Received: 15.09.2025
Accepted: 03.10.2025
Area of Expertise: Medical Genetics (Excl. Cancer Genetics)

Title: Evaluating ACMG/ClinGen PP3/BP4 recommendations for missense variants in CSNK2A1

Short title: Evaluating ACMG/ClinGen PP3/BP4 recommendations for missense variants in CSNK2A1.

Abstract

Purpose: Assessing variant pathogenicity in genes related to rare genetic disorders is a challenging task. While populational databases aid, additional methods are imperative when those genes are also constrained against variation, i.e. many potential variants are also absent from population databases. Many computational prediction algorithms (*in silico* tools) have been developed considering the protein and amino acid characteristics, and cross species conservation for assessing a variant pathogenicity. Some of those *in silico* tools are widely utilized by clinical and molecular geneticists and endorsed by professional organizations such as ACMG and ClinGen. However, their performance may not be the same on every gene and their variants.

Materials and methods: In this study, the performance characteristics of ACMG/ClinGen endorsed *in silico* tools for pathogenic/likely pathogenic (reported in affected individuals) and benign/likely benign (high population allele frequency) missense variants in *CSNK2A1* are evaluated to identify the most reliable prediction tool(s) in aiding variant pathogenicity assessment.

Results: Among the endorsed in silico tools AlphaMissense is the best predictor for variant pathogenicity followed by MutPred2, VARITY_R, and ESM1b; while REVEL, VEST4, and BayesDel do not seem to be good predictors for PP3. Conversely, REVEL and BayesDel are the most reliable predictors for variant benignity compared to the rest of the predictors.

Conclusion: Although the diagnostic laboratories are recommended to select one *in silico* predictor to utilize genome-wide variant predictions, every gene might benefit their own *in silico* predictor evaluations, even different predictors for pathogenic vs benign predictions might be better utilized.

Keywords: CSNK2A1, molecular genetics, in silico, ACMG, ClinGen.

Makale başlığı: ACMG/ClinGen PP3/BP4 önerilerinin CSNK2A1 missense varyantları için değerlendirmesi.

Kısa başlık: ACMG/ClinGen PP3/BP4 onerilerinin CSNK2A1 missense varyantlari için değerlendirmesi

Öz

Amaç: Nadir genetik bozukluklarla ilişkili genlerdeki varyant patojenitesini değerlendirmek zorlu bir iştir. Popülasyon veri tabanları yardımcı olsa da, bu genler varyasyona karşı kısıtlandığında, yani birçok potansiyel varyant popülasyon veri tabanlarında da bulunmadığında ek yöntemlere ihtiyaç duyulur. Protein ve amino asit özelliklerini ve türler arası korumayı dikkate alan birçok hesaplamalı tahmin algoritması (*in silico* araçlar) geliştirilmiştir ve bir varyant patojenitesini değerlendirmek için kullanılmaktadır. Bu *in silico* araçlardan bazıları klinik ve moleküler genetikçiler tarafından yaygın olarak kullanılmakta ve ACMG ve ClinGen gibi profesyonel kuruluşlar tarafından önerilmektedir. Ancak, bu *in silico* araçların performansları her gen ve varyantları üzerinde aynı olmayabilir.

Gereç ve yöntem: Bu çalışmada, CSNK2A1 genindeki patojenik/olası patojenik (etkilenen bireylerde bildirilen) ve benign/olası benign (yüksek popülasyon allel frekansı) missense varyantlar için ACMG/ClinGen tarafından önerilen *in silico* araçlarının performans özellikleri ve varyant patojenite değerlendirmesindeki en güvenilir in silico aracı/araçlarını belirlemek amacıyla değerlendirilmiştir.

Bulgular: Önerilen *in silico* araçlar arasında AlphaMissense, varyant patojenitesi için en iyi prediktör olup, bunu MutPred2, VARITY_R ve ESM1b takip etmektedir; REVEL, VEST4 ve BayesDel ise PP3 için iyi prediktörler olarak görünmemekle birlikte REVEL ve BayesDel, diğer *in silico* araçlara kıyasla varyant benign prediksiyon için en güvenilir araçlardır.

