gibco

24 days
of stem cells

Shape the future of stem cell innovation
October 1 - November 1, 2019

Join us for 24 Days of Stem Cells; a premiere virtual event featuring the latest advances in stem cell research.

This year's format will feature a new hour of cutting edge content every week day starting October 1st. Attend the sessions that are most relevant to your work - at your convenience and at your pace.

During the 24-day long event, you can:

- Access leading scientific presentations from thought leaders around the world
- Watch live training demonstrations from our stem cell experts
- Download key stem cell tools and resources
- Complete weekly challenges to earn points towards certification and prizes

Register today at
www.24daysofstemcells.com

Thermo Fisher
SCIENTIFIC

WILEY

WILEY HGVS HUMAN GENOME VARIATION SOCIETY

Specifications of the ACMG/AMP variant curation guidelines for the analysis of germline *CDH1* sequence variants

Kristy Lee[1] | Kate Krempely[2] | Maegan E. Roberts[3] | Michael J. Anderson[4] |
Fatima Carneiro[5] | Elizabeth Chao[2,6] | Katherine Dixon[7] | Joana Figueiredo[5] |
Rajarshi Ghosh[8] | David Huntsman[7] | Pardeep Kaurah[7] | Chimene Kesserwan[9] |
Tyler Landrith[2] | Shuwei Li[2] | Arjen R. Mensenkamp[10] | Carla Oliveira[5] |
Carolina Pardo[4] | Tina Pesaran[2] | Matthew Richardson[7] | Thomas P. Slavin[11] |
Amanda B. Spurdle[12] | Mackenzie Trapp[1] | Leora Witkowski[13] | Charles S. Yi[2] |
Liying Zhang[14] | Sharon E. Plon[8] | Kasmintan A. Schrader[7] | Rachid Karam[2]

[1]Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina

[2]Ambry Genetics, Aliso Viejo, California

[3]GeneDx, Gaithersburg, Maryland

[4]Invitae Corporation, San Francisco, California

[5]i3S, Istituto de Investigação e Inovação em Saúde & Ipatimup, Institute of Molecular Pathology and Immunology of the University of Porto, Porto, Portugal

[6]University of California Irvine, Irvine, California

[7]University of British Columbia, Vancouver, BC, Canada

[8]Baylor College of Medicine, Houston, Texas

[9]St. Jude Children's Research Hospital, Memphis, Tennessee

[10]Radboud University Medical Centre, Nijmegen, the Netherlands

[11]Department of Medical Oncology and Therapeutics Research, City of Hope, Duarte, California

[12]QIMR Berghofer Medical Research Institute, Brisbane, Australia

[13]Laboratory for Molecular Medicine, Partners Healthcare Personalized Medicine, Cambridge, Massachusetts

[14]Memorial Sloan Kettering, Memorial Sloan Kettering Cancer Center, New York, New York

Correspondence
Rachid Karam, Ambry Genetics, 15 Argonaut, Aliso Viejo, CA 92656.
Email: rkaram@ambrygen.com

Abstract

The variant curation guidelines published in 2015 by the American College of Medical Genetics and Genomics and the Association for Molecular Pathology (ACMG/AMP) provided the genetics community with a framework to assess variant pathogenicity; however, these rules are not gene specific. Germline pathogenic variants in the *CDH1* gene cause hereditary diffuse gastric cancer and lobular breast cancer, a clinically challenging cancer predisposition syndrome that often requires a multidisciplinary team of experts to be properly managed. Given this challenge, the Clinical Genome Resource (ClinGen) Hereditary Cancer Domain prioritized the development of the *CDH1* variant curation expert panel (VCEP) to develop and implement rules for *CDH1* variant classifications. Here, we describe the *CDH1* specifications of the ACMG/AMP guidelines, which were developed and validated after a systematic evaluation of variants obtained from a cohort of clinical laboratory data encompassing ∼827,000 *CDH1* sequenced alleles. Comparing previously reported germline variants that were classified using the 2015 ACMG/AMP guidelines to the *CDH1* VCEP recommendations resulted in reduced variants of uncertain significance and facilitated resolution of variants with conflicted assertions in ClinVar. The ClinGen *CDH1* VCEP recommends the use of

these *CDH1*-specific guidelines for the assessment and classification of variants identified in this clinically actionable gene.

KEYWORDS

ACMG/AMP Variant Curation Guidelines, *CDH1*, ClinGen, ClinVar, hereditary diffuse gastric cancer, lobular breast cancer

1 | INTRODUCTION

Germline pathogenic variants in the *CDH1* gene (NM_004360.3) cause a predisposition to hereditary diffuse gastric cancer (HDGC; MIM# 137215) and lobular breast cancer (LBC; MIM# 137215). The *CDH1* gene is located on chromosome 16q22.1, contains 16 exons, and encodes E-cadherin, a trans-membrane adhesion protein that acts as a tumor suppressor by inhibiting cell invasion (Kourtidis, Lu, & Anastasiadis, 2017). Penetrance estimates derived from pathogenic variant-positive HDGC families indicate that individuals heterozygous for a germline *CDH1* truncating pathogenic variant present an estimated lifetime risk of 56%–70% of developing gastric cancer, and women also have an approximate 42% risk of developing LBC (Hansford et al., 2015). However, it should be noted that these risk estimates could be somewhat inflated as they were based on severely affected families with gastric cancer, which could have introduced an ascertainment bias (Boland & Yurgelun, 2017). *CDH1* is considered a clinically actionable gene since pathogenic variants predispose to lethal cancers for which there are risk-reducing medical management recommendations expected to decrease mortality (van der Post et al., 2015). However, due to the significant risks associated with these recommended medical interventions, which include prophylactic gastrectomy, it is imperative to detect and correctly classify variants in *CDH1* (van der Post et al., 2015).

The implementation of high-throughput next-generation sequencing (NGS) has resulted in a dramatic increase in the identification of germline variants by clinical diagnostic laboratories (LaDuca et al., 2014). The increased availability of genetic testing has also been accompanied by new challenges in sequence interpretation (Plon et al., 2008). In this context, the American College of Medical Genetics and Genomics and the Association for Molecular Pathology (ACMG/AMP) substantially updated the previous guidelines for the interpretation of germline sequence variants (Richards et al., 2015). These recommendations describe the process for classifying variants into five categories—pathogenic (P), likely pathogenic (LP), variant of uncertain significance (VUS), likely benign (LB), and benign (B)—based on criteria using multiple types of evidence (e.g., population data, computational and predictive data, functional data, phenotype/family history data). However, the ACMG/AMP guidelines do not take into consideration gene-specific factors, such as disease-specific incidence and prevalence rates. There are a number of other evidence types which require specification for the gene in question. With this in mind, the National Institute of Health (NIH) funded the Clinical Genome Resource (ClinGen), dedicated to building an authoritative central resource that defines the clinical relevance of genes and variants for use in precision medicine and research (Karam, Pesaran, & Chao, 2015; Nussbaum, Rehm, & ClinGen, 2015; Rehm et al., 2015). Several key goals support ClinGen's overall mission, including optimization of gene-specific clinical annotation, interpretation of germline variants, and sharing of genomic data through a centralized database for clinical and research use. The ClinGen Hereditary Cancer Clinical Domain Working Group selected *CDH1* as a high priority gene and convened a *CDH1* variant curation expert panel (VCEP).

The *CDH1* VCEP was assembled from *CDH1* and variant curation experts with the primary goal of making specifications to the 2015 ACMG/AMP variant interpretation guidelines that are specific to *CDH1*, and then using these specifications to support expert panel variant classification and submission of the curated variants to ClinVar. The *CDH1* VCEP rule specifications to the ACMG/AMP guidelines were developed after a systematic evaluation of a pilot series of test variants obtained from a cohort of clinical laboratory data encompassing aggregated ~827,000 *CDH1* sequenced alleles. The *CDH1* VCEP recommendations for pathogenic and benign criteria include stand-alone and strong benign *CDH1*-specific allele frequency cutoffs, specifications for *CDH1* alleles predicted to result in loss-of-function (LoF), recommendations for the use of computational and functional evidence, and documentation of acceptable HDGC diagnostic criteria. These rules were validated by comparing variant assertions made by biocurators, who were trained in variant curation, but were not *CDH1* gene experts, to variant assertions made by expert members participating in the *CDH1* VCEP.

2 | METHODS

In 2015, the *CDH1* VCEP was convened and consisted of experts selected due to their expertise as clinicians, scientists, and clinical laboratory diagnosticians with regards to HDGC and/or variant classification. As several members of our VCEP are employed by clinical genetic testing laboratories that offer hereditary cancer test panels, we obtained and noted potential conflicts of interest for all VCEP members as required by ClinGen. The *CDH1* VCEP elected to divide rules into four types of evidence: (i) population data, (ii) computational/predictive data, (iii) functional data, and (iv) clinical evidence. Criteria in each evidence category have differing strength levels, with defined ACMG/AMP pathogenic criteria, including very strong (PVS), strong (PS), moderate (PM), and supporting (PP) levels, and stand-alone (BA), strong (BS), and supporting (BP) evidence levels for benign criteria. Work tasks were assigned and information on each evidence

type was presented on monthly call meetings to ensure that all members had the comprehensive background knowledge necessary for rule review and for informed decisions regarding the utility and potential to modify the strength of individual criteria.

Members of the VCEP agreed to nominate *CDH1* test variants and provide and share deidentified data relevant to their assessment. VCEP members were asked to nominate variants with sufficient amounts of underlying data to allow testing of a variety of the *CDH1* gene specifications. A series of 50 previously characterized *CDH1* variants were selected for rule pilot testing, with the goal of having approximately two-thirds expected to be known/suspected pathogenic or known/suspected benign, and the remaining being uncertain clinical significance or variants with conflicting assertions of interest in ClinVar (Supp. Table S1).

