Fitting a naturally scaled point system to the ACMG/AMP variant classification guidelines

Sean V Tavtigian, PhD[1,3] Kenneth M. Boucher, PhD[2,3] Leslie G. Biesecker, MD[4]

1. Department of Oncological Sciences, University of Utah School of Medicine, Salt Lake City, UT, USA.

2. Department of Internal Medicine, Division of Epidemiology, University of Utah School of Medicine, Salt Lake City, UT, USA.

3. Huntsman Cancer Institute at the University of Utah, Salt Lake City, UT, USA.

4. Medical Genomics and Metabolic Genetics Branch, National Human Genome Research Institute, Bethesda, MD, USA.

Correspondence to:

Sean V Tavtigian, PhD

Tel: +1 801 587 4258

Email: sean.tavtigian@hci.utah.edu

Grant Support

S.V.T is supported in part by R01 CA121245. K.M.B. and the Huntsman Cancer Institute's Biostatistics Core are supported in part by P30 CA042014. L.G.B. is supported by NHGRI grant HG200359.

Conflict of Interest Notification

Disclosure: S.V.T. holds Illumina stock in a personally managed account. L.G.B. is an uncompensated member of an Illumina Advisory Board, receives in-kind research support from ArQule Inc., and honoraria from Cold Spring Harbor Press.

Abstract (150 words max)

Recently, we demonstrated that the qualitative American College of Medical Genetics and Genomics/ Association for Medical Pathology (ACMG/AMP) guidelines for evaluation of Mendelian disease gene variants are fundamentally compatible with a quantitative Bayesian formulation. Here, we show that the underlying ACMG/AMP "strength of evidence categories" can be abstracted into a point system. These points are proportional to Log(odds), are additive, and produce a system that recapitulates the Bayesian formulation of the ACMG/AMP guidelines. Strengths of this system are its simplicity and that the connection between point values and odds of pathogenicity allows empirical calibration of strength of evidence for individual data types. Weaknesses include that a narrow range of prior probabilities is locked in, and that the Bayesian nature of the system is inapparent. We conclude that a points-based system has useful attributes of user friendliness and can be useful so long as the underlying Bayesian principles are acknowledged.

Key Words

Bayesian framework

ACMG

Points-based classification system

Scoring metric

Medical genetics

Variant classification

Unclassified variants

Variants of uncertain significance

VUS

Main Text (3-5 pages)

Recently, we demonstrated that the qualitative American College of Medical Genetics and Genomics/ Association for Medical Pathology (ACMG/AMP) guidelines for the evaluation of Mendelian disease gene variants are fundamentally compatible with a quantitative Bayesian formulation [Richards, 2015; Tavtigian, 2019]. However, actual use of that Bayesian formulation can be challenging for some users because of the required calculations. Through the following brief analysis, we further demonstrate a natural conversion from that Bayesian formulation into a points-based system.

Within the ACMG/AMP variant classification guidelines, thresholds for variant classification are defined by probabilistic boundaries that were set by community consensus [Plon, 2008; Richards, 2015]. Taking into account that the use of inequalities at the threshold boundaries should be symmetric around the broad VUS (Variant of Uncertain Significance) category, these are given in Table 1.

These community agreements in place, the strengths of the various ACMG/AMP rules for combining evidence criteria [Richards, 2015], can be expressed as odds in favor of pathogenicity via a single exponential equation [Tavtigian, 2018]. Here we cite "equation 5" from that publication, using the same variable definitions from that analysis:

equation 1: $$OP = O_{PVSt}^{\left(\frac{N_{PSu}}{8}+\frac{N_{PM}}{4}+\frac{N_{PSt}}{2}+\frac{N_{PVSt}}{1}-\left[\frac{N_{BSu}}{8}+\frac{N_{BSt}}{2}\right]\right)}$$

where *OP* are the calculated odds of pathogenicity; O_{PVSt} are the odds of pathogenicity assigned to the "Very Strong" evidence of pathogenicity category; N_P and N_B are the number of invocations of a specific pathogenic or benign evidence strength level, respectively, by a specific classification rule; and *Su*, *M*, *St*, and *VSt* are "Supporting", "Moderate", "Strong", and "Very Strong" strength of evidence strength level categories, respectively.

