Evaluation of Bayesian Classification Framework on the Variant Classification of Hereditary Cancer Predisposition Genes

Authors: Mohammad K. Eldomery[1,*,#,], Jamie L. Maciaszek[1,*], Taylor Cain[1], Victor Pastor Loyola[2], Suraj Sarvode Mothi[3], David A. Wheeler[2], Li Tang[3], Lu Wang[1,] Jeffery M. Klco[1], Patrick R. Blackburn[1]

Author affiliations

[1]Department of Pathology, St. Jude Children's Research Hospital, Memphis, Tennessee 38105, USA.

[2]Department of Computational Biology St. Jude Children's Research Hospital, Memphis, Tennessee, USA.

[3]Department of Biostatistics, St. Jude Children's Research Hospital, Memphis, Tennessee, USA.

*Contributed equally (Mohammad.Eldomery@stjude.org) and Jamie L. Maciaszek (Jamie.maciaszek@stjude.org)

[#]Correspondence should be addressed to:
Mohammad K. Eldomery, MD

Assistant Member of Molecular Pathology

Department of Pathology

St. Jude Children's Research Hospital

Email: mohammad.eldomery@stjude.org

A full list of author affiliations appears at the end of the paper.

Abstract

Purpose: To assess the differences in the variants classifications using the ACMG/AMP 2015 guidelines and the Bayesian point-based classification system (here referred to as point system) in 115 hereditary cancer predisposition genes and explore the utility of the point system in variant reanalysis.

Methods: Germline variant classifications from 721 pediatric patients were evaluated using the two scoring systems and compared with our reported classification.

Results: 2376 unique variants were identified. The point system exhibited a propensity to decrease the rate of variants of unknown significance (VUS) to 15% compared to 36% by the ACMG/AMP 2015 (Cochran-Armitage with Z-score of -16.686; p-value < 0.001). This reduction in VUS rate is attributed to 1) single benign supporting evidence (12%); 2) single benign strong evidence (4%), each of which independently could downgrade a VUS to likely benign in the point system; and 3) resolving conflicting evidence or evidence not recognized by the ACMG/AMP 2015 guidelines (5%). Examination of the point system scores of the reported VUS (28%) facilitated tiering and prioritizing the variants for reanalysis.

Conclusion: The point system facilitates the reduction of the VUS rate and provides a systematic way for periodic reanalysis of VUS in hereditary cancer predisposition genes.

Keywords

Variant classification, ACMG/AMP 2015 guidelines, Bayesian point-based classification system, hereditary cancer predisposition

INTRODUCTION

Genetic variant curation and analysis is an essential aspect of the practice of genomic medicine. Efforts led by the American College of Medical Genetics and Genomics and the Association for Molecular Pathology (ACMG/AMP) provided a framework for efficient germline variant curation in 2015 (here referred to as ACMG/AMP 2015 guidelines).[1] Despite the success of this framework, discordance in variant classification between laboratories remained and it was evident that additional clarification and refinement of these guidelines was needed.[2] Several studies have documented the necessity of periodic and systematic reanalysis of germline variants to improve patient care[3-5] and to accelerate disease-gene discovery in research settings.[6]

In 2018, the Clinical Genome Resource (ClinGen) Sequence Variant Interpretation (SVI) Working Group (https://clinicalgenome.org/working-groups/sequence-variant-interpretation/) presented a Bayesian classification framework that aims to provide a quantitative approach to variant classification.[7] Subsequently, Tavtigian and colleagues proposed a model to transform the Bayesian framework to a point-based classification system (referred to as the point system) to facilitate the integration of the Bayesian framework with the ACMG/AMP 2015 guidelines.[8] Building on these developments, the objective of our study is to assess the concordance in germline variant classifications using the ACMG/AMP 2015 guidelines and the point system utilizing data from a cohort of pediatric cancer patients evaluated by a 115-gene cancer predisposition germline panel. We assess the degree of concordance between

the two scoring methods and identify discrepancies and potential limitations of each scoring system. Furthermore, we explore the utility of the point system in tiering the variants and the systematic reanalysis of genetic results.

MATERIALS AND METHODS

Samples and data collection

Variant analysis was performed on reported germline variant calls from 721 patients who underwent hereditary cancer panel testing between June 2021 and May 2023. Dual Genome Sequencing (GS) and Exome Sequencing (ES) were performed on germline samples (peripheral blood or skin biopsy) to assess for single nucleotide variants, insertions, deletions, and copy number variants in 115 genes associated with the risk of inherited cancer syndromes (see Supplementary Table 1 for complete gene list and preferred transcripts). Sequencing and bioinformatics analysis were performed as previously described.[9,10]

Within our cohort, most evidence codes were applied in accordance with the strength originally proposed in the ACMG/AMP 2015 guidelines. However, a subset of variants had modified strengths either based on our committee review or recommendations from the ClinGen SVI Working Group. As such, results were split into variants with the original strengths according to ACMG/AMP 2015 guidelines and variants with modified strengths in order to make direct comparisons. The ClinGen-modified evidence codes include: BA1, PVS1, PS2/PM6, PS3/BS3, PM2, PM3, and PP3/BP4. Since the BA1-stand-alone evidence code was not included in the point system, variants meeting the criteria for BA1-stand-alone were excluded.[7]

Dataset assembly and scoring methodologies

To assemble our dataset, we queried our internal clinical database and retrieved variants classified by the Molecular Pathology/Clinical Genomics laboratory at St. Jude Children's Research Hospital between June 2021 and May 2023. The corresponding evidence codes and strength of each variant were scored using the ACMG/AMP 2015 guidelines and the point system. Eventually, we correlated with our reported classification. Our analysis did not include copy number variations.