The pilot rule testing was performed by a group of six biocurators trained to perform variant assessment using both the standard 2015 ACMG/AMP guidelines and the *CDH1* VCEP specifications. For a given variant, the curators selected the criteria based on the evidence provided by VCEP members and data retrieved from the literature and population databases, and then combined them to choose a classification from the five-tier system as previously described (Richards et al., 2015). Classifications were made without modifying the rules for combining evidence codes (Richards et al., 2015). The curations were then reviewed by the VCEP to achieve group consensus regarding the variant classification and any subsequent further optimization of the *CDH1* specifications. VUS or variants with conflicting interpretations by multiple submitters in ClinVar were reassessed by the VCEP using the *CDH1* specifications of the ACMG/AMP rules to eliminate conflicting classifications whenever possible. The rules specifications were then presented to the ClinGen Sequence Variant Interpretation (SVI) working group, which provided preliminary review and feedback on the criteria. In April 2018, the *CDH1* VCEP held a group meeting at the ACMG annual meeting to evaluate and implement, when applicable, the SVI recommendations to the *CDH1* rules specifications and reach the final specifications. Variants are annotated using GenBank *CDH1* reference sequence NM_004360.3 and NP_004351 (GRCh37/hg19).

3 | RESULTS

3.1 | *CDH1* variant curation specification

The *CDH1* VCEP specifications to the ACMG/AMP variant curation criteria are summarized in Table 1. There were four criteria that the *CDH1* VCEP members felt needed no specifications, which were PS1 (same amino acid change as a previously established pathogenic variant), PM4 (protein length changes as a result of in-frame nucleotide deletions/insertions or stop-loss variants), BS4 (lack of segregation in affected members of a family), and BP5 (variant found in a case with an alternate molecular basis for disease). These codes should be applied as outlined in the original description of the ACMG/AMP variant curation guidelines. To the remaining criteria, *CDH1*-specific specifications and/or modifications to the strength of evidence were made based on data evaluated by the VCEP, as discussed below.

3.2 | Population data

The frequency of a variant in the general population is useful evidence to assess its potential pathogenicity. This can be accomplished by searching publicly available population databases. The *CDH1* VCEP has decided to use the Genome Aggregation Database (gnomAD) as its reference population database for allele frequency (gnomAD: https://gnomad.broadinstitute.org/). The database aggregates and harmonizes exome and genome sequencing data from a variety of large-scale sequencing projects. To date, the data set provided on this Website spans 123,136 exomes and 15,496 genomes from unrelated individuals sequenced as part of various disease-specific and population genetic studies. In its first release, which contained exclusively exome data, it was known as the Exome Aggregation Consortium (ExAC) (Lek et al., 2016).

3.2.1 | BA1: Allele frequency is greater than expected for disorder

BA1 is a stand-alone criterion. HDGC is a rare disorder with an estimated prevalence of <0.1 per 100,000 (Oliveira, Seruca, & Carneiro, 2009; Oliveira, Seruca, Hoogerbrugge, Ligtenberg, & Carneiro, 2013). The original ACMG/AMP criteria that recommended applying the criterion for alleles identified above 5% frequency appeared to be much higher than necessary. To obtain a *CDH1* specific population allele frequency threshold (AFT) we utilized the Whiffen/Ware calculator (Whiffin et al., 2017) (https://cardiodb.org/allelefrequencyapp/). The AFT cutoff for BA1 was obtained using a conservative unascertained penetrance estimate of 30% (Lowstuter et al., 2017) and the prevalence of LBC (~one in 800 individuals), which is the more commonly occurring cancer that can be attributed to pathogenic variants in *CDH1* (Li, Anderson, Daling, & Moe, 2003). Using the estimate of 30% for penetrance and the prevalence of LBC in the general population, the inferred AFT obtained for BA1 was 2.08333E−03. Therefore, we recommend application of BA1 to an allele of one of the seven major population groups in ExAC/gnomAD (European, Latin, South Asian, East Asian, African, European Finnish, and Ashkenazi) with an allele frequency equal to or higher than 0.002 (0.2%, 99.99% CI). However, since benign alleles may be more frequent in specific populations but rare overall, for this criterion to be applied there must be a minimum of five alleles in the sub-population to eliminate potential erroneous variant calls.

3.2.2 | BS1: Allele frequency is greater than expected for disorder

Using the same approach described above, the *CDH1* VCEP recommends the use of BS1 for a variant with an allele frequency equal to or higher than 0.001 (0.1%, 99.99% CI) in any of the seven major population groups in ExAC/gnomAD (European, Latin, South Asian, East Asian, African, European Finnish, and Ashkenazi). Similar to above, a sub-population must include five or more alleles to allow using this criterion. The Whiffen/Ware AFT cutoff for BS1 was obtained using the conservative penetrance estimate of 30% and the Japanese gastric cancer prevalence (~one in 1,250 individuals), which is one of the highest in the world (Guggenheim & Shah, 2013; Inoue & Tsugane, 2005).

TABLE 1 CDH1 rule specifications for the ACMG/AMP variant curation guidelines

ACMG/AMP criteria codes	Original ACMG/AMP rule summary	CDH1 rule specifications					Comments
		Stand alone	Very strong	Strong	Moderate	Supporting	
PVS1	Null variant in a gene where LoF is a known mechanism of disease	—	Per ClinGen SVI guidelines with the exception of canonical splice sites	Per ClinGen SVI guidelines Other CDH1 caveats: - Use the strong strength of evidence for canonical splice sites - CDH1 Exonic deletions or tandem duplications of in-frame exons. - Truncations in NMD-resistant zone located upstream the most 3′ well-characterized pathogenic variant c.2506G>T (p.Glu836Ter). Use PVS1_moderate if premature stop is downstream of this variant	Per ClinGen SVI guidelines Other CDH1 caveats: - G to non-G variants disrupting the last nucleotide of an exon - Canonical splice sites located in exons demonstrated experimentally to result in in-frame partial skipping/insertion (e.g., Exon 3 donor site)	Per ClinGen SVI guidelines	RNA analysis is recommended for splicing alterations, and if the RNA evidence does not support the prediction, the strength should be updated PP3 cannot be applied for canonical splice sites
PS1	Same amino acid change as a previously established pathogenic variant regardless of nucleotide change	—	—	Per original ACMG/AMP guidelines	—	—	Variant must not impact splicing
PS2	De novo (both maternity and paternity confirmed) in a patient with the disease and no family history	—	≥Two patients with DGC &/or LBC w/parental confirmation	One patient with DGC &/or LBC w/parental confirmation	—	—	Use ClinGen's de novo point system for a highly specific phenotype (see Table S2)
PS3	Well-established in vitro or in vivo functional studies supportive of a damaging effect on the gene or gene product	—	—	RNA assay demonstrating abnormal out-of-frame transcripts	—	RNA assay demonstrating abnormal in-frame transcripts	This rule can only be applied to demonstrate splicing defects.
PS4	Prevalence of variant in affected individuals is significantly increased compared to controls	—	Sixteen families meet HDGC criteria	Four families meet HDGC criteria	Two families meet HDGC criteria	One meets HDGC criteria	This rule assumes 30% penetrance in individuals with pathogenic variants. For example, if the variant is observed in three families, at least one of those families need to meet criteria for HDGC in order to apply this rule. PS4 cannot be applied to variants that meet BS1 or BA1

(Continues)

TABLE 1 (Continued)

ACMG/AMP criteria codes	Original ACMG/AMP rule summary	CDH1 rule specifications					Comments
		Stand alone	Very strong	Strong	Moderate	Supporting	
PM1	Located in a mutational hot spot and/or critical and well-established functional domain without benign variation	—	—	—	—	—	Do not use for this gene
PM2	Absent in population databases	—	—	—	<One out of 100,000 alleles in gnomAD cohort; if present in ≥ 2 individuals, must be present in <one out of 50,000 alleles within a sub-population	—	Use gnomAD to determine allele frequency. Beware of technical limitations that can inaccurately represent allele frequency in this population database
PM3	For recessive disorders, detected in trans with a pathogenic variant	—	—	—	—	—	Does not apply to this gene
PM4	Protein length changes as a result of in-frame deletions/insertions in a non-repeat region or stop-loss variants	—	—	—	Per original ACMG/AMP guidelines	—	No rule specification proposed. Variant example - CDH1 c.2647T>C (p.Ter883GlnEXT*29)
PM5	Novel missense change at amino acid residue where a different missense variant is pathogenic	—	—	—	—	—	Do not use rule at this time
PM6	Assumed de novo, but w/o confirmation of paternity and maternity	—	≥Four patients with DGC &/or LBC w/o parental confirmation	≥Two patients with DGC &/or LBC w/o parental confirmation	One patient with DGC &/or LBC w/o parental confirmation	—	Use ClinGen's de novo point system for a highly specific phenotype (See Supp. Table S2)
PP1	Cosegregation in multiple affected family members in a gene definitively known to cause the disease	—	—	≥Seven meioses across ≥ 2 families	Five to six meioses across ≥ 1 families	Three to four meioses across ≥ 1 families	Based strength of rule code on number of meioses across one or more families
PP2	Missense variant in a gene with a low rate of benign missense variation & where missense variants are a common mechanism of disease	—	—	—	—	—	Do not use rule at this time

(Continues)

TABLE 1 (Continued)

