is supported by the Intramural Research Program of
11 the NIH, National Library of Medicine. We would like to thank the following groups and
12 individuals for contributing their disease expertise to review the examples included in
13 this manuscript: Alan Beggs Ph.D.; Alison Bertuch, M.D., Ph.D.; Rebecca H. Buckley,
14 M.D.; Eugene Chung, M.D.; Bill Craigen M.D., Ph.D.; Jennifer M. Puck, M.D.; Sharon A.
15 Savage, M.D.; Fergus J. Couch Ph.D. and the ClinGen Hereditary Breast and Ovarian
16 Cancer gene curation working group; Birgit H. Funke Ph.D. and the ClinGen
17 Cardiomyopathy gene curation working group; and the ClinGen RASopathy curation
18 working group. Input on the framework was also provided by the ClinGen Hereditary
19 Breast and Ovarian Cancer gene curation working group and Ray Hershberger, Mike
20 Gollob and the ClinGen Channelopathy gene curation working group. We would also
21 like to thank Scott Goehringer for his invaluable help in preparing the curated examples
22 for the ClinGen website and appendix.

23

1 **WEB RESOURCES:**

2 Clinical Genome Resource: www.clinicalgenome.org

3 Gene Curation Working Group members: [http://clinicalgenome.org/about/working-](http://clinicalgenome.org/about/working-groups/gene-curation/)
4 [groups/gene-curation/](http://clinicalgenome.org/about/working-groups/gene-curation/)

5 Link to the standard operating procedure for the clinical validity framework described in

6 this manuscript: <http://bit.ly/2clingenGCSOP>

REFERENCES:

1. Online Mendelian Inheritance in Man, OMIM®. In. (Baltimore, MD, McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University
2. Chong, J.X., Buckingham, K.J., Jhangiani, S.N., Boehm, C., Sobreira, N., Smith, J.D., Harrell, T.M., McMillin, M.J., Wiszniewski, W., Gambin, T., et al. (2015). The Genetic Basis of Mendelian Phenotypes: Discoveries, Challenges, and Opportunities. *American journal of human genetics* 97, 199-215.
3. Haddow, J., Palomacki, G. (2003). ACCE: A Model Process for Evaluating Data on Emerging Genetic Tests. In *Human Genome Epidemiology: A Scientific Foundation for Using Genetic Information to Improve Health and Prevent Disease*, M. Khoury, Little, J., Burke, W., ed. (Oxford University Press), pp 217-233.
4. Wilfert, A.B., Chao, K.R., Kaushal, M., Jain, S., Zollner, S., Adams, D.R., and Conrad, D.F. (2016). Genome-wide significance testing of variation from single case exomes. *Nat Genet* 48, 1455-1461.
5. Eisenberger, T., Di Donato, N., Baig, S.M., Neuhaus, C., Beyer, A., Decker, E., Murbe, D., Decker, C., Bergmann, C., and Bolz, H.J. (2014). Targeted and genomewide NGS data disqualify mutations in MYO1A, the "DFNA48 gene", as a cause of deafness. *Human mutation* 35, 565-570.
6. Rehm, H.L., Berg, J.S., Brooks, L.D., Bustamante, C.D., Evans, J.P., Landrum, M.J., Ledbetter, D.H., Maglott, D.R., Martin, C.L., Nussbaum, R.L., et al. (2015). ClinGen--the Clinical Genome Resource. *The New England journal of medicine* 372, 2235-2242.
7. Lander, E., and Kruglyak, L. (1995). Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet* 11, 241-247.
8. Sham, P.C., and Purcell, S.M. (2014). Statistical power and significance testing in large-scale genetic studies. *Nature reviews Genetics* 15, 335-346.
9. Lek, M., Karczewski, K.J., Minikel, E.V., Samocha, K.E., Banks, E., Fennell, T., O'Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B., et al. (2016). Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 536, 285-291.
10. MacArthur, D.G., Manolio, T.A., Dimmock, D.P., Rehm, H.L., Shendure, J., Abecasis, G.R., Adams, D.R., Altman, R.B., Antonarakis, S.E., Ashley, E.A., et al. (2014). Guidelines for investigating causality of sequence variants in human disease. *Nature* 508, 469-476.
11. Landrum, M.J., Lee, J.M., Riley, G.R., Jang, W., Rubinstein, W.S., Church, D.M., and Maglott, D.R. (2014). ClinVar: public archive of relationships among sequence variation and human phenotype. *Nucleic acids research* 42, D980-985.
12. Fokkema, I.F., Taschner, P.E., Schaafsma, G.C., Celli, J., Laros, J.F., and den Dunnen, J.T. (2011). LOVD v.2.0: the next generation in gene variant databases. *Human mutation* 32, 557-563.