Statistical analysis

Cohen's kappa statistic was used to assess pairwise concordance between the variant classification generated by the ACMG/AMP 2015 guidelines and the point system, as well as to compare our reported classification against each model. The variant classification was binned by comparing the variant of unknown significance (VUS) across the ACMG/AMP 2015, the point system and our reported classification against the remaining categories (Benign 'B', Likely Benign 'LB', Likely Pathogenic 'LP', or Pathogenic 'P') grouped into a single category labeled as 'other categories'. This facilitated the utilization of two-by-three tables for testing and contrasting trends of association using the Cochran-Armitage test and McNemar paired proportion tests. All statistical analyses used a two-sided alpha level of 0.05 and were performed using the R software (version 4.3.1 (2023-06-16 ucrt)).

RESULTS

Clinical Cohort Profile and Variant Distribution

Approximately 34% (244/721) of patients had a solid tumor primary diagnosis, ~33% (240/721) had central nervous system (CNS) tumors, and the remaining ~33% (237/721) had various hematolymphoid tumors (Supplementary Figure 1). One hundred and three P/LP variants in autosomal dominant genes and 18 P/LP heterozygous variants in genes associated with an autosomal recessive inheritance pattern (Supplementary Figure 1) were detected. The overall prevalence of P and LP variants in our cohort is ~17% (121/721), which is comparable to previous pediatric cancer predisposition cohorts.[10]

After excluding duplicate germline variants and eliminating variants meeting BA1-stand-alone criteria, a total of 2376 unique variants were identified in our clinical cohort. These comprised 1288 missense, 660 silent, 309 splice-associated variants, 45 frameshift indels, 41 in-frame indels, 32 nonsense, and one variant in a non-coding gene (*TERC*). The 309 splice-associated variants comprised 11 canonical splice site variants and 298 in splicing regulatory regions extending from the last three bases of an exon to the first base of the following exon, not including the canonical splice site.

Comparative Analysis of Variant Classification System

The analysis of the variability in genetic variant classifications among ACMG/AMP 2015, the point system, and our reported classification reveals distinct distributions within the various variant categories (Figure 1a). Evaluation of overall pairwise concordance using Cohen's Kappa statistic identified the lowest agreement

between the ACMG/AMP 2015 guidelines and the point system, with a Kappa coefficient of 0.63 (z = 49.7, p-value < 0.001; 95% CI: 0.59-0.68). A higher level of agreement was observed between the ACMG/AMP 2015 classification and our reported classification, with a Kappa coefficient of 0.82 (z = 58.7, p-value < 0.001; 95% CI: 0.77-0.86), while the agreement between the point system and our reported classification exhibited a Kappa coefficient of 0.72 (z = 53.7, p-value < 0.001; 95% CI: 0.67-0.76). Trend analysis using the Cochran-Armitage test demonstrated a significant reduction in VUS when applying the point system by shifting the variant classification towards LB/B and LP/P (Z = -16.686, p-value < 0.001) (Figure 1b; Supplementary Table 2).

Analysis of Discordant Results

To elucidate the main patterns of disagreement between the scoring systems and identify potential reasons for the decrease in VUS by the point system, an in-depth analysis was conducted on discordant calls between the ACMG/AMP 2015 guidelines and the point system, where the lowest agreement was noted. Among the 2376 unique variants analyzed, 23.5% (n=559) exhibited discordant classifications (Supplementary Table 3). These discordant variants were further divided into those that retained the originally defined evidence code strength from the ACMG/AMP 2015 guidelines (n= 445) and those that were assigned a modified evidence code strength (n = 114).

A) Variants with Original ACMG Evidence Codes

The predominant impact was on VUS, but it also affected P and LP variants (Figure 2). Among VUS, ~18% (422/2376) variants were subsequently downgraded to

LB, ~0.8% (18/2376) variants were reclassified as B, and notably, one variant was upgraded to LP. Furthermore, two variants initially categorized as LP were upgraded to P. Conversely, the application of the point system resulted in two variants being downgraded from P to LP.

Analysis of the evidence codes/categories, strengths, and corresponding point system scores within the discrepant categories (Supplementary Figure 2; Supplementary text) revealed 12% (289/2376) of variants with a score of -1, the minimum threshold in the point system for calling LB. The BP4 supporting evidence code was applied to most variants (283/2376), including 275 coding substitutions and eight intronic variants. Our application of PP3/BP4 criteria was dependent on the REVEL, where a value ≤0.290 is applied for BP4 (-1 point), ≥0.644 is applied for PP3 (1 point), and 0.3-0.63 would be inconclusive (0 point).[11] A total of 275 single nucleotide variants received BP4 based on their REVEL score, while six variants had an inconclusive score between 0.31 and 0.43.

Approximately 6% (151/2376) of VUS, according to ACMG/AMP 2015, had a score ≤ -2 using the point system. These variants were designated as VUS by ACMG/AMP 2015 guidelines because they did not meet a specific criterion and/or exhibited contradictory benign and pathogenic criteria. However, the variants achieved the minimum threshold score of -2 or -7 for classification as LB and B by the point system, respectively. Approximately 4% (98/2376) of variants were seen at score -4, driven by one benign strong evidence type (i.e., BS1, BS2, or BS3 each provides -4 points), with BS1 alone applied alone 87 times. The remaining ~2% (53/2376) exhibited contradictory benign and pathogenic criteria (Supplementary Figure 2).

The point system resulted in one variant, *CHEK2* NM_007194.4:c.190G>A (p.Glu64Lys), being upgraded from VUS to LP (score from 6 to 9 points). The *CHEK2* variant received PS3 Strong (4 points), PS4 Strong (4 points) and BP4 Supporting (-1 point). Additionally, two variants were upgraded from LP to P (score equal to or greater than 10 points). Those included *BLM* NM_000057.4:c.2250_2251insAAAT (p.Leu751LysfsTer25) and *POT1* NM_015450.2:c.1087C>T (p.Arg363Ter), each received a PVS1 Very Strong (8 points) and one moderate evidence code (either PM3 or PM5 provides 2 points), resulting in a final score of 10 points. Conversely, due to a final score of 9 points, driven by two strong evidence codes (PS3 Strong and PS4 Strong) and one supporting evidence code (either PP1 Supporting or PP3 Supporting provides 1 point), the point system downgraded two variants, *MUTYH* NM_001128425.1:c.1187G>A (p.Gly396Asp) and *MITF* NM_198159.3:c.1255G>A (p.Glu419Lys), from P to LP (9 points).