Does equation 1 imply a natural point system for variant classification?

Noting that by definition in our previous work $O_{PVSt} = O_{PSu}^{8}$, we can re-write equation 1 as:

equation 2: $$OP = O_{PSu}^{(1N_{PSu}+2N_{PM}+4N_{PSt}+8N_{PVSt}-[1N_{BSu}+4N_{BSt}])}$$

Taking the Log_{10} and then dividing by the Log_{10} (O_{PSu}), we have:

equation 3: $$Log_{10}(OP) \Big/ Log_{10}(O_{PSu}) = \mathbf{1}N_{PSu} + \mathbf{2}N_{PM} + \mathbf{4}N_{PSt} + \mathbf{8}N_{PVSt} - \left[\mathbf{1}N_{BSu} + \mathbf{4}N_{Bst}\right]$$

Inspecting the bolded integers 1, 2, 4, and 8 that emerge on the right side of equation 3, it is evident that the ACMG/AMP strength of evidence categories can be abstracted into a point system, given in Table 2. We emphasize that these points are proportional to Log(odds) rather than odds, and are therefore additive. Indeed, the odds corresponding to an individual rule for combining evidence criteria are easily retrieved, because $OP = O_{PSu}^{(Points)}$

While framing the ACMG/AMP as *OP* expresses a Bayesian point of view, actual application of Bayes' rule arrives when a prior probability of pathogenicity (*P1*) is combined with the *OP* to obtain a posterior probability of pathogenicity (*P2*). Two relevant expressions of Bayes' rule are:

equation 4: $$P2 = \left(OP \times P1\right) \div \left[\left(OP - 1\right) \times P1 + 1\right]$$

equation 5: $$OP = \left[P2 \times \left(1 - P1\right)\right] \div \left[\left(1 - P2\right) \times P1\right]$$

With the prior probability set at 0.10 per the preferred scenario of [Tavtigian, 2018], application of equation 5 shows the odds in favor of pathogenicity (odds) thresholds for Likely Pathogenic and Pathogenic are >81 and >891, respectively. Similarly, the odds thresholds for Likely Benign and Benign are <1.00 and <0.00901, respectively.

With O_{PVSt} set at 350 (and thus O_{PSu} set at 2.08) per that same preferred scenario, the points required to reach the classification thresholds are simply Threshold=2.08^{Points}. Rounding up to the nearest integers, the thresholds for Pathogenic and Likely Pathogenic are 10 points and 6 points, respectively. Rounding down to the nearest integers, the thresholds for Likely Benign

and Benign are -1 and -7 points, respectively. The resulting point-based categorical ranges are given in Table 3.

The principal strength of such a point system is that using it requires only addition and subtraction. The weakness is that the Bayesian nature of the system is hidden. Specific choices of prior probability, odds of pathogenicity, and posterior probability are locked in, and the very concepts of probabilities and odds removed from view. It is important to reiterate, however, that the points described here are intentionally proportional to Log(odds), and simply a shorthand representation of equations 1-3. Consequently, the odds of pathogenicity can be calculated from any evidence combination, then combined with a prior probability using Bayes' rule (i.e., equation 4). As strength of evidence increases in either the pathogenic or benign directions, the resulting posterior probabilities will asymptotically approach 1.00 or 0.00, respectively.