13. Collins, F.S., and Varmus, H. (2015). A new initiative on precision medicine. *The New England journal of medicine* 372, 793-795.

1 **FIGURE TITLES AND LEGENDS:**

2 **Figure 1: ClinGen clinical validity classifications and qualitative descriptions.** The
3 suggested minimum criteria needed to obtain a given classification are described for
4 each clinical validity classification. The types of evidence comprising these criteria are
5 described in the text. The default classification for genes without a convincing human
6 disease-causing variant is “No Reported Evidence.” The level of evidence needed for
7 each supportive gene-disease association category builds upon the previous category
8 (i.e. “Limited” builds upon “Moderate”). Gene-disease associations classified as
9 “Contradictory” likely have supporting evidence as well as opposing evidence, but are
10 described separately from the classifications for supportive gene-disease associations.

11 **Figure 2: Classes of genetic evidence and their relative weights used in the**
12 **ClinGen clinical validity framework.** Genetic evidence is separated into two main
13 categories: case-level data and case-control data. While a single publication may
14 include both case-level and case-control data, individual cases should NOT be included
15 in both categories. Each category is assigned a range of points with a maximum score
16 that can be achieved. Case-Level Data is derived from studies describing individuals
17 and/or families with qualifying variants in the gene of interest. Points should be assigned
18 to each case based on the variant’s inheritance pattern, molecular consequence and
19 evidence of pathogenicity in disease. In addition to variant evidence points, a gene-
20 disease pair may also receive points in the segregation evidence category for
21 compelling segregation analysis (see Figure S1). Case-Control Data: Studies utilizing
22 statistical analysis to evaluate variants in cases compared to controls. Case-control
23 studies can be classified as either single variant analysis or aggregate variant analysis,

1 however the number of points allowable for either category is the same. Points should
2 be assigned to case-control studies according to the overall quality of each study based
3 on these criteria: variant detection methodology, power, bias and confounding factors,
4 and statistical power. Additional details included in Supplemental Methods. Note that
5 the maximum total scores allowed for different types of Case-Level data are not
6 intended to add up to the total points allowed for Genetic Evidence as a whole. This
7 permits different combinations of evidence types to achieve the maximum total score.

8 **Figure 3: Types of gene-level experimental evidence and their relative weights**
9 **used in the ClinGen clinical validity framework.** Experimental evidence types used in
10 the ClinGen gene curation framework are modified from MacArthur, et al. 2014.
11 Evidence types are divided into three categories based on their relative contribution to
12 the overall clinical validity of a gene-disease pair giving more weight to *in vivo* data.
13 Each category is assigned a range of points with a maximum score that can be
14 achieved, allowing more weight to be given to *in vivo* data (e.g. Models & Rescue) over
15 *in vitro* experimental data. Evidence within the function category is given the least
16 weight and is comprised of the following types of evidence: biochemical function,
17 interactions, and expression. Functional alteration experiments in cells from patients
18 carrying candidate pathogenic variants are given more weight than the function
19 category. Finally model systems and phenotypic rescue experiments are given the most
20 weight in our framework. Note that the maximum total scores allowed for different
21 categories of Experimental Evidence are not intended to add up to the total allowable
22 points. This permits different combinations of evidence types to achieve the maximum
23 total score.