B) Variants With Modified ACMG Evidence Codes

Approximately 5% of discordant variants (114/2376) received at least one modified evidence code that differed from the original ACMG/AMP 2015 guidelines (Figure 3). The modified evidence code was either due to ClinGen recommendations or made by our internal committee (Supplementary Figure 3; Supplementary text). Approximately 3% (65/2376) were classified as VUS by the ACMG/AMP 2015 guidelines due to contradictory criteria between benign and pathogenic evidence codes or evidence strength combinations not recognized in the original guidelines (e.g., PM3 Very Strong). Of these, 51 variants were downgraded to LB by the point system,

including 32 variants with a score of -1 achieved by two benign supporting evidence codes with one pathogenic supporting (e.g., BP4 Supporting; BP7 Supporting; PM2 Supporting) (Supplementary Figure 4a). The 14 VUS called by the ACMG/AMP 2015 included three that were upgraded to P and 11 variants that were upgraded to LP (Supplementary Figure 4b). PM3 Very Strong (8 points) evidence was applied in 4 variants based on the ClinGen guidance on variant phasing (https://clinicalgenome.org/working-groups/sequence-variant-interpretation/) and was combined with PVS1 Very Strong in three variants and therefore regarded as P. PM3 Very Strong was also combined with a pathogenic supporting evidence code (1 point) in one variant and therefore met the criteria for LP by the point system. The remaining upgraded LP variants comprised eight variants with PVS1 Very Strong with one pathogenic supporting evidence code; two variants were upgraded based on a combination of different pathogenic evidence codes that are not recognized by the ACMG/AMP 2015 guidelines.

Interestingly, a variant in *BRCA2* NM_000059.3:c.9699_9702del that was classified as P by ACMG/AMP 2015 was subsequently downgraded to LP using the point system based upon the following evidence codes totaling 8 points: PS4 Strong; PVS1 Strong. The *BRCA2* variant is in the last exon and is not predicted to trigger nonsense-mediated-decay (NMD); therefore, PVS1 Very Strong was downgraded to PVS1 Strong according to the recommendations from Abou Tayoun and colleagues.[12]

Approximately 2% (47/2376) of variants with modified evidence codes applied that were categorized as LB based on ACMG/AMP 2015 were further downgraded to B by the point system. With the exception of BA1, the original ACMG/AMP 2015

recommendations typically required ≥ 2 strong evidence codes for benign and either one strong and one supporting evidence code or ≥ 2 supporting evidence codes for LB. In our cohort, 45 unique variants with a score of -7 explained the downgrade from LB to B using the point system framework (Supplementary Figure 4c).

Comparing the Point System with Reported Classification

Overall, the point system is associated with a ~21% reduction in the number of VUS compared to the ACMG/AMP 2015. This reduction is due to single benign supporting evidence (~12%) or single benign strong evidence code (~4%) being applied, as well as resolving conflicting evidence codes or inclusion of evidence not recognized by the ACMG/AMP 2015 guidelines (~5%) (Figure 4). Subsequent analysis was performed to explore the impact of those patterns on our reported variants classification. Among 321 unique variants receiving a point system score of -1, ~12% (292/2376) were reported as VUS (with BP4 in 278) and 29 variants were reported as B/LB. Approximately 4% (97/2376) of variants with a score of -4 driven by a single benign strong evidence code were reported as LB (Supplementary Figure 5). Of the ~4% of VUS with conflicting evidence (as determined by the ACMG/AMP 2015 guidelines), 86 were reported as B/LB variants with a point system score range of -1 to -11 and 14 variants were reported as P/LP (Supplementary Table 3; Supplementary text). Less than 1% (14/ 2376) were reported as VUS and one variant was upgraded during our analysis.

VUS Tiering Using the Point System

The point system facilitated the tiering of the reported VUS by examining the associated point system scores and variant type (Figure 5; Supplementary Table 3). The point system demonstrated concordance in classifying 351 reported VUS. In this category, missense variants are more prevalent, with a high frequency observed 0- and 1-point system scores (121 and 129 occurrences respectively), but with a notable decrease at lower point system scores with 2, 3, 4, and 5 detected at frequencies of 56, 4, 3, and 2, respectively. For variants with a point system score of 0, both a pathogenic and benign supporting evidence type were applied, leading to a net score of 0, with the most commonly occurring combinations being BP4 Supporting with PM2 Supporting (n=101) and BP4 Supporting with PP2 Supporting (n=13). The majority of VUS with a score of 1 harbored a single pathogenic supporting evidence (e.g., PP3 Supporting, n = 67; PM2 Supporting, n = 41). Variants with a score of 2 were frequently associated with two pathogenic supporting evidence codes (e.g., PM2 Supporting; PP3 Supporting). Seventeen variants with scores of 3 and above had either several pathogenic supporting evidence codes or evidence codes with moderate or strong strength.

Two variants in *ATM* and *POLE* with scores of 9 and 12 were reported as VUS. During our reevaluation, *ATM* NM_000051.3:c.7135C>G (p.Leu2379Val) was upgraded to LP. The variant was initially classified as VUS with a score of 9. It has a ~0.038% subpopulation frequency in gnomAD v2.1.1, causes partial exon 49 skipping, and is reported in individuals with breast/ovarian cancer.[13] While this variant meets the point system classification criteria for LP, the initial version (v1.1) of the *ATM* classification rules specified by ClinGen restricted LP calls for one very strong with PM2 supporting (internal communication). As such, the variant was initially regarded as VUS.

Importantly, the v1.2 *ATM* curation guidelines now permit this combination for LP and the reported variant was upgraded to LP accordingly. A truncating variant was identified in *POLE* with a score of 12; however, the variant was reported as a VUS with respect to *POLE*-associated polyposis (Supplementary text).