We know of multiple efforts that have developed or are developing point-based systems that are intended to contribute to sequence variant classification. One early effort was a point system for evaluation of rare missense substitutions observed in the mismatch repair (MMR) genes responsible for Lynch syndrome [Barnetson, 2007]. Their system captured six kinds of data, with points ranging from -6 (variant present in controls) to +6 (variant present in more than five family members). The system also had the unusual feature, from a 2020 point of view, that concordance in favor of pathogenicity from several computation tools could generate a higher score than loss of function in an appropriate functional assay. Additionally, no point-thresholds were defined above which a variant could be called pathogenic, nor below which one could be called benign. Indeed, the authors noted that theirs was "a simple arbitrary scoring system that we devised for defining the likelihood of a variant being pathogenic". Although the types of data captured in this system contribute to MMR gene variant classification today, the actual point system that they devised does not.

A much more recent effort involves standards for interpretation of copy number variants [Riggs, 2019]. In this point system, total scores of 0.90 and 0.99 are the thresholds for likely pathogenic and pathogenic, respectively because "variants interpreted as pathogenic should have a 99% level of confidence and variants interpreted as likely pathogenic should have a 90% level of confidence". That is, Riggs *et al*. considered that their score thresholds resemble probabilities of pathogenicity. Within this point system, individual pieces of evidence in favor of pathogenicity receive between 0.10 and 1.00 points, and all of the data for a single sequence variant are added together to arrive at a score for that variant. Focusing the pathogenic side of the Riggs *et al*. system, we would point out three considerations. Firstly, as those authors admit, there is no derived, fitted, trained, or otherwise calibrated connection between evidence types and the points accorded to them (Riggs *et al*. noted "that these numbers have not been statistically derived."). This makes it difficult to calibrate a point scale. Secondly, under Bayes' rule (equation 4), conditional odds of 11.0:1 are required to move from a posterior probability of 0.90 to 0.99. Using the point system that we derived above, bridging that gap–moving from the threshold of likely pathogenic to pathogenic–requires 4 points, i.e., at least four supporting, two moderate, two supporting and one moderate, or one strong piece of pathogenic evidence. Yet in the Riggs *et al*. system, one element of supporting pathogenic evidence would be sufficient. This means that in the Riggs *et al*. point system, the difference between likely pathogenic and pathogenic is very small; in all likelihood, either the likely pathogenic boundary is too strong or the pathogenic boundary too weak. Thirdly, since summing across the data for a single sequence variant can easily result in total scores that exceed 1.0, the thresholds of 0.90 and 0.99 are cannot be considered as posterior probabilities.

A common argument against classification based on point scales is that the scales and classification thresholds tend to be arbitrary. Of course, an arbitrary point-based classification system, if thoughtfully designed, may be operationally satisfactory. While the qualitative

ACMG/AMP variant classification system itself has components that may be considered arbitrary, it was thoughtfully enough designed that an internally consistent Bayesian formulation could be fitted to it. The point system derived here flows naturally from that Bayesian formulation. Indeed, upon examination of the Richards et al. combining rules (their Table 5), simply allowing one point for each invocation of supporting pathogenic evidence, two points for each invocation of moderate pathogenic evidence, etc., could lead one to propose this point system, with the same caveats about the strength of the rules Likely Pathogenic (i) and Pathogenic (iii) that we noted previously [Tavtigian, 2018].

In a more abstract sense depicted in Figure 1, the ACMG/AMP qualitative classification schema provided a scaffold that could be combined with Bayes' rule to produce its Bayesian formulation. Bidirectional feedback between the qualitative classification schema and its Bayesian formulation, with a particular focus on empirical measurement of strength of evidence attributable to existing or new data types, should steadily improve the rigor of sequence variant classification. The point scaled derived here automatically inherits these features.

Variant interpretation is a new and rapidly developing science. All of us are learning and developing novel approaches at a rapid pace and the integration of mathematical, statistical, and computational techniques into our practices will benefit testing laboratories and, ultimately, patient care. Going forward, we recommend that developers of all variant assessment schemes examine their proposed scoring scales, classification thresholds, and underlying logic to see how well they comport with a Bayesian probabilistic framework. This assessment should determine how naturally they flow from the parent ACMG/AMP variant classification guidelines, which were pioneering and insightful and provide a solid foundation for future development efforts.