1 **Figure 4. Final summary matrix used to provisionally classify gene-disease**
2 **associations.** A summary matrix was designed to generate a “provisional” clinical
3 validity assessment using a point system consistent with the qualitative descriptions of
4 each classification. Genetic Evidence: total number of points (not exceeding 12)
5 obtained using the scoring metric in Fig. 2. If no human mutations have been found
6 within the literature, then the default classification is “No Reported Evidence.”
7 Experimental Evidence: total number of points (not exceeding 6) derived from each of
8 the experimental categories in Fig. 3. Replication Over Time – Yes, if more than three
9 years has passed since the publication of the first paper reporting the gene-disease
10 relationship AND more than two publications with human mutations exist. Contradictory
11 Evidence – No points are assigned to this category. Instead, the curator should provide
12 a summary of contradictory information. Scoring - The sum of the quantified evidence
13 from each category can be used to determine a “provisional” classification using the
14 scale at the bottom of the figure. If a curator does not agree with this classification,
15 he/she may provide a different suggested classification along with appropriate
16 justification.

17 **Figure 5. Comparison of provisional clinical validity classifications and**
18 **associated matrix scores for selected gene-disease pairs evaluated by multiple**
19 **curators.** Of the 33 gene-disease pairs (y-axis) curated to validate the clinical validity
20 curation framework, 31 were classified using the summary matrix (2 gene-disease pairs,
21 *PMS2*:pancreatic cancer and *ARSD*:chondrodysplasia punctata, were classified as “No
22 evidence reported” and are not shown). Genetic evidence (grey bars) and experimental
23 evidence (black bars) were evaluated by two independent curators (C1-C9) to arrive at

1 a provisional classification (x-axis). Gene-disease relationships scoring between 12-18
2 points can be “Strong” or “Definitive,” depending on whether the association has been
3 replicated over time (indicated by the squared “r/t”), in which case the preliminary
4 classification is “Definitive”. Clinical validity classifications that were discordant between
5 preliminary curators are represented with a dashed background. Gene-disease pairs in
6 which conflicting evidence was reported are represented by diagonal lines through the
7 evidence bars and a grey background. The letter “C” in a triangle indicates that the
8 curators classified the gene-disease pair as “Conflicting Evidence Reported”. Each
9 gene-disease pair was ultimately evaluated by an expert in the field for a final
10 classification (far right column). Final expert classifications that differed from the
11 preliminary classification are indicated by italics and asterisks.