In addition to 292 variants with a score of -1 reported as VUS, 13 variants with lower scores were regarded by the point system as B/LB, of which seven had contradictory evidence, while the remaining six variants were re-reviewed during our study and the classification remained unchanged.

DISCUSSION:

We compared the scoring rules of qualitative ACMG/AMP 2015 guidelines and the quantitative Bayesian point-based system and their respective effects on germline variant classification. Our analysis highlighted the tendency of the point system to decrease the overall number of VUS by either downgrading to LB/B or upgrading to LP/P. We observed a discrepancy in variant scoring between the two methodologies in ~23.5% of the variants analyzed. Among discrepant variants, ~21% were attributed to a decrease in VUS rates, driven by three primary modifications: 1) ~12% reduction in VUS with a point system score of -1, primarily due to a single benign supporting evidence code (most notably BP4 Supporting in 283 variants); 2) ~4% reduction in VUS with a point system score of -4 linked to a single benign strong evidence code (e.g., BS1-BS3 in 98 variants); and 3) ~5% reduction in VUS with conflicting evidence codes or criteria not acknowledged by the current ACMG/AMP 2015 guidelines. The remaining ~2.5% of discrepant variants were related to the reclassification from LB to B by the point system

and, to a lesser extent, changes between LP and P classifications. Approximately 8% of variants classified by the point system aligned with our reported findings, either by resolving conflicting/contradictory evidence (4%) or through a single benign strong evidence code being applied (4%).

To elucidate the primary causes of the discrepancies observed between these two classification systems, it is essential to consider the seminal insights from the research undertaken by Tavtigian and colleagues.[7] Analysis of a prior probability (Prior_P) of 0.10, with odds of pathogenicity of 350, produced a posterior probability in alignment with the ACMG/AMP 2015 criteria, with two notable exceptions: instances where two pathogenic strong classifications are interpreted as likely pathogenic (LP) and cases where one very strong and one moderate piece of evidence are considered pathogenic (P). Subsequent inference of the point-system scale derived from a Prior_P of 0.1 resulted in a single piece of benign evidence or equivalent conflicting evidence with a point score of -1 is capable of reclassifying a genetic variant from a VUS to LB – a critical difference from the criteria set forth in the ACMG/AMP 2015 guidelines.[8]

Approximately 12% of discordant classifications were associated with a point system score of -1. While this can raise some uncertainty regarding the impact of the point system, a detailed examination of the reported missense VUS and associated point system scores, ranging from 0 to 5, revealed that the 250 of VUS had scores of 1 or 2. In order to change a classification from LB (with a score of -1) to LP (with a score of 6), a minimum of 8 pathogenic points would be required. While this might seem to be a rare situation for a variant to acquire eight pathogenic points, missense or synonymous variants that might activate a cryptic splice site, and alter the coding frame

of proteins, could be assigned PVS1 Very Strong (8 points) evidence code, potentially resulting in reclassification from LB to LP. It may, therefore, become necessary for clinical laboratories to conduct a systematic assessment of all substitutions/missense variants by employing computational predictions that encompass not only protein structure but also potential changes in splicing. Such a comprehensive approach is vital to ensure the accurate application of the BP4 criterion in the variant classification process with the point system. The point system score of -1 might be helpful in resolving conflicting evidence such as BP4 Supporting, BP7 Supporting, and PM2 Supporting. In the context of cancer predisposition and incidental findings, it might be challenging to classify a variant as LB based on a single benign supporting evidence code due to the paucity of comprehensive clinical information.[14] Conversely, in non-cancerous Mendelian disorders, with a detailed and comprehensive phenotypic evaluation, the application of a threshold that allows the acceptance of a single piece of benign supporting evidence for the reclassification of VUS to LB may be feasible.

The modeling of the ACMG/AMP 2015 classification guidelines as a Bayesian classification framework has yielded other pivotal insights on variants with a single piece of strong benign evidence with a score of -4 (i.e., BS1-BS3). Tavtigian and colleagues uncovered a distinct pattern within the ACMG/AMP 2015 criteria where one criterion from a stronger evidence category could equivalently be replaced by two criteria from the next lower strength category or by four criteria from two subsequently lower categories.[7] In the ACMG/AMP 2015, the LB can be achieved by at least two supporting benign evidence types; within the point system, one strong piece of benign evidence (e.g., BS1-4) would be equivalent to four pieces of supporting benign evidence

(BP1-BP7), thereby positing that a single strong evidence type could suffice to reclassify a variant from VUS to LB.

The point system resolved ~5% of VUS due to conflicting evidence codes or with a combination of evidence codes that are not directly recognized by ACMG/AMP 2015 guidelines. While the ACMG/AMP 2015 guidelines recommend that experts should use their judgment when they encounter conflicting/contradictory evidence codes, the ability of the point system to address these conflicts suggests a path for large-scale automation of variants.[1,7] Furthermore, a combination of one very strong and one supporting evidence code can achieve a posterior probability greater than 90%, aligning with the criteria to meet LP. The significance of this finding is particularly relevant in light of ClinGen's recommendation to downgrade the strength of PM2 from moderate to supporting. Similarly, combinations like BP4 Supporting, BP7 Supporting, and PM2 Supporting would be LB (-1 point), while it would be VUS (0 point) using PM2 Moderate. Clinical laboratories adopting the latest ClinGen guidelines must be aware of these implications.

The point system can shed light on a systematic way for sub-tiering the variants.[15] In our cohort, reevaluation of the reported VUS (28%) was facilitated by examination of the point system score and associated variant type, which eventually upgraded one variant in *ATM* to LP based on the most updated ClinGen guidelines. A recent comprehensive study revealed the frequency of these VUS among non-cancer patients to be ~32.6% when assessed with multi-gene panels and ~22.2% when evaluated using exome or genome sequencing.[15] While our current pediatric cohort is focused on cancer predisposition genes, we found a VUS frequency of 28% that sits

between the VUS rates reported from broader exome or genome sequencing and those from multi-gene panel testing in Mendelian non-cancer cohorts. Hypothetically, with the application of the point system to our data, the VUS rate drops to 15%, which is lower than even the 18.9% VUS rate observed when trio analysis is used.[15] Future studies to explore the utility of trio analysis with the point system might be warranted to study the effects on VUS rate.