ACMG/AMP criteria codes	Original ACMG/AMP rule summary	CDH1 rule specifications					Comments
		Stand alone	Very strong	Strong	Moderate	Supporting	
PP3	Multiple lines of computational evidence support a deleterious effect on the gene or gene product	—	—	—	Variants affecting the same splice site as a well-characterized variant with similar or worse in silico/ RNA predictions	At least three in silico splicing predictors in agreement (HSF, Maximum Entropy (MaxEnt), Berkeley Drosophila Genome Project (BDGP), or ESEfinder)	Rule code is only for non-canonical splicing variants. Code also does not apply to last nucleotide of exon 3 (c.387G). Do not use protein-based computational prediction models for missense variants
PP4	Patient's phenotype or family history is highly specific for a disease with a single genetic etiology	—	—	—	—	—	Use PS4 in place of PP4
PP5	Reputable source recently reports variant as pathogenic	—	—	—	—	—	Do not use rule at this time
BA1	Allele frequency is greater than expected for disorder	MAF cutoff of 0.2%	—	—	—	—	99.99% CI; subpopulation must have a minimum of five alleles present
BS1	Allele frequency is greater than expected for disorder	MAF cutoff of 0.1%	—	—	—	—	99.99% CI; subpopulation must have a minimum of five alleles present
BS2	Observed in a healthy adult individual for a dominant disorder with full penetrance expected at an early age	—	—	Variant seen in ≥10 individuals w/o DGC, SRC tumors, or LBC & whose families do not suggest HDGC	—	Variant seen in ≥3 individuals w/o DGC, SRC tumors, or LBC & whose families do not suggest HDGC	
BS3	Well-established in vitro or in vivo functional studies show no damaging effect on protein function or splicing	—	—	Functional RNA studies demonstrating no impact on transcript composition	—	—	This rule can only be used to demonstrate lack of splicing and can be downgraded based on quality of data
BS4	Lack of segregation in affected members of a family	—	—	Per original ACMG/AMP guidelines	—	—	Beware of the presence of phenocopies (e.g., breast cancer) that can mimic lack of segregation. Also, families may have more than one pathogenic variant contributing to another AD disorder

(Continues)

WILEY Human Mutation | 1559

TABLE 1 (Continued)

ACMG/AMP criteria codes	Original ACMG/AMP rule summary	CDH1 rule specifications					Comments
		Stand alone	Very strong	Strong	Moderate	Supporting	
BP1	Missense variant in a gene for which primarily truncating variants are known to cause disease	—	—	—	—	—	Does not apply to this gene
BP2	Observed in a healthy homozygous individual, or in *trans* with a pathogenic variant for a fully penetrant dominant gene/disorder or observed in *cis* with a pathogenic variant	—	—	Variant observed *in trans* w/known pathogenic variant (phase confirmed) **OR** observed in the homozygous state in individual w/o personal &/or family history of DGC, LBC, or SRC tumors	—	Variant is observed in *cis* (or phase is unknown) w/ a pathogenic variant	Evidence code is dependent on strength of data. Take into consideration the quality of sequencing data when applying code. Note that code requires knowledge of individuals' phenotype. Therefore, data from population databases should only be used when phenotypic info is available
BP3	In-frame deletions/insertions in a repetitive region without a known function	—	—	—	—	—	Do not use rule at this time
BP4	Multiple lines of computational evidence suggest no impact on gene/gene product	—	—	—	—	Splicing predictions only. At least three *in silico* splicing predictors in agreement (HSF, Maximum Entropy (MaxEnt), Berkeley Drosophilia Genome Project (BDGP), or ESEfinder)	This rule can only be used when splicing prediction models suggest no impact on protein. Do not use protein based computational prediction models for missense variants
BP5	Variant found in a case with an alternate molecular basis for disease	—	—	—	—	Per original ACMG/AMP guidelines	This applies if a P/LP variant is identified in an alternate gene known to cause HDGC (e.g., CTNNA1)
BP6	Reputable source recently reports variant as benign	—	—	—	—	—	Do not use rule at this time
BP7	Synonymous variant which splicing prediction algorithms predict no impact to the splice consensus sequence nor the creation of a new splice site & the nucleotide is not highly conserved.	—	—	—	—	Synonymous variants where nucleotide is not highly conserved; variant is the reference nucleotide in one primate and/or >3 mammal species	Note the *CDH1* rule specification does not require a benign *in silico* splice prediction. This allows use with BP4, as appropriate, to classify variants meeting both criteria as likely benign

3.2.3 | PM2: Absent from population databases

The original PM2 criterion was developed by ACMG/AMP prior to the availability of ever larger databases of exome and genome sequence data. In addition, these databases are not designed to exclude every individual with disease (e.g., the ExAC online server includes individuals from the TCGA cancer project (Lek et al., 2016)) or every individual that is a young adult and might develop disease later in life. Thus, pathogenic alleles can be found in these databases, and in order to use the PM2 criterion for *CDH1*, the VCEP specifies that the variant must be present in fewer than one in 100,000 alleles in the gnomAD cohort. If the variant is present in two or more alleles in the gnomAD cohort, it must be present in <1 in 50,000 alleles within a sub-population. The variant cannot be observed in the homozygous state in population databases. The cutoff for use of this pathogenic evidence code was validated using rare *CDH1* LoF pathogenic alleles identified in gnomAD. Currently, the most common known pathogenic variant identified in gnomAD is *CDH1* c.1003C>T (p.Arg335Ter) with an overall allele frequency of three out of 277,002 (0.001%), a South Asian population allele frequency of one out of 30,780 (0.003%), and a non-Finnish European population allele frequency of two out of 126,548 (0.002%).

3.3 | Computational and predictive data

3.3.1 | PVS1: Predicted null variant in a gene where LoF is a known mechanism of disease

LoF is the mechanism of disease in *CDH1* (Oliveira, Pinheiro, Figueiredo, Seruca, & Carneiro, 2015). *CDH1* LoF variants are pathogenic, in part, because the majority of these germline alterations result in premature stop codons (van der Post et al., 2015), most of which are predicted to elicit nonsense-mediated decay (NMD), an mRNA quality control pathway that destabilizes abnormal transcripts containing premature stop codons (Holbrook, Neu-Yilik, Hentze, & Kulozik, 2004). Indeed, *CDH1* variants that result in transcripts with premature stop codons have been shown to be targeted by the NMD pathway and to result in haploinsufficiency in gastric tissue (Karam et al., 2008). Therefore, the pathogenic very strong criterion, PVS1, should be applied following SVI recommendations (Abou Tayoun, Harrison, *submitted to this issue*) to variants that are predicted to result in LoF and elicit NMD, including nonsense, frameshift, initiation codon, out-of-frame single or multi exon deletions, out-of-frame tandem single or multi exon duplications, and whole gene deletions. However there are scenarios specific to *CDH1* that required additional rule specifications, as described below.

Variants in *CDH1* that may not trigger NMD include alterations that result in premature stop codons located downstream of the *CDH1* NMD boundary, which is approximately 55 nucleotides upstream of the last exon-exon junction in the penultimate exon of the gene, at amino acid 795 of the E-cadherin protein (Karam et al., 2008; Krempely & Karam 2018). Transcripts that possess premature termination codons past this boundary are predicted to escape detection by the NMD machinery (Holbrook et al., 2004; Karam et al., 2008; Rivas et al., 2015), potentially encoding for C-terminal truncated E-cadherin proteins retaining partial function (Sasaki, Lin, Morin, & Longo, 2000). The *CDH1* VCEP recommends adjusting the strength of the PVS1 evidence code for variants resulting in a premature stop codon downstream of the NMD boundary at amino acid 795 (see Table 1). In these cases, we recommend adjusting the strength of the evidence from "very strong" to "strong" (PVS1_strong), if the predicted truncation is located downstream of the NMD boundary, but upstream of the most 3′ well-characterized pathogenic variant, *CDH1* c.2506G>T (p.Glu836Ter) (Krempely & Karam, 2018). This variant is predicted to result in truncation of E-cadherin's C-terminal region and to disrupt the cytoplasmic domain. E-cadherin's cytoplasmic domain includes the catenin-binding domain, which stabilizes cell adhesion (Gul, Hulpiau, Saeys, & van Roy, 2017); therefore, truncations that disrupt this region may result in increased cell invasion and impairment of cellular adhesion. However, if the premature stop codon is located downstream of *CDH1* c.2506G>T (p.Glu836Ter), in codons 837–882, use PVS1 as moderate evidence (PVS1_moderate), due to the uncertainty surrounding the clinical relevance of the amino acids located downstream of amino acid 836 (Krempely & Karam, 2018).

Other types of alterations that do not elicit NMD and may not result in LoF include in-frame exon duplications proven to be in tandem (Richardson et al., 2018) and deletions of in-frame exons: exon 4 (144 nt.), exon 5 (156 nt.), exon 8 (129 nt.), exon 9 (183 nt.), exon 12 (225 nt.), exon 13 (228 nt.), and exon 15 (144 nt.). For in-frame exon deletions or for tandem duplications predicted to result in in-frame exon gains, adjust the strength of the evidence from "very strong" to "strong" (PVS1_strong) as these deletions and duplications would effect key E-cadherin functional domains.

For splicing variants, we recommend using the PVS1_strong rule code for all canonical ±1/2 splice site variants. This recommendation is based on the fact that there are *CDH1* canonical splicing variants of uncertain deleteriousness (Yelskaya, Somar et al., 2016). By applying PVS1_strong to canonical splice sites, additional evidence, such as RNA data (PS3) and clinical history (PS4), are necessary to reach a pathogenic classification. As a result, the classification of splicing variants will not rely too heavily on the LoF attribute, encouraging the use of important complimentary data, such as RNA analysis, to confirm LoF. For canonical ±1/2 splice site variants located in exons with in-frame cryptic acceptor and donor splice sites demonstrated to result in small in-frame amino acid deletions/insertions (less than 10% of the protein) use PVS1 as moderate evidence (PVS1_moderate). For example, the PVS1 evidence level should be downgraded to moderate for any canonical splicing alteration located at exon's 3 native donor splice site, since the variant *CDH1* c.387+1G>A has been demonstrated experimentally to result in activation of cryptic in-frame donor splice sites (Yelskaya, Somar et al., 2016). Finally, there are well-characterized pathogenic variants located in the last nucleotide of many *CDH1* exons that are known to disrupt splicing (Frebourg et al., 2006; Guilford et al., 1998; Karam et al., 2008). With this in mind, the EP is recommending to use PVS1_moderate for "G" to "non-G" nucleotide changes in the last nucleotide of an exon. Note, quantitative and qualitative RNA analysis of variants predicted to affect splicing is recommended in order to confirm whether the predictions are accurate (Farber-Katz et al., 2018).