Sonuç: Tanı laboratuvarlarının genom çapında varyant tahminlerinde kullanmak için yalnızca bir adet *in silico* aracı seçmeleri önerilmekle birlikte, her gen için farklı *in silico* prediktör kullanmak, hatta patojenik ve benign prediksiyonlar için farklı araçlar kullanmak daha doğru sınıflamalara yol açabilir.

Anahtar kelimeler: CSNK2A1, moleküler genetik, in silico, ACMG, ClinGen.

Introduction

Classification of missense variants in a gene related with nonspecific clinical findings (i.e. neurodevelopmental disorders) is a challenging task in the absence of functional studies. Many computational prediction algorithms (*in silico* tools) have been developed to predict whether a given variant may be damaging to the encoded amino acid/protein, hence pathogenic, and they have been formally endorsed for supporting evidence (PP3) in 2015 by American College of Medical Genetics (ACMG) and Association for Molecular Pathology (AMP) [1]. Recently, an updated scaled point system has been proposed and endorsed by ClinGen Sequence Variant Interpretation Working Group, in which nominal classification criteria now correspond to the integers based on their weighted strength levels [2, 3]. In line with that, *in silico* tools were evaluated based on their performance on the high confidence pathogenic variants in ClinVar and tool-specific threshold recommendations were provided to adjust a criterion's strength level [4, 5].

CSNK2A1-related Okur-Chung neurodevelopmental syndrome (OCNDS) [MIM#617062] is an autosomal dominant non-specific neurodevelopmental disorder that was first identified in 2016 in six individuals who were found to carry *de novo* missense (n= 5) and canonical donor splice site (n= 1) variants [6]. Since then, missense variants account for the majority of reported variants in *CSNK2A1* in individuals with OCNDS; of which 90% are located in functional domains/residues rather than randomly distributed across the gene [7]. Although some of those variants are recurrently detected in individuals and confidently classified as pathogenic/likely pathogenic, there are still many novel variants being inquired about by physicians, molecular geneticists, and families. The lack of high throughput functional study settings along with hypotheses for existence of different molecular mechanisms, i.e. loss-of-function and altered substrate specificity, renders the role of *in silico* prediction tools more significant.

In this study, the performance of ACMG/ClinGen endorsed *in silico* prediction tools for the previously reported pathogenic/likely pathogenic and precurated benign/likely benign missense variants were evaluated to assess the utility of proposed thresholds when assessing the pathogenicity of missense variants identified in *CSNK2A1*.

Materials and methods

No private health information was used for this study. The literature and ClinVar reviews were performed manually by a clinical and laboratory geneticist to identify previously reported individuals with checking for cross-referencing to prevent double counting of affected individuals. All previously observed and potential missense variants

in *CSNK2A1* were downloaded from gnomAD v4.1, All of Us, and TOPMed databases to retrieve their population frequencies and REVEL's website, respectively, by using *CSNK2A1* genomic coordinates on hg38 (chr20:472,498-543,790) obtained from the National Center for Biotechnology Information (NCBI) webpage (NCBI ID:1457). The curated and retrieved variants data were converted to VCF format and was run on Ensembl Variant Effect Predictor (VEP) [8] web interface (release 114) on 07/10/2025 to re-annotate the variants with population frequency data and *in silico* prediction scores including precomputed splice AI scores. An independent run with CRAVAT (run.opencravat.org) was also performed to retrieve VEST4 scores and cross-reference the concordance across selected prediction tool scores provided by VEP. dbNSFP v5.2 (https://www.dbnsfp.org) was run to retrieve MutPred2 raw scores. MANE (Matched Annotation from NCBI and EMBL-EGI) select transcript (NM_177559.3) annotations were chosen for final evaluation. BayesDel (noAF), MutPred2, REVEL, VEST4, AlphaMissense, ESM1b, and VARITY_R were compared as recommended for PP3 (Pathogenic supporting) and/or BP4 (Benign supporting) by ACMG/ClinGen-endorsed publications [4, 5, 9-15].