Our analysis has several limitations and we recognize the potential for variability in the interpretation of variants and the application of evidence types across different clinical laboratories, as highlighted by the literature.[2] To mitigate this, our approach involved a direct comparison of the ACMG/AMP 2015 guidelines and the point system, further stratifying the data by original evidence codes and modified ones. Our analysis was concentrated on genes associated with hereditary cancer syndromes and did not address the influence of the point system on genes related to constitutional non-cancer conditions or on the somatic variant curation.[16] Subsequent large-scale studies will be critical for validating our findings and for discerning additional levels of agreement or disparity between the two models.

In conclusion, our study underscores the need to continually refine the scoring systems used in genetic variant classification. The differences in variant classification between the ACMG/AMP 2015 guidelines and the point system are remarkable. By comparing the ACMG/AMP 2015 guidelines against the point system, our study demonstrates that a combination of the point system with new ClinGen recommendations may add precision to variant classification, which could impact clinical decisions and assist with further standardization of variant analysis.

Web Resources

ClinVar, https://www.ncbi.nlm.nih.gov/clinvar/

Genome Aggregation Database (gnomAD), https://gnomad.broadinstitute.org/

REVEL, https://sites.google.com/site/revelgenomics/

ClinGen: https://clinicalgenome.org/working-groups/sequence-variant-interpretation/

Figure 1. A) The overall variant classification rate across the ACMG/AMP 2015, point system and our reported classification. The classification under strict ACMG/AMP 2015 guidelines and point system depicts a 21% difference in the variant of uncertain significance rate (VUS), highlighting a lower level of agreement (Kappa coefficient of 0.63) compared to 8% difference in VUS rate between ACMG/AMP 2015 guidelines and reported classification (Kappa coefficient of 0.82), and 12% difference in VUS rate between point system and reported classification (Kappa coefficient of 0.72). B) Classification trends for ACMG/AMP 2015 and the point system highlight the overall tendency of point system to reduce VUS and increase the number of likely benign variant calls (Cochran-Armitage test: $Z = -16.686$, p-value < 0.001)

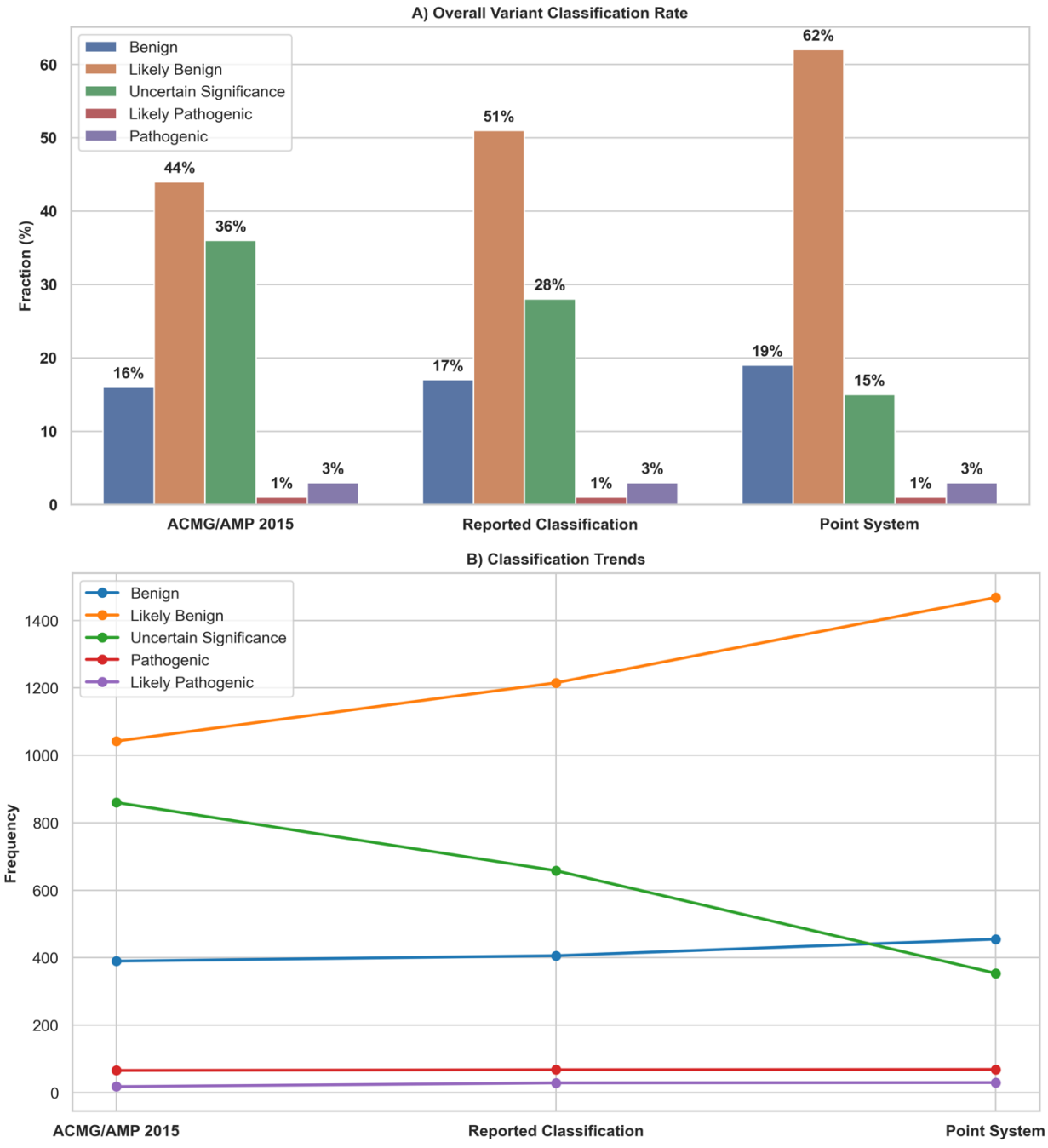


Figure 2. Sankey figure summarizing the classification changes between the ACMG/AMP 2015 guidelines and the point system. In variants classified using the original ACMG/AMP 2015 evidence code strengths (n=445), ~18% (422/2376) of VUS w downgraded to LB using the point system, ~0.8% (18/2376) was downgraded to B,

and one variant was upgraded to LP. Two variants classified as LP were upgraded to P, and conversely, two variants were downgraded from P to LP.

