In Silico Variant Effect Prediction for Whole Exome Analysis: Clinical Caveats

#1828

S. Chervitz, J. Anderson, S. Galasinski, K. Chen, J. Harris, R. Chen.
Personalis, Inc. 1330 O'Brien Drive, Menlo Park, CA 94025.

Contact:
steve.chervitz@personalis.com

Background

Computational tools such as SIFT, PolyPhen and others provide in silico prediction of the pathogenicity of missense sequence variants based on sequence information. Some of these tools, such as SIFT and PolyPhen, have been in use since the early 2000's, before the completion of the human genome.

	Disease Present	No Disease
Damaging Call	TP	FP
Benign Call	FN	TN

$PPV = TP/(TP+FP)$
$NPV = TN/(TN+FN)$
$Sn = TP/(TP+FN)$
$Sp = TN/(TN+FP)$
$FPR = FP/(TN+FP)$

Current versions of SIFT & PolyPhen2 have sensitivity 70-90%, but they suffer from low specificity (Flanagan et al 2010; Thusberg et al 2011; Adzhubei et al 2013). Given a 15% false positive rate (FPR) for SIFT and PolyPhen2, such tools would incorrectly flag 19 of 20 missense variants as damaging, for a positive predictive value (PPV) of only 5%, assuming 1% of missense variants are damaging (the true rate of damaging missense mutations is not known). The following table shows the PPV and the proportion of incorrect calls at different mutation prevalence values.

Damaging mutation prevalence	TP	TN	FP (15%)	PPV at 15% FPR	Incorrect calls
1/10,000	10	99990	18748	0.0005	100/100
1/1,000	100	99900	18731	0.0053	99/100
1/100	1000	99000	18563	0.0511	95/100
42379	10000	90000	16875	0.3721	63/100
Sample size(calls) = 100,000					

The clinical utility of these tools, given this performance is questionable at best. Efforts to improve the performance of in silico tools are on-going (Masica & Karchin 2016, Jordan et al 2015), but given their wide use in the clinical community, we need to recognize and understand the limitations of these pathogenicity prediction tools.

Gene	Disease	Prevalence	Penetrance	Inheritance Pattern	Pathogenic Allele Frequency
FBN1	Marfan syndrome	1:5,000	100%	AD	0.0001
GLA	Fabry disease	1:50,000 males	100%	X-linked	0.00002
SOS1	Noonan syndrome	1:1,000 to 1:2,500	incomplete	AD	0.0004-0.001
GAA	glycogen storage disease, type II	1:40,000	100%	AR	0.005
PCCA	propionic acidemia	1:100,000	100%	AR	0.00316
USH2A	Usher syndrome, type 2A	1:20,000	100%	AR	0.00707
NEFL	Charcot-Marie-Tooth disease	1:3,300	Unknown	Multiple	≤0.0174
EDAR	hypohidrotic ectodermal dysplasia	1:5,000	Unknown	Multiple	≤0.0141
PAX9	selective tooth agenesis, type 3	1:1,000	incomplete	AD	0.001
TTN	dilated cardiomyopathy, muscular dystrophy, proximal myopathy, etc.	Unknown	incomplete	Multiple	Unknown

Prevalence estimates of a cross-section of known human genetic diseases. Pathogenic allele frequency data are estimated assuming 50% penetrance for predicted allele frequency data for disorders with incomplete penetrance. Population used for estimate is either worldwide or US. Sources: OMIM, ClinVar

Here, we assess the reliability of SIFT, PolyPhen2 and other variant effect predictions on a set of variants observed in clinical whole exome samples processed by Personalis, Inc. We examine the correlation between the damaging/tolerated calls of the tools compared to a reference set of clinical significance values assessed by genetic counselors using a variety of newer resources. We highlight the risk of misinterpretation of results from commonly used in silico tools, particularly for variants of uncertain significance (VUS), and underscore the need of the clinicians to understand the limitations of such tools when assessing variant results from whole exome analyses.