Results

A total of 42 unique missense variants in *CSNK2A1* that meet pathogenic/likely pathogenic (P/LP) classification by ACMG/AMP variant curation guidelines have been reported in the literature and ClinVar (Supplementary Table 1). At least one affected individual has been reported for 28 unique variants in the literature, while the inheritance was not known in one individual for one variant and another variant was detected in affected individuals of a single family (mother and their two affected sons) within ATP/GTP binding loop, a hot-spot region. Four variants within important domains/residues were deposited in ClinVar with affected and *de novo* status provided. Eight variants were deposited in ClinVar with affected status provided but de novo status not provided. Since all of those 8 variants are also located within important domains/residues [6, 7], they were also classified as likely pathogenic.

While four reported missense variants were observed once in heterozygous state across population databases, the most commonly reported c.593A>G p.(Lys198Arg) variant is observed in heterozygous state in six individuals across population databases. The penetrance of OCNDS is complete but shows variable clinical expressivity. While observance of six individuals in population databases can be against pathogenicity, the clinical evidence (>50 published and unpublished individuals) outweigh the population allele frequency evidence for the p.(Lys198Arg) variant's pathogenicity. Furthermore,

OCNDS can follow a milder course in some individuals, and those individuals might have been overlooked in the healthcare system and/or their challenges might have been attributed to other, i.e. socioeconomic, factors. Thus, the allele count threshold for population heterozygote observance is set at five; a conservative threshold given that other variants with these allele counts would also be expected to be identified in diagnostic settings should they also be pathogenic.

None of the missense variants in *CSNK2A1* is seen in homozygous state nor had either total or population specific variant allele frequency over 0.5% on gnomAD v4.1; hence BA1 (>5% population allele frequency) is not met for any variant. A total of 299 missense variants has been observed across population databases at least once. All variants on gnomAD v4.1 website were manually re-checked for not having quality and/or annotation flags, for passing filtering parameters, and not having allele bias on provided IGVs. A total of 47 variants, excluding the p. (Lys198Arg) variant, are observed in heterozygous state in at least five individuals across population databases, and were pre-curated as likely benign (LB) (Supplementary Table 1). The remaining variants with less than five heterozygotes across the population databases are pre-curated as Variants of Uncertain Significance (VUS) and were not included in downstream analyses.

Of the P/LP missense variants, AlphaMissense provided strong (PP3_S) level pathogenic prediction for 40/42, moderate-strong (PP3_MS) for 1/42, and moderate (PP3_M) for 1/42 variant; MutPred2 provided PP3_S level prediction for 32/42 variants, PP3_MS for 3/42, PP3_M for 4/42, and no prediction score for 3/42 variants; VARITY_R provided PP3_S level prediction for 12/42, PP3_MS for 21/42, PP3_M for 8/42, and supporting (PP3) for 1/42 variant; ESM1b provided PP3_MS for 34/42, PP3_M for 7/42, and indeterminate for 1/42 variant; REVEL provided PP3_S for 8/42, PP3_MS for 3/42, PP3_M for 2/42, PP3 for 13/42, and indeterminate for 16/42 variants; VEST4 provided PP3_S for 3/42, PP3_MS for 13/42, PP3_M for 8/42, PP3 for 4/42, indeterminate for 11/42, BP4_M for 1/42, and BP4 for 2/42 variants; and BayesDel provided PP3_S for 2/42, PP3_MS for 9/42, PP3_M for 6/42, PP3 for 10/42, indeterminate for 13/42, and BP4 for 2/42 variants (Figure 1). Interestingly, only MutPred2 provided PP3_S level prediction for the most common p. (Lys198Arg) variant.