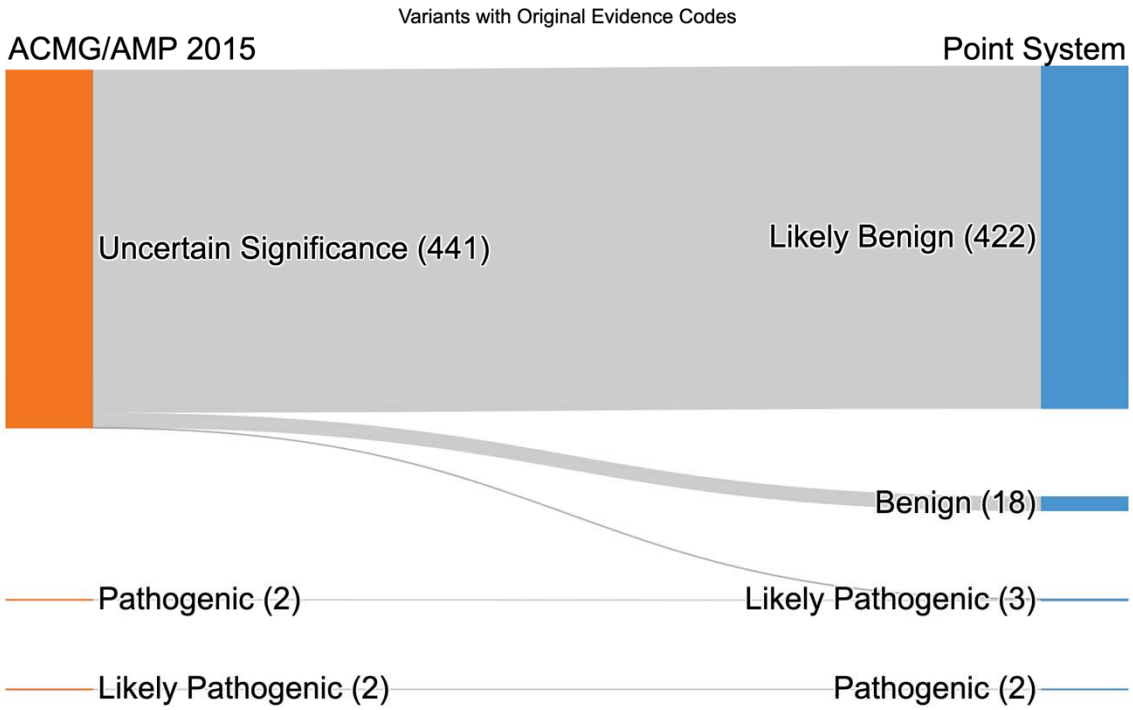


Figure 3. In variants with modified evidence code strengths (n=114), 2% (47/2376) of variants that were classified as LB using ACMG/AMP 2015 were downgraded to B, and 3% (65/2376) of VUS were reclassified as follows: 51 variants were downgraded to LB, 11 variants were upgraded to LP, and three variants were upgraded to P.