Results

1) Assessment of mutation effect prediction tools

Predicted effects of mutations from various in silico tools observed in a subset of clinical exome samples processed by Personalis compared to the reported classification of the same variants by genetic counselors and variant scientists.

Reported Class	SIFT		PolyPhen		Mutation Taster		LRT	
	Damaging	Benign	Damaging	Benign	Damaging	Benign	Damaging	Benign
Damaging	7/10	3/10	6/7	1/7	8/8	0/8	8/10	2/10
Benign	1/4	3/4	2/4	2/4	2/4	2/4	1/2	1/2
VUS	3/9	6/9	5/8	3/8	8/9	1/9	3/8	4/8

Results from the in silico tools were available to the reviewers during their review process, along with a variety of other resources, some of which are described in the table below. In silico tool results were obtained from dbNSFP v2.7.

2) Modern resources for variant interpretation

Publicly available resources from the increasing amount of whole genome and whole exome sequence data can provide powerful tools for modern variant interpretation. These resources include allele frequency data from control and patient populations and case observations. Growing use of transgenic animal models as well as the improvement and wider usage of in vitro functional studies will also increase the number of useful resources available for variant classification.

Resource Type	Name	Size	Released	Updated	URL
Allele frequency	ExAC	60,000 exomes	2014	No	exac.broadinstitute.org
Classified Variants	ClinVar	234,000 submissions	2013	Daily	ncbi.nlm.nih.gov/clinvar
Allele frequency	1000Genomes	80M variants, 2500 genomes	2008	Ad hoc	browser.1000genomes.org/index.html
Allele frequency	NHLBI Exome Sequencing Project	6500 exomes	2012	No	evs.gs.washington.edu/EVS/
Published Variants	HGMD	192,000 variants	2007	Ad hoc	www.hgmd.cf.ac.uk/ac/index.php
Gene & Variant Data	OMIM	15,000 genes	1995	Frequently	www.omim.org/
Classified Variants	EmVClass	Unknown	Unknown	Frequently	geneticslab.emory.edu/emvclass/emvclass.php

Conclusion

In silico tools with low positive predictive values offer little clinical utility. The growing amount of publicly available data from individual whole exome and whole genome sequences is increasingly making the use of in silico tools obsolete for variant classification in a clinical context.

Current ACMG guidelines warn about the low specificity of in silico tools which leads them to overpredict deleterious missense changes. The ACMG recommends the use of multiple, complementary algorithms, taking the combined results of these tools together as a single piece of evidence which should not be used as the sole source of evidence to make a clinical assertion (Richards et al. 2015). Furthermore, it is important to understand the positive predictive value of any genetic variant assessment method when interpreting results (Morley 2010).

References

1. Adzhubei, I et al. (2013). Predicting Functional Effect of Human Missense Mutations Using PolyPhen-2. Current Protocols in Human Genetics. 76:7.20:7.20.1–7.20.41.
2. Flanagan et al. (2010). Using SIFT and PolyPhen to predict loss-of-function and gain-of-function mutations. Genet Test Mol Biomarkers 14(5):730.
3. Jordan DM et al. (2015). Identification of cis-suppression of human disease mutations by comparative genomics. Nature. 524(7564): 225–229.
4. Masica DL, Karchin R (2016) Towards Increasing the Clinical Relevance of In Silico Methods to Predict Pathogenic Missense Variants. PLoS Comput Biol 12(5): e100472
5. Morley K. (2010). How well can a screening test predict disease risk? Genomes Unzipped. http://genomesunzipped.org/2010/08/predictive-capacity-of-screening-tests.php
6. Richards, S et al. (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. 2015 May; 17(5): 405–424.
7. Thusberg J et al. (2011). Performance of mutation pathogenicity prediction methods on missense variants. Hum Mutat. 32: 358–368.

Personalis®
Pioneering Genome-Guided Medicine