Of the LB missense variants, AlphaMissense provided moderate-strong (BP4_MS) level benign prediction for 6/47, moderate (BP4_M) level prediction for 11/47, supporting (BP4) for 7/47, indeterminate for 16/47, PP3 for 3/47, PP3_MS for 2/47, and PP3_S for 2/47 variants; REVEL provided BP4_MS level prediction for 4/47, BP4_M for 23/47, BP4 for 8/47, indeterminate for 10/47, and PP3 for 2/47 variants; BayesDel provided BP4_MS for 3/47, BP4_M for 11/47, BP4 for 20/47, indeterminate for 9/47, and PP3 for 4/47

variants; ESM1b provided BP4_M for 6/47, BP4 for 19/47, indeterminate for 12/47, PP3 for 6/47, PP3_M for 2/47, and PP3_MS for 2/47 variants; VARITY_R provided BP4_S for 2/47, BP4_MS for 2/47, BP4_M for 4/47, BP4 for 5/47, indeterminate for 20/47, PP3 for 6/47, PP3_M for 6/47, PP3_MS for 1/47, and PP3_S for 1/47 variant; MutPred2 provided BP4_M for 5/47, BP4 for 8/47, indeterminate for 11/47, PP3 for 4/47, PP3_M for 10/47, PP3_MS for 4/47, and PP3_S for 5/47 variants; and VEST4 provided BP4_M for 7/47, BP4 for 9/47, indeterminate for 22/47, PP3 for 7/47, PP3_M for 1/47, PP3_MS for 1/47 variant (Figure 2).

Discussion

While utilization of next generation sequencing technologies has expanded our understanding and knowledge about genetic disorders, it has also brought a so-called 'VUS (variant of uncertain significance) problem', where the number of reported VUS per a molecular diagnosis establishing pathogenic/likely pathogenic variant has increased exponentially. Evolution of variant curation guidelines towards a more conservative approach for some criteria contributed to this increase in the VUS reporting. For example, while a *de novo* ultra-rare variant in an individual was able to be classified as likely pathogenic per ACMG 2015 guidelines [1] by meeting PS2 and PM2 criteria, the iterative recommendations by ClinGen Sequence Variant Interpretation Working Group (https://clinicalgenome.org/working-groups/sequence-variant-interpretation/) have rendered the weight of these criteria lower and more criteria are needed to reach a likely pathogenic and pathogenic classification. This has ramifications for reliance on other criteria. Although functional studies are the optimal evidence to classify a variant as damaging, it is not feasible in real world experience, particularly when an ultra-rare variant is detected for the first time in the tested individual. And, in the absence of proximity criteria (PM1 and PM5), the remaining most important criterion becomes *in silico* tool prediction (PP3), in addition to inheritance (PS2) and population frequency (PM2) criteria, in classification of a variant.

Consulting *in silico* predictions has a long history in variant curation practice and a lot of work has been put into developing and improving prediction scores by training the computational algorithms with pathogenic/likely pathogenic variants in ClinVar. There are at least 40 *in silico* prediction tools developed so far. Starting with 2015 guidelines, diagnostic laboratories are recommended to select one *in silico* tool and use it genome-wide for any detected variant in any tested individual. In recent recommendations, *in silico* tool specific ranges were also published to adjust the PP3 strength level accordingly. However, this approach might not be applicable to every gene since most of

the training set variants in ClinVar have inherent biases such as the preponderance of hereditary cancer predisposition gene variants and inferred or demonstrated loss-of-function missense variants.

This study was conducted to analyze the performance characteristics of ACMG/ClinGen endorsed *in silico* prediction tools with their recommended ranges for pathogenic/likely pathogenic and likely benign missense variants in *CSNK2A1*, related with Okur-Chung Neurodevelopmental syndrome (OCNDS). Missense variants constitute the majority of the reported variants in *CSNK2A1*-related OCNDS. Furthermore, most of the variants are located in important domains. However, there are also multiple affected individuals reported to carry ultra-rare missense variants in non-domain regions of the gene such as p.(Glu27Lys) and p.(Arg312Trp). The gene itself is also constraint against the missense variation such that the missense Z-score is 5.33 in gnomAD v4.1. Hence it is imperative to have highly reliable additional criteria in variant assessments.