Figure 1

Evidence Level		Evidence Description
Supportive Evidence	DEFINITIVE	The role of this gene in this particular disease has been repeatedly demonstrated in both the research and clinical diagnostic settings, and has been upheld over time (in general, at least 3 years). No convincing evidence has emerged that contradicts the role of the gene in the specified disease.
	STRONG	The role of this gene in disease has been independently demonstrated typically in at least two separate studies providing strong supporting evidence for this gene's role in disease, usually including both of the following types of evidence: <ul style="list-style-type: none"> • Strong variant-level evidence demonstrating numerous unrelated probands with variants that provide convincing evidence for disease causality¹ as well as • Compelling gene-level evidence from different types of supporting experimental data². <p>In addition, no convincing evidence has emerged that contradicts the role of the gene in the noted disease.</p>
	MODERATE	There is moderate evidence to support a causal role for this gene in this disease, typically including both of the following types of evidence: <ul style="list-style-type: none"> • Several probands with variants that provide convincing evidence for disease causality¹ • Moderate experimental data² supporting the gene-disease association <p>The role of this gene in disease may not have been independently reported, but no convincing evidence has emerged that contradicts the role of the gene in the noted disease.</p>
	LIMITED	There is limited evidence to support a causal role for this gene in this disease, such as: <ul style="list-style-type: none"> • Fewer than three observations of variants that provide convincing evidence for disease causality¹ OR • Variants have been observed in probands, but none have sufficient evidence for disease causality. • Limited experimental data² supporting the gene-disease association <p>The role of this gene in disease may not have been independently reported, but no convincing evidence has emerged that contradicts the role of the gene in the noted disease.</p>
NO REPORTED EVIDENCE		Evidence for a causal role in disease has not been reported. These genes might be "candidate" genes based on linkage intervals, animal models, implication in pathways known to be involved in human diseases, etc., but no reports have directly implicated the gene in human disease cases.
Contradictory Evidence	CONFLICTING EVIDENCE REPORTED	Although there has been an assertion of a gene-disease association, conflicting evidence for the role of this gene in disease has arisen since the time of the initial report indicating a disease association. Depending on the quantity and quality of evidence disputing the association, the association may be further defined by the following two sub-categories: <ol style="list-style-type: none"> 1. Disputed <ol style="list-style-type: none"> a. Convincing evidence <i>disputing</i> a role for this gene in this disease has arisen since the initial report identifying an association between the gene and disease. b. Refuting evidence need not outweigh existing evidence supporting the gene:disease association. 2. Refuted <ol style="list-style-type: none"> a. Evidence refuting the role of the gene in the specified disease has been reported and significantly outweighs any evidence supporting the role. b. This designation is to be applied at the discretion of clinical domain experts after thorough review of available evidence
NOTES		
<p>¹Variants that disrupt function and/or have other strong genetic and population data (e.g. <i>de novo</i> occurrence, absence in controls, strong linkage to a small genomic interval, etc.) are considered convincing of disease causality in this framework.</p> <p>²Examples of appropriate types of supporting experimental data based on those outlined in MacArthur et al. 2014.</p>		

Figure 2

	Evidence Type		Case Information		Suggested Points/Case		Points Given	Max Score
					Default	Range		
Case-Level Data ¹	Variant Evidence	Autosomal Dominant OR X-Linked Disorder ²	Variant is <i>de novo</i> ³		2	0-3		12
			Proband with predicted or proven null variant ⁴		1.5	0-2		10
			Proband with other variant type with some evidence of gene impact ⁵		0.5	0-1.5		7
	Autosomal Recessive	Two variants in <i>trans</i> and at least one <i>de novo</i> ³ or a predicted/proven null variant ⁴		2	0-3		12	
		Two variants (not predicted/proven null) with some evidence of gene impact ⁵ in <i>trans</i>		1	0-1.5			
	Segregation Evidence	Evidence of segregation in one or more families ⁶	LOD Score Examples	3	5	0-7		7
		2		4				
		1.5		3				
		1		1.5				
Case-Control Data ⁷	Case-Control Study Type ⁸	Case-Control Quality Criteria ⁹		Suggested Points/Study		Points Given	Max Score	
	Single Variant Analysis ^{8a}	<ul style="list-style-type: none"> Variant Detection Methodology^{9a} Power^{9b} Bias and Confounding Factors^{9c} Statistical Significance^{9d} 		0-6			12	
	Aggregate Variant Analysis ^{8b}			0-6				
TOTAL ALLOWABLE POINTS for Genetic Evidence								12
General Notes <ul style="list-style-type: none"> All variants under consideration should be rare enough in the general population to be consistent with disease. Cohorts/cases should not be double counted. For example, individual cases included as part of case-control studies should not be given points from both the "Case Level Data" and "Case-Control Data" categories. Case-Level Data includes studies describing individuals or families with variation in the gene of interest Case-Control studies are those in which statistical analysis is used to evaluate variation in cases compared to controls. Numbered footnotes are included in the Supplemental Material. 								