3.3.2 | PP3: Multiple lines of computational evidence support a deleterious effect on the gene or gene product

This code can only be used for variants predicted by multiple lines of computational evidence to have a deleterious impact on splicing. At least three *in silico* splicing predictors, such as Human Splicing Finder (HSF), Maximum Entropy Scan (MaxEntScan), Berkeley Drosophila Genome Project (BDGP), and ESEfinder, must be in agreement to apply the supporting rule for variants likely to impact splicing. These include coding and non-coding variants that are predicted to either have an impact on the native site, or result in activation/creation of cryptic/novel splice sites. These four tools were used to perform splicing *in silico* analysis for three *CDH1* alterations reported in the literature to result in abnormal splicing (c.715G>A, c.1901C>T, c.2195G>A) (Kaurah et al., 2007; Pinheiro, Oliveira, Seruca, & Carneiro, 2014; van der Post et al., 2015). These three variants were predicted by the four *in silico* tools above to result in the creation of novel donor sites, which was confirmed by the RNA data described in the literature. We also performed analysis of the *CDH1* VCEP membership laboratories' cohort phenotypes associated with each of the variants. A summary of the clinical and population data for these variants is shown in Figure 1a, b, and in Supp. Table S1. To avoid double counting of evidence, this PP3 rule cannot be used for canonical splice site variants and should only be applied to splicing variants at locations other than the ±1/2 positions. Note, this code also does not apply to the last nucleotide of exon 3 (c.387G) (Yelskaya et al., 2016).

A PP3_moderate rule code can be used when the variant in question affects the same splice site as a well-characterized variant with similar or worse splicing *in silico* predictions. For example, c.1137G>T (p.Thr379=) impacts the last nucleotide of exon 8 and has similar *in silico* splicing predictions as the well-characterized pathogenic variant c.1137G>A (p.Thr379=); therefore PP3_moderate was used towards the classification of c.1137G>T (p.Thr379=) (Supp. Table S1).

The VCEP decided to not use PP3 for *CDH1* missense variants, after substantial testing of a variety of protein-based *in silico* algorithms (REVEL, VEST3, MetaSVM, CADD, SIFT, Polyphen2, and Provean). We found that computational predictions for deleteriousness were inconclusive with regards to accurately predicting the pathogenicity of the variant as determined by other lines of evidence, such as clinical data and population frequency (Figure 1a and b). Therefore the *CDH1* VCEP recommends to not use this evidence code for missense changes.

3.3.3 | BP4: Multiple lines of computational evidence suggest no impact on splicing

We recommend that this code only be used for synonymous or intronic splicing variants when at least three *in silico* tools (HSF, MaxEntScan, BDGP, and ESEfinder) agree that the variant is not predicted to affect splicing (defined as the scores from each program, on average, change less than 10% from the reference sequence). For missense variants, clinical validity of protein based computational predictions as a benign line of evidence was inconclusive (see above PP3), and therefore the panel recommends to not use this evidence code for missense changes.

3.3.4 | BP7: A synonymous (silent) variant and the nucleotide is not highly conserved

This evidence code can be applied to a synonymous variant when the nucleotide is not highly conserved. The *CDH1* VCEP recommends to use BP7 if the variant is the reference nucleotide in 1 primate and/or >3 mammal species.

3.4 | Functional data

3.4.1 | PS3: Well-established *in vitro* or *in vivo* functional studies supportive of a damaging effect on the gene or gene product

Abnormal splicing is a well-known mechanism of disease in *CDH1*. There are several well-characterized *CDH1* pathogenic variants demonstrated to result in abnormal splicing (Frebourg et al., 2006; Karam et al., 2008; Kaurah et al., 2007; Pinheiro, Oliveira, Seruca, & Carneiro, 2014; van der Post et al., 2015). We reviewed the available literature for six exonic *CDH1* variants that were demonstrated experimentally to result in abnormal out-of-frame transcripts (c.715G>A, c.1008G>A, c.1137G>A, c.1679C>G, c.1901C>T, c.2195G>A) (Frebourg et al., 2006; Karam et al., 2008; Kaurah et al., 2007; Pinheiro, Oliveira, Seruca, & Carneiro, 2014; van der Post et al., 2015; Yelskaya, Bacares et al., 2016). Based on this review, and on the curation of other splicing variants performed by the VCEP, we recommend using PS3 for variants demonstrated to result in abnormal out-of-frame transcripts. Note that this recommendation applies to any variants demonstrated to cause splicing defects (e.g., exonic missense/synonymous, canonical splice sites, other intronic variants). However, like many genes, *CDH1* pre-mRNA can undergo complex splicing, with several mechanisms that could impact the deleteriousness of a variant (e.g., in-frame skipping, usage of cryptic in-frame splice sites, etc.). We recommend using PS3_supporting for variants demonstrated to result in in-frame transcripts. In the future it will be important to perform a systematic analysis of splicing alterations in *CDH1* in order to address many unanswered questions, such as type of tissue that should be evaluated, type of assays that would be ideally performed, and abnormal splicing thresholds. With this in mind, the VCEP will continue to evaluate this specific subject with the intention of making further specifications for the use of PS3.

The *CDH1* VCEP also evaluated the use of PS3 for missense protein functional evidence. E-cadherin is a transmembrane glycoprotein expressed on epithelial tissue and is responsible for calcium-dependent cell-to-cell adhesion and establishing and maintaining polarized and differentiated epithelia through intercellular adhesion complexes (Berx et al., 1995; Takeichi, 1991). E-cadherin deregulation is correlated with the infiltrative and metastatic ability of tumors, with the consequent loss of cell adhesion and concomitant increase in cell motility and invasion (Bruner, et al., 2018). Missense variants in E-cadherin have been described across the protein, and account for ~20% of all variants reported in the literature so far (Corso, Marrelli, & Roviello, 2013; Pinheiro et al., 2014; van der Post, Vogelaar, Carneiro et al., 2015; van der Post, Vogelaar, Manders et al., 2015). We reviewed the literature for 14 missense variants (not
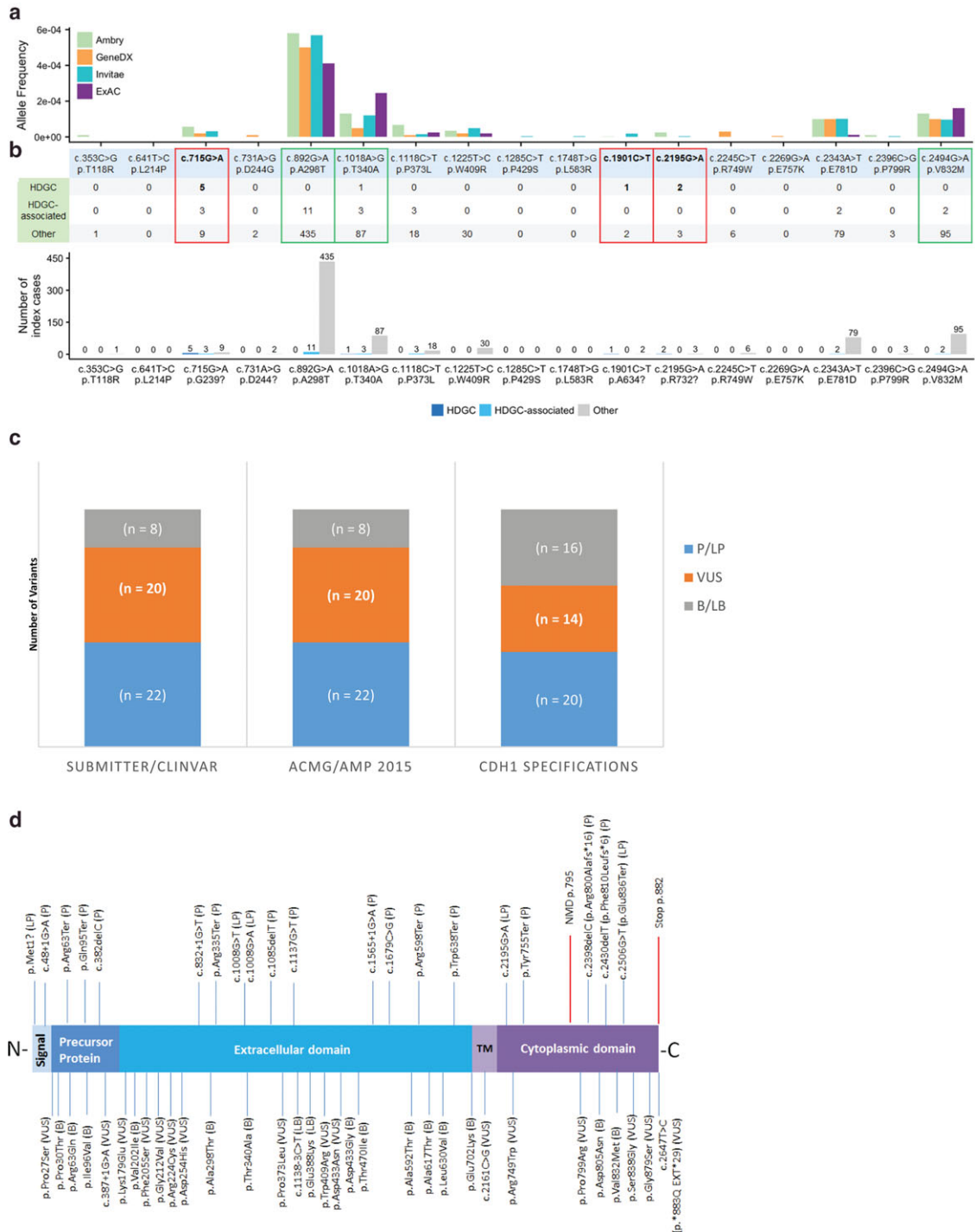
FIGURE 1 Pilot assessment of *CDH1* variants. (a) Clinical and population evidence of selected *CDH1* missense variants. Allele frequencies for the selected variants in the ExAC database (~120,000 alleles) and in the cohorts of the three largest data sharing clinical genetic laboratories in the U.S.A. (~827,000 alleles total). (b) Laboratory cohort phenotypes associated with each of the specific variants. HDGC, family meets International Gastric Cancer Linkage Consortium (IGCLC) criteria; HDGC-associated tumors, Diffuse Gastric Cancer, Signet Ring Cell tumors (e.g., Krukenberg tumor), Lobular Breast Cancer, but does not meet IGCLC criteria; others are either reported unaffected or having non-*CDH1* tumors (e.g., breast invasive ductal carcinoma). CDH1 c.715G>A, c.1901C>T, and c.2195G>A were demonstrated experimentally by RNA analysis to activate cryptic splicing sites; CDH1 p.A298T, p.T340A, and p.V832M met benign population allele-frequency cutoffs. (c) Comparing variant classifications from VCEP or ClinVar submitters and from the original ACMG/AMP criteria versus applying *CDH1* evidence code specifications demonstrates a reduction of VUS. P, pathogenic; LP, likely pathogenic; VUS, variant of uncertain significance; LB, likely benign; B, benign. ClinVar assertions were accessed on April 16, 2018. (d) E-cadherin schematic representation demonstrating its major domains in relation to the 50 pilot variants curated by the VCEP. E-cadherin is a highly conserved transmembrane protein that possess a signal peptide, a precursor protein, the extracellular domain, which is the "adhesive" ectodomain composed of tandem cadherin repeats, the transmembrane (TM) domain, and the cytoplasmic domain, which interacts with the armadillo catenins, p120 and β-catenin (Gul et al., 2017). Clinically actionable variants (P, LP) are shown at the top; VUS, LB, and B variants are shown at the bottom