The analysis herein demonstrated that some predictors such as AlphaMissense outperform others in predicting the damaging effect (pathogenicity) of a variant. Interestingly, predictors performing poorly for the damaging effect such as REVEL and BayesDel have better sensitivity for predicting the absence of a damaging effect (benignity) of a variant. Therefore, if a diagnostic laboratory chose one of these predictors for their genome-wide variant assessment as recommended, some variant classifications will be over- or underinflated.

There are limitations to this study. First, the analysis is limited to the endorsed in silico tools by ACMG/ClinGen publications. There might very well be other tools performing better for both pathogenicity and benignity predictions. Second, all the prediction tools undergo iterations, i.e. versions, to improve their sensitivity and specificity, hence a poor-performing tool in this analysis might show improvement with the subsequent versions. Thus, iterative updates to this evaluation are also warranted in the future.

In conclusion, the performance of ACMG/ClinGen endorsed *in silico* tools have been evaluated for the reported pathogenic/likely pathogenic and precurated likely benign missense variants in *CSNK2A1*. The qualitative and semi-quantitative analyses indicate that AlphaMissense has the highest sensitivity for predicting a variant being pathogenic. While REVEL and BayesDel do not perform well for predicting pathogenicity (PP3), they might have superior sensitivity for benignity (BP4). Future re-iterations of this analysis including additional pathogenic/likely pathogenic and benign/likely benign variants and additional predictors are warranted to have dynamic recommendations for *in silico* tool sensitivity and specificity.

Funding: None

Author contributions: V.O. conceived the study, conducted the analyses, wrote and reviewed the manuscript.

Conflict of interest: No conflict of interest was declared by the author.

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. doi:10.1038/gim.2015.30
2. 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. doi:10.1002/humu.24088
3. 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. doi:10.1038/gim.2017.210
4. 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. doi:10.1016/j.ajhg.2022.10.013
5. Bergquist T, Stenton SL, Nadeau EAW, et al. Calibration of additional computational tools expands ClinGen recommendation options for variant classification with PP3/BP4 criteria. *Genet Med*. 2025;27(6):101402. doi:10.1016/j.gim.2025.101402
6. Okur V, Cho MT, Henderson L, et al. De novo mutations in CSNK2A1 are associated with neurodevelopmental abnormalities and dysmorphic features. *Hum Genet*. 2016;135(7):699-705. doi:10.1007/s00439-016-1661-y
7. Bagatelas ED, Khan MM, Rushing GV. OCNDS core features are conserved across variants, with loop-region mutations driving greater symptom burden. *Front Hum Neurosci*. 2025;19:1589897. doi:10.3389/fnhum.2025.1589897
8. McLaren W, Gil L, Hunt SE, et al. The ensembl variant effect predictor. *Genome Biol*. 2016;17(1):122. doi:10.1186/s13059-016-0974-4
9. Feng BJ. PERCH: A unified framework for disease gene prioritization. *Hum Mutat*. 2017;38(3):243-251. doi:10.1002/humu.23158
10. Pejaver V, Urresti J, Lugo Martinez J, et al. Inferring the molecular and phenotypic impact of amino acid variants with MutPred2. *Nat Commun*. 2020;11(1):5918. doi:10.1038/s41467-020-19669-x
11. Ioannidis NM, Rothstein JH, Pejaver V, et al. REVEL: an ensemble method for predicting the pathogenicity of rare missense variants. *Am J Hum Genet*.