Tables

Table 1. Variant classification categories and their probabilistic boundaries

Category	Posterior-Probability (PP) based boundaries
Pathogenic	PP > 0.99
Likely Pathogenic	0.99 ≥ PP > 0.90
Uncertain	0.10 ≤ PP ≤ 0.90
Likely Benign	0.001 ≤ PP < 0.10
Benign	PP < 0.001.

Table 2. Point values for ACMG/AMP strength of evidence categories

Evidence Strength	Point Scale Pathogenic	Benign
Supporting	1	-1
Moderate	2	-2 †
Strong	4	-4
Very Strong	8	-8 †

† Note is made that Richards et al did not specify benign evidence at the moderate or very strong levels. Nevertheless, the point system would readily support the addition of such criteria.

Table 3. Point based variant classification categories

Category	Point ranges
Pathogenic	≥ 10
Likely Pathogenic	6 – 9 ¥
Uncertain	0 – 5
Likely Benign	-1 – -6 ¥
Benign	≤ -7

¥ Operationally, the prior probability should be understood to be infinitesimally greater than 0.10. This has two effects. First, it makes the posterior probability of the likely pathogenic categories infinitesimally greater than 0.90, so that the likely pathogenic classifiers work properly. Second, it enforces a requirement for some evidence of benign effect for sequence variants to be classified as likely benign. One could also argue that the point threshold for likely benign should really be -2. This would match the ACMG rule "Likely Benign (ii)" rather than the simple numerical requirement that the posterior probability be <0.10.

Figure Legends

Figure 1. Schematic relationship among Bayes' rule, the qualitative ACMG/AMP variant classification guidelines, the Bayesian formulation of those guidelines, and the point system derived here.

Data Sharing

Data sharing not applicable – no new data generated

Acknowledgments

S.V.T is supported in part by R01 CA121245. K.M.B. and the Huntsman Cancer Institute's Biostatistics Core are supported in part by P30 CA042014. L.G.B. is supported by NHGRI grant HG200359.

References

Barnetson R. A., Cartwright N., van Vliet A., Haq N., Drew K., Farrington S., ... Dunlop MG. Classification of ambiguous mutations in DNA mismatch repair genes identified in a population-based study of colorectal cancer. Hum Mutat, 29(3), 367–374, (2008). doi:10.1002/humu.20635. PMID: 18033691.

Plon, S. E., Eccles, D. M., Easton, D., Foulkes, W. D., Genuardi, M., Greenblatt, M. S., . . . Tavtigian, S. V. (2008). Sequence variant classification and reporting: recommendations for improving the interpretation of cancer susceptibility genetic test results. Hum Mutat, 29(11), 1282-1291. doi:10.1002/humu.20880. PMID: 18951446.

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., . . . Rehm, H. L. (2015). Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med, 17(5), 405-423. doi:10.1038/gim.2015.30. PMID: 25741868.

Riggs, E. R., Andersen, E. F., Cherry, A. M., Kantarci, S., Kearney, H., Patel, A., . . . ACMG. (2020). Technical standards for the interpretation and reporting of constitutional copy-number variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics (ACMG) and the Clinical Genome Resource (ClinGen). Genet Med, 22(2), 245-257. doi:10.1038/s41436-019-0686-8. PMID: 31690835.

Tavtigian, S. V., Greenblatt, M. S., Harrison, S. M., Nussbaum, R. L., Prabhu, S. A., Boucher, K. M., . . . ClinGen Sequence Variant Interpretation Working Group (ClinGen, S. V. I. (2018). Modeling the ACMG/AMP variant classification guidelines as a Bayesian classification framework. Genet Med, 20(9), 1054-1060. doi:10.1038/gim.2017.210. PMID: 29300386.