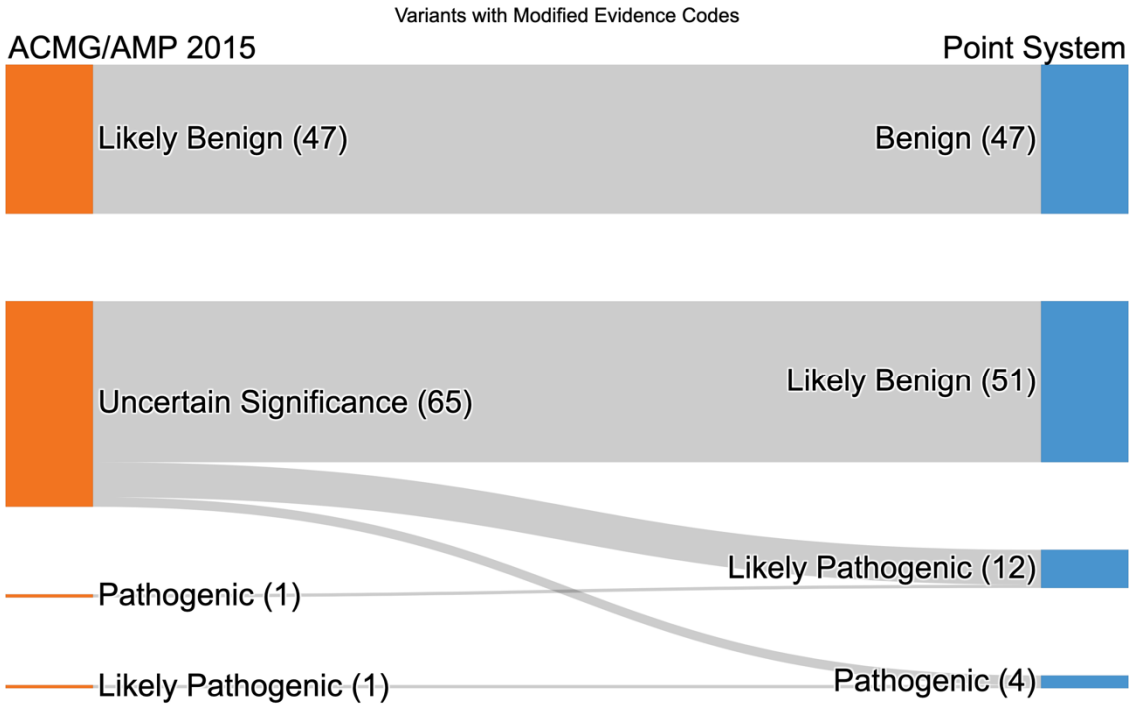


Figure 4. Main changes on the variant classification using the point system compared to ACMG/AMP 2015 guidelines. A discrepancy of 23.5% was observed between the point system and the ACMG/AMP 2015 guidelines. Among these discrepancies, ~21% were attributed to a decrease in the number of VUS, primarily due to ~12% from Single Benign Supporting evidence (mostly BP4 Supporting), ~4% from Single Benign Strong evidence (mostly BS1 Strong), and ~5% from resolving Conflicting or Contradictory criteria according to the current ACMG/AMP 2015 guidelines. The remaining ~2.5% of differences related to changes between LB and B or LP and P classifications (for instance, the application of two Pathogenic Strong Evidence codes leading to a downgrade from P to LP using the point system).

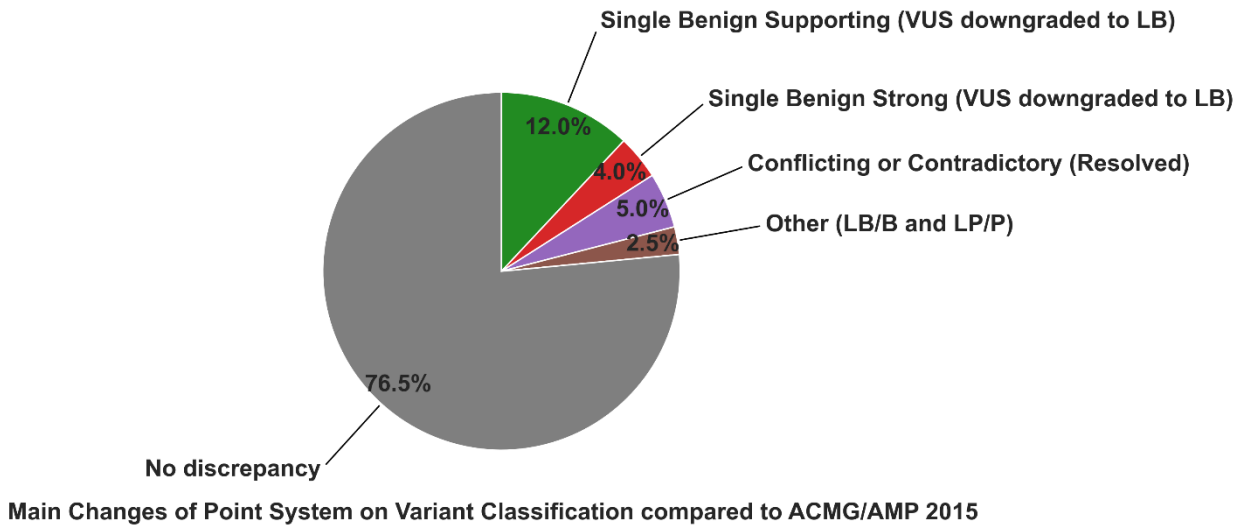
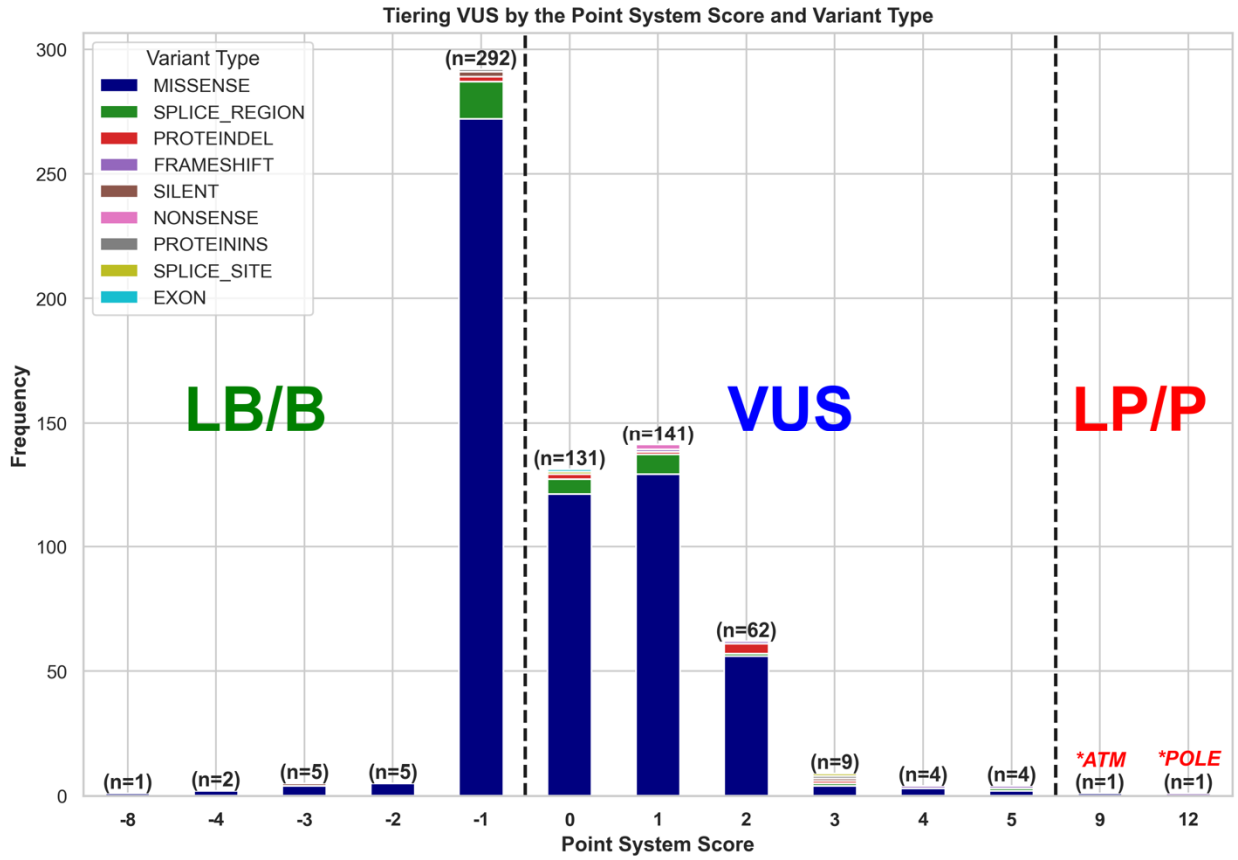