predicted to impact splicing). All 14 of these variants are reported as having computational and two or more *in vitro* and/or *in vivo* functional lines of evidence to support a pathogenic classification (Figueiredo et al., 2013; Fitzgerald & Caldas, 2004; Pinheiro et al., 2014; Simoes-Correia et al., 2012; Simoes-Correia et al., 2008; Suriano, Mulholland et al., 2003; Suriano, Oliveira, Ferreira et al., 2003; Suriano, Oliveira, Huntsman et al., 2003; Suriano, Seixas, Rocha, & Seruca, 2006). To estimate clinical validity of these protein functional assays, we then compared the available functional evidence for these variants with the de-identified clinical data obtained from members of the *CDH1* VCEP, encompassing an aggregated cohort of clinical laboratory data of ~827,000 *CDH1* alleles and ExAC's ~120,000 alleles (Figure 1a and b). In this large dataset, only one individual with a missense alteration (p.Thr340Ala) met the 2015 diagnostic testing criteria for HDGC (Figure 1b). However, this individual was heterozygous for a co-occurring *CDH1* truncating variant (p.Gln771Ter), and clinical history was attributed to this pathogenic alteration. Overall, no individuals with a missense variant met diagnostic criteria for HDGC (Figure 1b), despite reported functional evidence of abrogated protein function. Based on the clinical and population evidence, the majority of the *CDH1* VCEP agreed that none of the currently published *in vitro* and *in vivo* functional studies could be confidently used to predict pathogenicity of E-cadherin missense variants, and therefore functional assays for missense alterations should not be used for *CDH1* variant classification. However, this topic will be under continued examination, as novel research in the future may generate assays that are able to better predict the clinical behavior of E-cadherin missense variants.

3.4.2 | BS3: Well-established *in vitro* or *in vivo* functional studies show no damaging effect on protein function or splicing

Similar to above, this evidence code should only be used for splicing. It applies to intronic and synonymous variants that are demonstrated to not result in abnormal transcripts. In gastric tissue, an abnormal out-of-frame transcript may not be detectable due to NMD, and therefore allele-specific expression is recommended in these samples (Karam et al., 2008). Finally, the *CDH1* VCEP did not further evaluate the use of BS3 as a benign line of evidence for missense alterations at this time.

3.5 | Clinical data

Two decades ago, linkage analysis identified germline *CDH1* LoF variants as the genetic cause of HDGC (Guilford et al., 1998). A year later, the association of LBC and germline *CDH1* variants was first reported (Keller et al., 1999). Since these original findings, HDGC has been established as a cancer predisposing syndrome with increasingly reliable estimates of cancer risk for carriers of LoF P/LP variants (Lowstuter et al., 2017; Oliveira et al., 2015; van der Post et al., 2015). Recent penetrance estimates were derived from 75 families that were ascertained by having met clinical criteria for HDGC. The number of pathogenic variants seen in 3,858 individuals (353 with DGC and 89 with LBC) demonstrated that the cumulative incidence of DGC by 80 years was 70% (95% CI, 59%–80%) for male participants and 56%

(95% CI, 44%–69%) for female participants. Additionally, the risk of LBC for female participants was 42% (95% CI, 23%–68%) by 80 years (Hansford et al., 2015). Based on the *CDH1* highly specific phenotype, the VCEP recommendation is to use the updated 2015 criteria (see below) to evaluate whether patients meet HDGC clinical criteria in a semi-quantitative fashion—the more patients/families that meet criteria, the stronger the evidence for ACMG/AMP codes that rely on phenotype (Table 1).

3.5.1 | PS4: Prevalence of the variant in affected individuals is significantly increased compared with the prevalence in controls

The 2015 ACMG/AMP guidelines allow the observation of a variant in multiple unrelated individuals with the same phenotype, and its absence in controls, to be used as evidence of pathogenicity when case–control studies are unavailable (Richards et al., 2015). Most P/LP *CDH1* variants are rare, and since there are limited case-control data available for most variants, the VCEP recommends counting individual cases meeting diagnostic criteria for PS4. Following SVI recommendations and to create consistency across Expert Panels, the PS4 rule code is being used in conjunction with the factors associated with PP4 (patient's phenotype or family history is highly specific for a disease with a single genetic etiology). Since phenotype and family history are now integrated with the PS4 rule code, we eliminated the use of the PP4 to avoid double counting of evidence and applied the stronger rule code to take into account all appropriate probands. This code has a sliding weight scale to account for the number of individuals who meet criteria.

In order to apply evidence code PS4, all probands being counted, including both published and unpublished cases harboring the variant, should have a personal and/or family history (first- and second-degree relatives) meeting one of the recommended *CDH1* testing criteria below (van der Post et al., 2015):

- Two gastric cancer cases in a family, regardless of age, at least one confirmed DGC;
- One case of DGC diagnosed <40 years;
- Personal or family history of DGC and LBC; one diagnosed <50 years;
- Bilateral LBC or family history of two or more cases of LBC <50 years;
- A personal or family history of cleft lip/palate in a patient with DGC;
- *In situ* signet ring cell (SRC) and/or pagetoid spread of SRC detected in pathology specimens.

Following SVI recommendations, the strength of the evidence code increases exponentially based on the number of families meeting HDGC criteria. For instance, having one family meeting criteria would correlate to PS4_supporting, which is the same weight as using the original PP4 code. Having two to three families correlates to PS4_moderate, 4–15 correlates to PS4, and 16 or more families would allow for use of PS4_very strong.

However, more recently the penetrance of phenotypes associated with pathogenic variants in *CDH1* in unascertained individuals was estimated to be 30% (Lowstuter et al., 2017). Due to the estimated 30% penetrance, we can expect that approximately 70% of the families identified with a specific pathogenic variant will not meet the strict testing criteria described above. Since pathogenic variants are expected to be identified in unaffected individuals, pathogenicity is not excluded if detected in families not meeting criteria. Therefore, if the variant is observed in three separate families, at least one of those families must meet criteria for HDGC in order to apply this rule as supporting (PS4_supporting). As follows, one would use this rule code as moderate evidence if this variant was identified in six separate families and at least two met criteria for HDGC, and so forth. As we better understand the penetrance of *CDH1* pathogenic variants, the panel will continue to evaluate this approach, and will consider developing a partial weighting scheme for cases that do not meet full criteria.

3.5.2 | PS2: *De novo* (both maternity and paternity confirmed) in a patient with the disease and no family history

CDH1 de novo variants are rare, with only a few reports in the literature (Krempely & Karam, 2018; Shah et al., 2012; Sugimoto et al., 2014). To apply this evidence code, the proband with the variant must have DGC or LBC. Both parents must have tested negative for the variant, and paternity and maternity must be confirmed. Following the ClinGen SVI's *de novo* point system for highly specific phenotype, the strength of the rule increases based on the number of confirmed *de novo* cases (Supp. Table S2). For example, if the variant is observed once and confirmed as a *de novo* occurrence in a patient with DGC, apply strong pathogenic evidence; if seen and confirmed twice in patients with DGC, use very strong evidence (PS2_very strong) for pathogenicity. Refer to Supp. Table S2 to determine the strength of this evidence code when assessing *de novo CDH1* variants.