2016;99(4):877-885. doi:10.1016/j.ajhg.2016.08.016

12. Carter H, Douville C, Stenson PD, Cooper DN, Karchin R. Identifying Mendelian disease genes with the variant effect scoring tool. *BMC Genomics*. 2013;14(Suppl 3):S3. doi:10.1186/1471-2164-14-S3-S3
13. Cheng J, Novati G, Pan J, et al. Accurate proteome-wide missense variant effect prediction with AlphaMissense. *Science*. 2023;381(6664):eadg7492. doi:10.1126/science.adg7492
14. Brandes N, Goldman G, Wang CH, Ye CJ, Ntranos V. Genome-wide prediction of disease variant effects with a deep protein language model. *Nat Genet*. 2023;55(9):1512-1522. doi:10.1038/s41588-023-01465-0
15. Wu Y, Li R, Sun S, Weile J, Roth FP. Improved pathogenicity prediction for rare human missense variants. *Am J Hum Genet*. 2021;108(10):1891-1906. doi:10.1016/j.ajhg.2021.08.012

	PP3_S	PP3_MS	PP3_M	PP3	Indeterminate /No score	BP4	BP4_M	Total	PP3_All%	BP4_All%
AlphaMissense	95.24%	2.38%	2.38%	NA	NA	NA	NA	100.00%	100.00%	NA
MutPred2	76.19%	7.14%	9.52%	NA	7.14%	NA	NA	100.00%	92.86%	NA
VARITY_R	28.57%	50.00%	19.05%	2.38%	NA	NA	NA	100.00%	100.00%	NA
ESM1b	NA	80.95%	16.67%	NA	2.38%	NA	NA	100.00%	97.62%	NA
REVEL	19.05%	7.14%	4.76%	30.95%	38.10%	NA	NA	100.00%	61.90%	NA
VEST4	7.14%	30.95%	19.05%	9.52%	26.19%	4.76%	2.38%	100.00%	66.67%	7.14%
BayesDel_noAF	4.76%	21.43%	14.29%	23.81%	30.95%	4.76%	NA	100.00%	64.29%	4.76%

Figure 1. Comparison of the prediction subcategories and overall PP3 sensitivity for the ACMG/ClinGen endorsed *in silico* tools

	BP4_S	BP4_MS	BP4_M	BP4	Indeterminate/ No score	PP3	PP3_M	PP3_MS	PP3_S	Total	BP4_All%	PP3_All%
REVEL	NA	8.51%	48.94%	17.02%	21.28%	4.26%	NA	NA	NA	100.00%	74.47%	4.26%
BayesDel_noAF	NA	6.38%	23.40%	42.55%	19.15%	8.51%	NA	NA	NA	100.00%	72.34%	8.51%
ESM1b	NA	NA	12.77%	40.43%	25.53%	12.77%	4.26%	4.26%	NA	100.00%	53.19%	21.28%
AlphaMissense	NA	12.77%	23.40%	14.89%	34.04%	6.38%	NA	4.26%	4.26%	100.00%	51.06%	14.89%
VEST4	NA	NA	14.89%	19.15%	46.81%	14.89%	2.13%	2.13%	NA	100.00%	34.04%	19.15%
VARITY_R	4.26%	4.26%	8.51%	10.64%	42.55%	12.77%	12.77%	2.13%	2.13%	100.00%	27.66%	29.79%
MutPred2	NA	NA	10.64%	17.02%	23.40%	8.51%	21.28%	8.51%	10.64%	100.00%	27.66%	48.94%

Figure 2. Comparison of the prediction subcategories and overall BP4 sensitivity for the ACMG/ClinGen endorsed *in silico* tools

Okur V. Evaluating ACMG/ClinGen PP3/BP4 recommendations for missense variants in CSNK2A1. Pam Med J 2026;19:…-…

Okur V. ACMG/ClinGen PP3/BP4 önerilerinin CSNK2A1 missense varyantları için değerlendirmesi. Pam Tıp Derg 2026;19:…-…

Volkan Okur, M.D. New York Genome Center, United States, e-mail: vokur7@gmail.com (https://orcid.org/0000-0001-6461-0957) (Corresponding Author)