Figure 5. Analysis of VUS using the scoring system reveals that missense variants are predominantly observed at lower scores (0-1), with their frequency decreasing at higher scores (2-5). Dashed lines delineate the VUS (in blue) from those variants classified by the scoring system as likely pathogenic/pathogenic (LP/P) in red for scores greater than five or benign/likely benign (B/LB) in green for scores less than 0. Further analysis identified two variants in the genes ATM and POLE with scores of 9 and 12, respectively, indicated by asterisks and in red. This reevaluation resulted in the classification of the ATM variant as likely pathogenic.



Data Availability

All of the relevant data used in this study has been provided in the tables, or relevant supplementary information.

Acknowledgments

We thank all the patients and their families at St. Jude Children's Research Hospital (SJCRH) for their contribution to the biological specimens used in this study.

Funding Statement

This work was funded by the American Lebanese and Syrian Associated Charities of SJCRH.

Author Contributions

Conceptualization: M.K.E.; Data curation: M.K.E., J.L.M., P.R.B., T.C., V.P.; Formal analysis: M.K.E., J.L.M., P.R.B., T.C., V.P.; Supervision: M.K.E., J.L.M., P.R.B.; Statistical analysis: L.T., S.S.M.; Writing-original draft: M.K.E., J.L.M., P.R.B.,; Writing-review & editing: all authors.

Ethics Declaration

The study was reviewed by the St. Jude Children's Research Hospital Institutional Review Board, which determined that this project did not meet the criteria for human subjects research and, therefore, did not require full institutional review board approval. All patients consented to clinical analysis and individual reporting of germline sequencing data.

Conflicts of Interest

No conflicts of interest relating to this work are noted by the study authors.

References

1. Richards S, Aziz N, Bale S, et al. 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. 2015;17(5):405-424.
2. Amendola LM, Jarvik GP, Leo MC, et al. Performance of ACMG-AMP Variant-Interpretation Guidelines among Nine Laboratories in the Clinical Sequencing Exploratory Research Consortium. Am J Hum Genet. 2016;98(6):1067-1076.
3. Wenger AM, Guturu H, Bernstein JA, Bejerano G. Systematic reanalysis of clinical exome data yields additional diagnoses: implications for providers. Genet Med. 2017;19(2):209-214.
4. Hiatt SM, Amaral MD, Bowling KM, et al. Systematic reanalysis of genomic data improves quality of variant interpretation. Clin Genet. 2018;94(1):174-178.
5. Deignan JL, Chung WK, Kearney HM, et al. Points to consider in the reevaluation and reanalysis of genomic test results: a statement of the American College of Medical Genetics and Genomics (ACMG). Genet Med. 2019;21(6):1267-1270.
6. Eldomery MK, Coban-Akdemir Z, Harel T, et al. Lessons learned from additional research analyses of unsolved clinical exome cases. Genome Med. 2017;9(1):26.
7. Tavtigian SV, Greenblatt MS, Harrison SM, et al. Modeling the ACMG/AMP variant classification guidelines as a Bayesian classification framework. Genet Med. 2018;20(9):1054-1060.
8. Tavtigian SV, Harrison SM, Boucher KM, Biesecker LG. Fitting a naturally scaled point system to the ACMG/AMP variant classification guidelines. Hum Mutat. 2020;41(10):1734-1737.
9. Rusch M, Nakitandwe J, Shurtleff S, et al. Clinical cancer genomic profiling by three-platform sequencing of whole genome, whole exome and transcriptome. Nat Commun. 2018;9(1):3962.
10. Newman S, Nakitandwe J, Kesserwan CA, et al. Genomes for Kids: The Scope of Pathogenic Mutations in Pediatric Cancer Revealed by Comprehensive DNA and RNA Sequencing. Cancer Discov. 2021;11(12):3008-3027.

11. Pejaver V, Byrne AB, Feng BJ, et al. Calibration of computational tools for missense variant pathogenicity classification and ClinGen recommendations for PP3/BP4 criteria. Am J Hum Genet. 2022;109(12):2163-2177.
12. Abou Tayoun AN, Pesaran T, DiStefano MT, et al. Recommendations for interpreting the loss of function PVS1 ACMG/AMP variant criterion. Hum Mutat. 2018;39(11):1517-1524.
13. Rofes P, Menendez M, Gonzalez S, et al. Improving Genetic Testing in Hereditary Cancer by RNA Analysis: Tools to Prioritize Splicing Studies and Challenges in Applying American College of Medical Genetics and Genomics Guidelines. J Mol Diagn. 2020;22(12):1453-1468.
14. Zhang J, Walsh MF, Wu G, et al. Germline Mutations in Predisposition Genes in Pediatric Cancer. N Engl J Med. 2015;373(24):2336-2346.
15. Rehm HL, Alaimo JT, Aradhya S, et al. The landscape of reported VUS in multi-gene panel and genomic testing: Time for a change. Genet Med. 2023;25(12):100947.
16. Horak P, Griffith M, Danos AM, et al. Standards for the classification of pathogenicity of somatic variants in cancer (oncogenicity): Joint recommendations of Clinical Genome Resource (ClinGen), Cancer Genomics Consortium (CGC), and Variant Interpretation for Cancer Consortium (VICC). Genet Med. 2022;24(5):986-998.