3.5.3 | PM6: Assumed *de novo*, but without confirmation of paternity and maternity

In many cases, the laboratory may not have samples or appropriate genotypes to confirm parental relationships. For example, only a targeted region of *CDH1* might be sequenced in parental samples. Similar to the other related codes, the proband with the variant must have DGC or LBC to apply this evidence code. Family history must be consistent with *de novo* occurrence and both parents must have tested negative for the variant. Similar to PS2, the strength of the rule increases based on the number of assumed *de novo* cases. For example, if the variant is observed once as an assumed *de novo* occurrence (parental status not confirmed), apply moderate pathogenic evidence (PM6); if seen as an assumed *de novo* twice, use strong evidence for pathogenicity (PM6_strong); if seen four times as an assumed *de novo*, use very strong evidence for pathogenicity (PM6_very strong). As above, refer to Supp. Table S2 to determine the strength of this evidence code when assessing *de novo* variants.

3.5.4 | BS2: Observed in a healthy adult individual with full penetrance expected at an early age

The corresponding evidence code for individuals with a variant but without disease was also considered. Although originally written for diseases with full penetrance at an early age, we further specified the code to account for the reduced penetrance and later onset of *CDH1*. The strength depends on the number of adult individuals (older than 18 years old) in a clinical cohort reported as either unaffected or having non-*CDH1* tumors (e.g., breast invasive ductal carcinoma) where the pathology is specified. In order to apply strong evidence, the variant must be seen in at least 10 adult individuals with a documented diagnosis that does not include a known personal or family history of DGC, LBC, or tumors with SRC histology. Individuals should not be included in this count if the pathology of their cancer is unknown or unspecified. The fact that this condition is not fully penetrant was taken into consideration when specifying this evidence code. We used the allele frequency of known pathogenic variants to establish this threshold. For example, a variant identified 10 times in a laboratory cohort of >100,000 alleles (0.01%) would still be 5–10 times more common than the most frequently observed known pathogenic variant in gnomAD, c.1003C>T (p.Arg335Ter), and unlikely to be an undiscovered founder pathogenic variant. BS2_supporting should be applied if there are 3 to 9 adult individuals reported as unaffected, or having no personal or family history of DGC, LBC, or tumors with SRC histology. This analysis is based on testing cohorts of predominantly adult patients and would not necessarily apply to other younger large cohorts (e.g., exome or genomes of children with developmental disorders).

3.5.5 | BP2: Observed in *trans* with a pathogenic variant for a fully penetrant dominant gene/disorder or observed in *cis* with a pathogenic variant in any inheritance pattern

This rule is applicable if the variant is observed in the homozygous state in an individual without *CDH*1-associated phenotype, or observed in *trans* or in *cis* with a pathogenic variant in a patient with a *CDH1*-associated phenotype. The evidence strength depends on the data (Table 1). Since biallelic pathogenic variants in *CDH1* are purported to be embryonically lethal (Riethmacher et al., 1995), BP2_strong applies if observed in *trans* with a known pathogenic alteration and phase is confirmed. BP2_strong also applies if observed in a homozygous state in an individual with no personal and/or family history of *CDH1* cancers. BP2_supporting is appropriate if observed in *cis*, or phase unknown, with a known pathogenic alteration that has been reported separately in an individual with a personal and/or family history of *CDH1* cancers.

3.6 | ACMG evidence codes not used for *CDH1* variant classification

Eight of the twenty-eight published 2015 ACMG/AMP criteria were not relevant to *CDH1* and were therefore not used for variant classification. The reasons for each evidence code are briefly described here. As a gene associated with an autosomal dominant condition, PM3 (detected in *trans* with a pathogenic variant in a recessive gene) does not apply. With virtually no missense variants classified as P or LP,

we could not validate PM5 (missense change at the same codon as another pathogenic missense variant) or PP2 (missense variant in a gene with low rate of benign missense variation), and therefore are recommending to not apply PM5 and PP2 to *CDH1* missense variants. With regards to PM1 (located in mutational hot spot or domain without benign variation), the EP has decided to not use this code because there are benign variants in every domain of the protein (https://gnomad.broadinstitute.org/gene/ENSG00000039068). Following recommendations from the ClinGen SVI committee, the VCEP decided to not use PP4 (patient's phenotype or family history is highly specific for a disease with a single genetic etiology). As previously mentioned, instead of PP4, the PS4 rule code (significantly increased presence in cases versus controls) should be used, since it allows one to account for multiple individuals with the same variant and phenotype. The VCEP also decided not to use BP1 (missense variant in gene where primarily LoF causes disease) based on the possibility that there are missense P/LP variants that have not yet been identified. Similarly, *CDH1* does not contain a repetitive region without known function used to apply BP3. In the case of PP5 (reputable source reports as pathogenic) and BP6 (reputable source reports as benign), these criteria were removed based on recommendation from the ClinGen SVI working group (Biesecker & Harrison, 2018).

3.7 | Pilot evidence code testing

For pilot testing, 22 suspected pathogenic variants, seven suspected benign variants, and 12 VUS were recommended for inclusion by *CDH1* VCEP members. Most variants were previously classified by a participating clinical laboratory, but a few variants were recommended from a participating research group with their variant interpretation. We also selected nine variants from ClinVar, with a bias toward VUS (eight VUS/conflicted variants, one LB). This gave us a total of 50 variants for pilot testing. The VCEP biocurators applied the original ACMG/AMP guidelines and the *CDH1* specified criteria to each variant. Our pilot testing allowed for an iterative process by which experts could comment on the weight of certain lines of evidence and biocurators could provide feedback on the usability of the evidence codes. A list of all pilot variants, the ClinVar and the VCEP submitters' variant classifications, the classifications made by our biocurators using the original ACMG/AMP guidelines, and the *CDH1* VCEP variant assessment utilizing the *CDH1* rule specifications with evidence codes are listed in Supp. Table S1.

Twenty of 22 variants previously listed as P/LP in ClinVar and/or by the VCEP submitters remained at that assertion after applying the *CDH1* specifications. Of note, two of these 20 variants (c.1679C>G and c.2195G>A), which were demonstrated to result in abnormal out-of-frame splicing (Kaurah et al., 2007; Yelskaya et al., 2016), had conflicting variant assertions in ClinVar, as they were each listed as VUS by one ClinVar submitter. Following the *CDH1* specifications, the VCEP classified these variants as P and LP respectively based on the *CDH1*-specified codes that were applied. Two of 22 variants submitted by a VCEP member as P/LP were downgraded to VUS after applying the *CDH1* specifications. These variants were p.Lys179Glu and p.Leu721Val, and they were classified by the VCEP as VUS due to

lack of sufficient benign or pathogenic evidence. Six of the seven variants that were submitted by VCEP members as B/LB remained in the B/LB classification after applying the *CDH1* specifications; the remaining variant submitted as LB was classified as VUS by the VCEP due to lack of appropriate evidence codes (p.Arg224Cys). Twelve variants were previously interpreted as VUS by our VCEP members; four of these 12 were classified as benign, and the remaining eight remained as VUS. In addition, out of the nine variants selected from ClinVar (eight VUS/conflicted variants, one LB), six were classified as benign, while three remained as VUS. Overall, applying the *CDH1* specifications to the VUS submitted by VCEP members and to the VUS we selected from ClinVar (20 total) resulted in a reduction in VUS classifications from 40% to 28% (Figure 1c). In addition, 17 variants with assertions in ClinVar were VUS with conflicting classifications. After applying the *CDH1* specifications, 14 were reclassified to P/LP or B, while only three remained as VUS. Adjusting the *CDH1*-specific allele frequency cutoffs (BA1 and BS1), along with sharing clinical laboratory data (Supp. Table S3), made the most impact on moving variants from VUS/conflicting to benign.

When the VCEP biocurators compared classifications between using the ACMG/AMP guidelines and the *CDH1* specifications, we found similar results. Out of 20 variants that were classified as VUS with the ACMG/AMP guidelines, nine were classified as benign using the *CDH1* specifications, one was classified as pathogenic after applying *CDH1* specifications (c.2398delC), and the remaining 10 stayed as VUS. Overall, applying the *CDH1* specifications to 20 VUS classified by both the VCEP submitter and confirmed with VCEP biocurator application of the ACMG/AMP guidelines resulted in a reduction in VUS classifications from 40% to 28% (Figure 1c). Also, three variants were classified as likely pathogenic using ACMG/AMP guidelines (mostly due to the use of PS3 for missense functional assays and PP3 for missense *in silico* predictions), that ended up being classified as VUS using the *CDH1* specifications (p.Phe205Ser, p.Asp254His, p.Pro373Leu). However, in each of those instances, the VCEP submitter and/or the ClinVar submitter(s) classified the variants as VUS (Supp. Table S1). A summary of the classifications of the pilot variants obtained utilizing the *CDH1* rule specifications is provided in Figure 1d.

4 | DISCUSSION

Sequencing technology has evolved rapidly with the advent of NGS, resulting in an ever increasing number of new variants identified in individuals from multiple ethnic backgrounds. Diagnostic laboratories are collectively accumulating evidence at an astounding rate and through collaborative efforts, such as sharing data in ClinVar, interlaboratory discrepancy resolution efforts (Garber et al., 2016; Harrison et al., 2018; Abou Tayoun, Harrison, submitted to this issue), and ClinGen expert panels (Kelly et al., 2018; Gelb et al., 2018), are substantially increasing our knowledge of the tolerance for variability for each gene. Since its implementation, the *CDH1* VCEP has been evaluating a series of specifications to the 2015 ACMG/AMP variant interpretation guidelines. This expert panel plans to continue curating *CDH1* variants for deposition into the ClinVar database for public use, while

maintaining a critical eye toward the rule specifications. Variant assessment is an ever-evolving field, and as improved computational modeling, functional assays, diagnostic tools, and/or population databases are developed, the *CDH1* VCEP will continue to reassess variant classification and the rules specifications in order to reflect the most current lines of evidence available.

