InSiGHT Variant Interpretation Committee: Mismatch Repair Gene Variant Classification Criteria

Rules for Variant Classification:
Rules describing the 5 class system for classification of MMR gene variants were devised and documented by Amanda Spurdle and Bryony Thompson in September 2009 for standardised classification of variants in the Colon Cancer Family Registry database, and revised by Amanda Spurdle, Bryony Thompson, Sean Tavtigian and Marc Greenblatt during April 2011, in collaboration with the InSiGHT Variant Interpretation Committee (VIC), modulated by input from committee members during ongoing VIC meetings. They are based on the following:

- The 5 class system described for quantitative assessment of variant pathogenicity in Plon et al.[1], using a multifactorial likelihood model[2-5] as applied to MMR gene variants[6,7];
- The 5 class system for interpretation of splicing variants and aberrations by Spurdle et al.[8];
- The classification of sequence changes according to standard clinical practice – that is, description of variants generally considered pathogenic (clinically relevant in a genetic counselling setting such that germline variant status is used to inform patient and family management) or non-pathogenic (significant evidence against being a dominant high-risk pathogenic mutation); and
- The documentation of non-quantitative methods that have been used to classify variants in the literature.

For a given class, a variant is required to satisfy all the criteria listed for at least one bullet-point that falls within that class. The symbol "✓" represents an "AND" statement. The footnote # describes the rationale for suggested sample numbers, and the rationale for use of indicative MSI or IHC information. The interpretation of functional assays is assisted by a flowchart developed for this purpose (Figure 1).

These criteria provide a baseline for standardized clinical classification of MMR gene sequence variation that may be linked to patient and family management in the genetic counseling arena according to published guidelines [1]. Use of the InSiGHT database, and associated interpretation relating to pathogenicity, is subject to User discretion and responsibility. Whereas InSiGHT has developed processes to assign pathogenicity utilizing high-level expertise available through its membership, such assignments are subject to change with the availability of new information and interpretation processes. Information submitted to the InSiGHT database is available for individual enquiry for clinical use, but any collective use of the data for research or other purposes transgresses agreements InSiGHT has with the valuable community of researchers, clinicians and diagnostic laboratories that generously support the database through their submissions. All users of the database are encouraged to submit their own variants to support the InSiGHT international collaboration to share gene variant data relating to gastrointestinal tumors. Submissions should be directed to the InSiGHT curator John-Paul Plazzer (johnpaul@variome.org).

Class 5 – Pathogenic

- Variants with probability of pathogenicity >0.99 using a multifactorial likelihood model
- Coding sequence variation resulting in a stop codon i.e. nonsense or frameshift pathogenic alteration that is predicted to result in interruption of known functional protein domains or highly conserved exonic sequences
- Variants, where mRNA assays in **patient RNA** indicate that the variant allele results in a splicing aberration (with evidence that the variant allele produces no wild-type transcript) leading to premature stop codon or in-frame deletion disrupting a functional domain or protein conformation
- Large genomic deletions
- Large genomic duplications shown by laboratory studies to result in a frameshift before the last splice junction
- Variants demonstrating all of the following characteristics with no conflicting results (combined evidence achieves estimated LR >100:1, posterior >0.99 with prior 0.5):
  - ✓ Variant-specific abrogated function in protein or mRNA based lab assays (see MMR functional assay supplementary material)
    - OR - for *MLH1* and *MSH2*, co-occurrence *in trans* with a known pathogenic sequence variant in the same gene in a patient with clinical features consistent with Constitutional Mismatch Repair Deficiency (CMMRD) and reported LS cancer in the parent of origin for the variant to be classified. If parental genotype is unknown, then both parents should have early-onset LS cancer
    - OR - presence of the variant on different haplotypes across families indicating that reported LS clinical features are not due to an undiscovered sequence change *in cis* with the variant
  - ✓ Evidence for co-segregation with disease where pedigree information provided allows calculation of likelihood ratio (LR) of ≥10:1*****
    - OR - at least one revised Amsterdam criteria [9] family with ≥4 affected carriers
    - OR - across ≥2 revised Amsterdam criteria families reported to show segregation with disease (no further information provided)
    - OR - ≥2 families with ≥3 affected non-proband carriers
  - ✓ ≥2 independent tumors with MSI using a standard panel of 5-10 markers**** and/or** loss of MMR protein expression consistent with the variant location (may include tumor information from proband)
  - ✓ evidence that variant is not an undescribed polymorphism at allele frequency >1% (minor allele frequency [MAF] >0.01) in healthy controls from an appropriate population (i.e. 1000 genomes)