Figure 3

Evidence Category	Evidence Type	Suggested Points		Points Given	Max
		Default	Range		
Function	Biochemical Function	0.5	0-2		2
	Protein Interaction		0-2		
	Expression		0-2		
Functional Alteration	Patient cells	1	0-2		2
	Non-patient cells	0.5	0-1		
Models & Rescue	Animal model	2	0-4		4
	Cell culture model system	1	0-2		
	Rescue in animal model	2	0-4		
	Rescue in engineered equivalent	1	0-2		
Total Allowable Points for Experimental Evidence					6

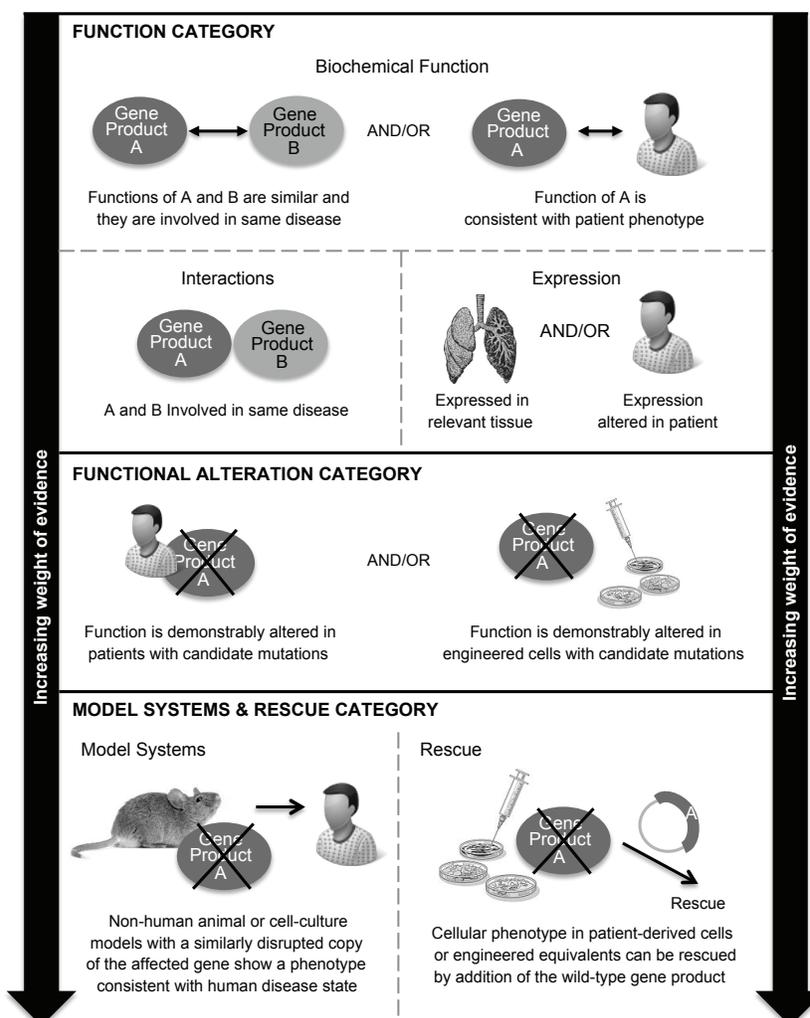


Figure 4

Clinical Validity Summary Matrix

GENE/DISEASE PAIR:				
Assertion criteria	Genetic Evidence (0-12 points)	Experimental Evidence (0-6 points)	Total Points (0-18)	Replication Over Time (Y/N)
Description	Case-level, family segregation, or case-control data that support the gene-disease association	Gene-level experimental evidence that support the gene-disease association	Sum of Genetic & Experimental Evidence	> 2 pubs w/ convincing evidence over time (>3 yrs)
Assigned Points				
CALCULATED CLASSIFICATION		LIMITED	1-6	
		MODERATE	7-11	
		STRONG	12-18	
		DEFINITIVE	12-18 & Replicated Over Time	
Valid contradictory evidence (Y/N)*	List PMIDs and describe evidence:			
CURATOR CLASSIFICATION				
FINAL CLASSIFICATION				

Figure 5