Assessment was performed using both the 2015 ACMG/AMP guidelines and the *CDH1* VCEP specifications (Figure 1c and Supp. Table S1). Variants were classified into the five-tier system described by Richards et al. (2015). Even though most variants tested were rare, access to large diagnostic laboratory cohorts allowed us to collect enough families to reclassify several variants using the *CDH1* specifications. For example, as in the case of the BS2 evidence code, internal laboratory data was often helpful to reclassify variants to LB or B. There were three variants for which we were able to utilize a summation of clinical evidence contributed by multiple clinical laboratories to classify as LB/B (Supp. Table S3). Thus, without sharing data, many of these variants would have remained as VUS. Being able to reclassify a variant from VUS has a significant impact on patient care, as it mitigates the ambiguity faced by patients and providers regarding cancer risks and the best course of clinical action.

Current recommendations for carriers of P/LP variants in *CDH1* include risk-reducing gastrectomy between ages 20 and 30 or annual upper endoscopy with multiple biopsies, following the dedicated Cambridge protocol for people who do not undergo gastrectomy (van der Post et al., 2015). For women, there are additional recommendations for breast MRI starting at the age of 30 and consideration of risk-reducing mastectomy (Corso et al., 2018; van der Post et al., 2015). Indeed, the complexities of assessing *CDH1* are reflected by the assessment of the ClinGen Actionability Work Group, a team that aims to take into account several facets of broader actionability for a variety of genes. In the case of *CDH1*, the assessment yielded a moderately actionable score for gastric cancer, which reflected the likely efficacy of risk-reducing gastrectomy but also its significant morbidity (https://actionability.clinicalgenome.org/redmine/projects/actionability_release/genboree_ac/ui/stg2SummaryRpt?doc=AC103).

As genomic testing becomes increasingly part of the clinical management of cancer patients and their families, it is fundamental to better understand the clinical significance of genomic variance, especially in clinically actionable genes such as *CDH1*. The effort described here to specify the criteria for *CDH1* variant classification and facilitate sharing of evidence across the ClinGen *CDH1* VCEP should accelerate accurate *CDH1* variant classification, improving our ability to identify the individuals and families that will benefit from cancer surveillance and risk-reducing surgery.

ORCID

Kate Krempely https://orcid.org/0000-0003-0311-8859
Fatima Carneiro https://orcid.org/0000-0002-1964-1006
Joana Figueiredo https://orcid.org/0000-0002-1590-1974
Arjen R. Mensenkamp https://orcid.org/0000-0003-3805-877X
Carla Oliveira https://orcid.org/0000-0001-8340-2264
Amanda B. Spurdle https://orcid.org/0000-0003-1337-7897
Kasmintan A. Schrader https://orcid.org/0000-0002-7413-4314
Rachid Karam https://orcid.org/0000-0002-5645-498X

REFERENCES

Berx, G., Cleton-Jansen, A. M., Nollet, F., de Leeuw, W. J., van de Vijver, M., Cornelisse, C., & van Roy, F. (1995). E-cadherin is a tumour/invasion suppressor gene mutated in human lobular breast cancers. *EMBO Journal*, *14*(24), 6107–6115.

Biesecker, L. G., & Harrison, S. M. (2018). The ACMG/AMP reputable source criteria for the interpretation of sequence variants. *Genetics in Medicine*, https://doi.org/10.1038/gim.2018.42

Boland, C. R., & Yurgelun, M. B. (2017). Historical perspective on familial gastric cancer. *Cellular and Molecular Gastroenterology and Hepatology*, *3*(2), 192–200. https://doi.org/10.1016/j.jcmgh.2016.12.003

Bruner, H. C., & Derksen, P. W. B. (2018). Loss of E-Cadherin-dependent cell-cell adhesion and the development and progression of cancer. *Cold Spring Harbor Perspectives in Biology*, *10*(3). pii, a029330. https://doi.org/10.1101/cshperspect.a029330

Corso, G., Marrelli, D., & Roviello, F. (2013). E-cadherin germline missense mutations in diffuse gastric cancer. *OA Cancer*, *201*(1), 4.

Corso, G., Figueiredo, J., La Vecchia, C., Veronesi, P., Pravettoni, G., Macis, R., … Galimberti, V. (2018). Hereditary lobular breast cancer with an emphasis on E-cadherin genetic defect. *Journal of Medical Genetics*, *55*(7), 431–441. https://doi.org/10.1136/jmedgenet-2018-105337

Farber-Katz, S., Hsuan, V., Wu, S., Landrith, T., Vuong, H., Xu, D., … Karam, R. (2018). Quantitative analysis of BRCA1 and BRCA2 germline splicing variants using a novel RNA-massively parallel sequencing assay. *Frontiers in Oncology*, https://doi.org/10.3389/fonc.2018.00286. eCollection 2018.

Figueiredo, J., Soderberg, O., Simoes-Correia, J., Grannas, K., Suriano, G., & Seruca, R. (2013). The importance of E-cadherin binding partners to evaluate the pathogenicity of E-cadherin missense mutations associated to HDGC. *European Journal of Human Genetics*, *21*(3), 301–309. https://doi.org/10.1038/ejhg.2012.159

Fitzgerald, R. C., & Caldas, C. (2004). Clinical implications of E-cadherin associated hereditary diffuse gastric cancer. *Gut*, *53*(6), 775–778.

Frebourg, T., Oliveira, C., Hochain, P., Karam, R., Manouvrier, S., Graziadio, C., … Seruca, R. (2006). Cleft lip/palate and *CDH1*/E-cadherin mutations in families with hereditary diffuse gastric cancer. *Journal of Medical Genetics*, 43(2), 138–142. https://doi.org/10.1136/jmg.2005.031385

Garber, K. B., Vincent, L. M., Alexander, J. J., Bean, L. J., Bale, S., & Hegde, M. (2016). Reassessment of genomic sequence variation to harmonize interpretation for personalized medicine. *American Journal of Human Genetics*, 99(3), 1140–1149.

Gelb, B. D., Cavé, H., Dillon, M. W., Gripp, K. W., Lee, J. A., Mason-Suares, H., … Vincent, L. M. (2018). ClinGen's RASopathy Expert Panel consensus methods for variant interpretation. *Genetics in Medicine*, https://doi.org/10.1038/gim.2018.3

Guggenheim, D. E., & Shah, M. A. (2013). Gastric cancer epidemiology and risk factors. *Journal of Surgical Oncology*, 107(3), 230–236. https://doi.org/10.1002/jso.23262

Guilford, P., Hopkins, J., Harraway, J., McLeod, M., McLeod, N., Harawira, P., … Reeve, A. E. (1998). E-cadherin germline mutations in familial gastric cancer. *Nature*, 392(6674), 402–405. https://doi.org/10.1038/32918

Gul, I. S., Hulpiau, P., Saeys, Y., & van Roy, F. (2017). Evolution and diversity of cadherins and catenins. *Experimental Cell Research*, 358(1), 3–9. https://doi.org/10.1016/j.yexcr.2017.03.001

Hansford, S., Kaurah, P., Li-Chang, H., Woo, M., Senz, J., Pinheiro, H., … Huntsman, D. G. (2015). Hereditary diffuse gastric cancer syndrome: *CDH1* mutations and beyond. *JAMA Oncology*, 1(1), 23–32. https://doi.org/10.1001/jamaoncol.2014.168

Harrison, S., Dolinsky, J. S., Johnson, Knight, A. E., Pesaran, T., Azzariti, D. R., Bale, S., … Rehm, H. L. (2017). Clinical laboratories collaborate to resolve differences in variant interpretations submitted to ClinVar. *Genetics in Medicine*, 19, 1096–1104.

Holbrook, J. A., Neu-Yilik, G., Hentze, M. W., & Kulozik, A. E. (2004). Nonsense-mediated decay approaches the clinic. *Nature Genetics*, 36(8), 801–808. https://doi.org/10.1038/ng1403

Inoue, M., & Tsugane, S. (2005). Epidemiology of gastric cancer in Japan. *Postgraduate Medical Journal*, 81(957), 419–424. https://doi.org/10.1136/pgmj.2004.029330

Karam, R., Carvalho, J., Bruno, I., Graziadio, C., Senz, J., Huntsman, D., … Oliveira, C. (2008). The NMD mRNA surveillance pathway downregulates aberrant E-cadherin transcripts in gastric cancer cells and in *CDH1* mutation carriers. *Oncogene*, 27(30), 4255–4260. https://doi.org/10.1038/onc.2008.62

Karam, R., Pesaran, T., & Chao, E. (2015). ClinGen and genetic testing. *New England Journal of Medicine*, 373(14), 1376–1377. https://doi.org/10.1056/NEJMc1508700

Kaurah, P., MacMillan, A., Boyd, N., Senz, J., De Luca, A., Chun, N., … Huntsman, D. (2007). Founder and recurrent *CDH1* mutations in families with hereditary diffuse gastric cancer. *JAMA*, 297(21), 2360–2372. https://doi.org/10.1001/jama.297.21.2360

Keller, G., Vogelsang, H., Becker, I., Hutter, J., Ott, K., Candidus, S., … Hofler, H. (1999). Diffuse type gastric and lobular breast carcinoma in a familial gastric cancer patient with an E-cadherin germline mutation. *American Journal of Pathology*, 155(2), 337–342. https://doi.org/10.1016/S0002-9440(10)65129-2

Kelly, M. A., Caleshu, C., Morales, A., Buchan, J., Wolf, Z., Harrison, S. M., … Funke, B. (2018). Adaptation and validation of the ACMG/AMP variant classification framework for MYH7-associated inherited cardiomyopathies: Recommendations by ClinGen's Inherited Cardiomyopathy Expert Panel. *Genetics in Medicine*, 20(3), 351–359. https://doi.org/10.1038/gim.2017.218

Kourtidis, A., Lu, R., & Anastasiadis, P. Z. (2017). A central role for cadherin signaling in cancer. *Experimental Cell Research*, 358, 78–85.

Krempely, K., & Karam, R. (2018). A novel de novo *CDH1* germline variant aides in the classification of carboxy-terminal E-cadherin alterations predicted to escape nonsense-mediated mRNA decay. *Cold Spring Harbor Molecular Case Study*, 5, a003012. https://doi.org/10.1101/mcs.a003012

LaDuca, H., Stuenkel, A. J., Dolinsky, J. S., Keiles, S., Tandy, S., Pesaran, T., … Chao, E. (2014). Utilization of multigene panels in hereditary cancer predisposition testing: Analysis of more than 2,000 patients. *Genetics in Medicine*, 16(11), 830–837. https://doi.org/10.1038/gim.2014.40

Lek, M., Karczewski, K. J., Minikel, E. V., Samocha, K. E., Banks, E., Fennell, T., … Exome Aggregation, C. (2016). Analysis of protein-coding genetic variation in 60,706 humans. *Nature*, 536(7616), 285–291. https://doi.org/10.1038/nature19057

Li, C. I., Anderson, B. O., Daling, J. R., & Moe, R. E. (2003). Trends in incidence rates of invasive lobular and ductal breast carcinoma. *JAMA*, 289(11), 1421–1424.

Lowstuter, K., Espenschied, C. R., Sturgeon, D., Ricker, C., Karam, R., LaDuca, H., … Gruber, S. B. (2017). Unexpected *CDH1* mutations identified on multigene panels pose clinical management challenges. *JCO Precision Oncology*, 1–12. https://doi.org/10.1200/PO.16.00021

Nussbaum, R. L., Rehm, H. L. & ClinGen. (2015). ClinGen and genetic testing. *New England Journal of Medicine*, 373(14), 1379. https://doi.org/10.1056/NEJMc1508700

Oliveira, C., Pinheiro, H., Figueiredo, J., Seruca, R., & Carneiro, F. (2015). Familial gastric cancer: Genetic susceptibility, pathology, and implications for management. *Lancet Oncology*, 16(2), e60–70. https://doi.org/10.1016/S1470-2045(14)71016-2

Oliveira, C., Seruca, R., & Carneiro, F. (2009). Hereditary gastric cancer. *Best Practice and Research: Clinical Gastroenterology*, 23(2), 147–157. https://doi.org/10.1016/j.bpg.2009.02.003

Oliveira, C., Seruca, R., Hoogerbrugge, N., Ligtenberg, M., & Carneiro, F. (2013). Clinical utility gene card for: Hereditary diffuse gastric cancer (HDGC). *European Journal of Human Genetics*, 21(8). https://doi.org/10.1038/ejhg.2012.247

Pinheiro, H., Oliveira, C., Seruca, R., & Carneiro, F. (2014). Hereditary diffuse gastric cancer: Pathophysiology and clinical management. *Best Practice and Research: Clinical Gastroenterology*, 28(6), 1055–1068. https://doi.org/10.1016/j.bpg.2014.09.007

Plon, S. E., Eccles, D. M., Easton, D., Foulkes, W. D., Genuardi, M., Greenblatt, M. S., … Group, I. U. G. V. W. (2008). Sequence variant classification and reporting: Recommendations for improving the interpretation of cancer susceptibility genetic test results. *Human Mutation*, 29(11), 1282–1291. https://doi.org/10.1002/humu.20880

Rehm, H. L., Berg, J. S., Brooks, L. D., Bustamante, C. D., Evans, J. P., Landrum, M. J., … ClinGen. (2015). ClinGen—The Clinical genome resource. *New England Journal of Medicine*, 372(23), 2235–2242. https://doi.org/10.1056/NEJMsr1406261

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., … Committee, A. L. Q. A. (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. *Genetics in Medicine*, 17(5), 405–424. https://doi.org/10.1038/gim.2015.30.

Richardson, M. E., Chong, H., Mu, W., Conner, B. R., … Elliott, A. (2018). DNA breakpoint assay reveals a majority of gross duplications occur in tandem reducing VUS classifications in breast cancer predisposition genes. *Genetics in Medicine*, https://doi.org/10.1038/s41436-018-0092-7

Riethmacher, D., Brinkmann, V., & Birchmeier, C. (1995). A targeted mutation in the mouse E-cadherin gene results in defective

preimplantation development. *Proceedings of the National Academy of Sciences in the United States of America*, 92(3), 855–859.

Rivas, M. A., Pirinen, M., Conrad, D. F., Lek, M., Tsang, E. K., Karczewski, K. J., … MacArthur, D. G. (2015). Human genomics. Effect of predicted protein-truncating genetic variants on the human transcriptome. *Science*, 348(6235), 666–669. https://doi.org/10.1126/science.1261877

Sasaki, C. Y., Lin, H., Morin, P. J., & Longo, D. L. (2000). Truncation of the extracellular region abrogates cell contact but retains the growth-suppressive activity of E-cadherin. *Cancer Research*, 60(24), 7057–7065.

Shah, M. A., Salo-Mullen, E., Stadler, Z., Ruggeri, J. M., Mirander, M., Pristyazhnyuk, Y., & Zhang, L. (2012). De novo *CDH1* mutation in a family presenting with early-onset diffuse gastric cancer. *Clinical Genetics*, 82(3), 283–287. https://doi.org/10.1111/j.1399-0004.2011.01744.x

Simoes-Correia, J., Figueiredo, J., Lopes, R., Stricher, F., Oliveira, C., Serrano, L., & Seruca, R. (2012). E-cadherin destabilization accounts for the pathogenicity of missense mutations in hereditary diffuse gastric cancer. *PLoS One*, 7(3), e33783. https://doi.org/10.1371/journal.pone.0033783

Simoes-Correia, J., Figueiredo, J., Oliveira, C., van Hengel, J., Seruca, R., van Roy, F., & Suriano, G. (2008). Endoplasmic reticulum quality control: A new mechanism of E-cadherin regulation and its implication in cancer. *Human Molecular Genetics*, 17(22), 3566–3576. https://doi.org/10.1093/hmg/ddn249

Sugimoto, S., Yamada, H., Takahashi, M., Morohoshi, Y., Yamaguchi, N., Tsunoda, Y., … Komatsu, H. (2014). Early-onset diffuse gastric cancer associated with a de novo large genomic deletion of *CDH1* gene. *Gastric Cancer*, 17(4), 745–749. https://doi.org/10.1007/s10120-013-0278-2

Suriano, G., Mulholland, D., de Wever, O., Ferreira, P., Mateus, A. R., Bruyneel, E., … Seruca, R. (2003). The intracellular E-cadherin germline mutation V832 M lacks the ability to mediate cell-cell adhesion and to suppress invasion. *Oncogene*, 22(36), 5716–5719. https://doi.org/10.1038/sj.onc.1206672

Suriano, G., Oliveira, C., Ferreira, P., Machado, J. C., Bordin, M. C., De Wever, O., … Seruca, R. (2003). Identification of *CDH1* germline missense mutations associated with functional inactivation of the E-cadherin protein in young gastric cancer probands. *Human Molecular Genetics*, 12(5), 575–582.

Suriano, G., Oliveira, M. J., Huntsman, D., Mateus, A. R., Ferreira, P., Casares, F., … Seruca, R. (2003). E-cadherin germline missense mutations and cell phenotype: Evidence for the independence of cell invasion on the motile capabilities of the cells. *Human Molecular Genetics*, 12(22), 3007–3016. https://doi.org/10.1093/hmg/ddg316

Suriano, G., Seixas, S., Rocha, J., & Seruca, R. (2006). A model to infer the pathogenic significance of *CDH1* germline missense variants. *Journal of Molecular Medicine (Berlin)*, 84(12), 1023–1031. https://doi.org/10.1007/s00109-006-0091-z

Takeichi, M. (1991). Cadherin cell adhesion receptors as a morphogenetic regulator. *Science*, 251(5000), 1451–1455.

van der Post, R. S., Vogelaar, I. P., Carneiro, F., Guilford, P., Huntsman, D., Hoogerbrugge, N., … Fitzgerald, R. C. (2015). Hereditary diffuse gastric cancer: Updated clinical guidelines with an emphasis on germline *CDH1* mutation carriers. *Journal of Medical Genetics*, 52(6), 361–374. https://doi.org/10.1136/jmedgenet-2015-103094

van der Post, R. S., Vogelaar, I. P., Manders, P., van der Kolk, L. E., Cats, A., van Hest, L. P., … Ligtenberg, M. J. (2015). Accuracy of hereditary diffuse gastric cancer testing criteria and outcomes in patients with a germline mutation in *CDH1*. *Gastroenterology*, 149(4), 897–906, e819. https://doi.org/10.1053/j.gastro.2015.06.003

Whiffin, N., Minikel, E., Walsh, R., O'Donnell-Luria, A. H., Karczewski, K., Ing, A. Y., … Ware, J. S. (2017). Using high-resolution variant frequencies to empower clinical genome interpretation. *Genetics in Medicine*, 19(10), 1151–1158. https://doi.org/10.1038/gim.2017.26

Yelskaya, Z., Bacares, R., Salo-Mullen, E., Somar, J., Lehrich, D. A., … Zhang, L. (2016). *CDH1* missense variant c.1679C>G (p.T560R) completely disrupts normal splicing through creation of a Novel 5' splice site. *PLoS One*, 11(11), e0165654. https://doi.org/10.1371/journal.pone.0165654

Yelskaya, Z., Somar, J., Bacares, R., Salo-Mullen1, E. E., Yang, C., … Zhang, L. (2016). *CDH1* splice site variant c.387+1G>A emphasizes the importance of functional studies. *Journal of Molecular Diagnostics*, 8(6), 941. https://doi.org/10.1016/S1525-1578(16)30178-7

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.