Class 4 – Likely pathogenic

- Variants with probability of pathogenicity between 0.95-0.99 using a multifactorial likelihood model
- Variants at IVS±1 or IVS±2, or G>non-G at last base of exon if first 6 bases of the intron are <u>not</u> GTRRGT, that are untested for splicing aberrations *in vitro* - *irrespective of bioinformatic predictions*
- Variants demonstrating (combined evidence achieves LR >20:1, posterior 0.95-0.99 with prior 0.5):
  - ✓ variant-specific abrogated function in protein or mRNA based lab assays (see MMR functional assay supplementary material)
    - OR - for *MLH1* and *MSH2*, co-occurrence *in trans* with a known pathogenic sequence variant in the same gene in a patient with clinical features consistent with CMMRD and reported LS cancer in the parent of origin for the variant to be classified. If parental genotype is unknown, then both parents should have early-onset LS cancer
    - OR - presence of the variant on different haplotypes across families indicating that reported LS clinical features are not due to an undiscovered sequence change *in cis* with the variant
  - ✓ **plus one of the following:**
    - co-segregation with disease assessed by available pedigree information allows calculation of LR of >5:1*****
    - OR - at least one revised Amsterdam criteria family with ≥3 affected carriers
    - OR - ≥2 families with ≥2 affected non-proband carriers
    - OR - ≥2 independent tumours with MSI using a standard panel of 5-10 markers**** and/or** loss of MMR protein expression consistent with the variant location (may include tumour information from proband)

Class 3 – Uncertain

- Variants with probability of pathogenicity between 0.05-0.949 using a multifactorial likelihood model
- Variants that have insufficient evidence (molecular or otherwise) to classify, which may include large genomic duplications not yet shown by laboratory studies to result in a frameshift before the last splice junction, missense alterations, small in-frame insertions/deletions, silent variants, intronic variants, promoter and regulatory region variants

Class 2 – Likely not pathogenic/little clinical significance

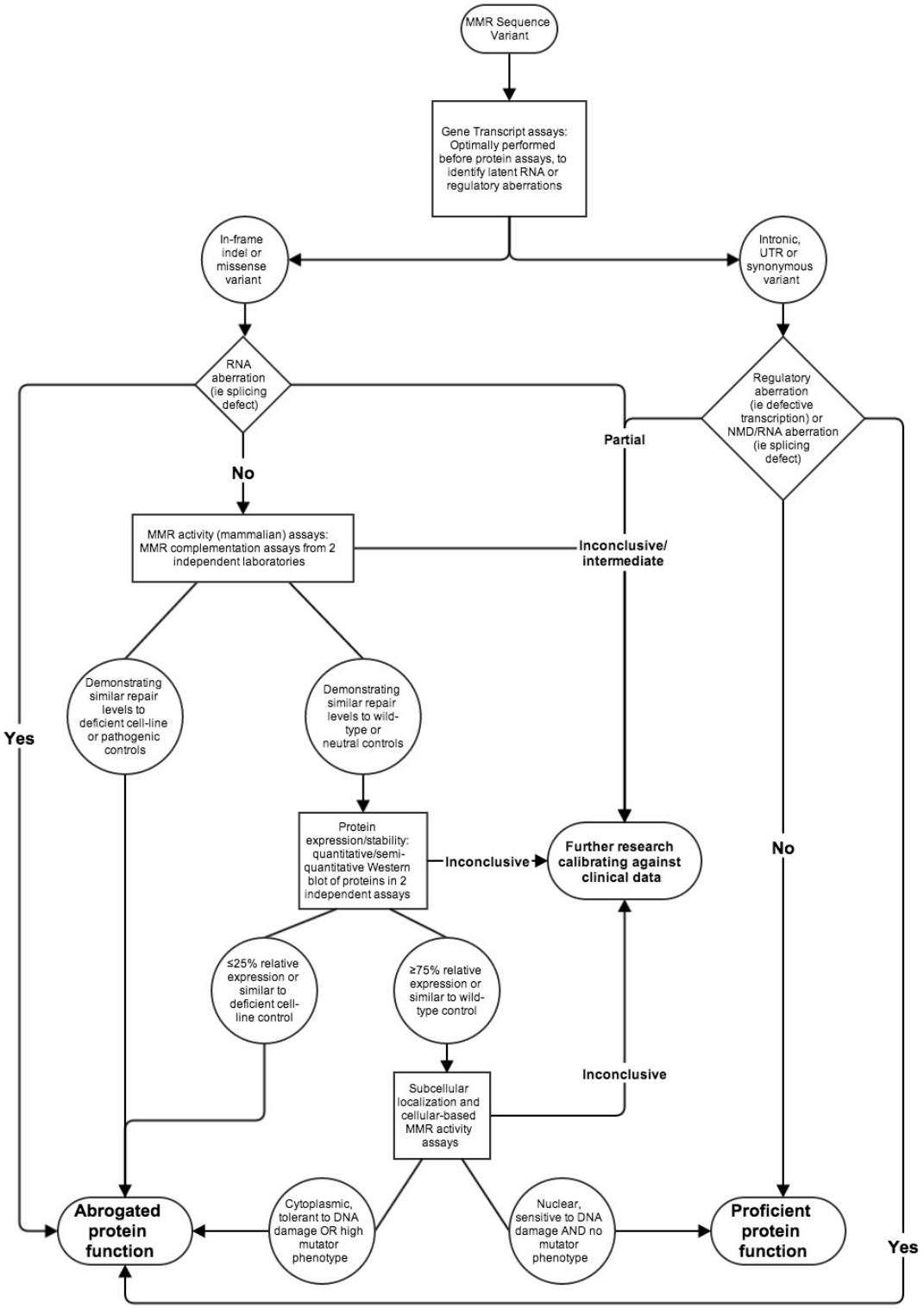
- Variants with probability of pathogenicity between 0.001-0.049 using a multifactorial likelihood model
- Synonymous substitutions and intronic variants with no associated mRNA aberration (either splicing or allelic imbalance) as determined by laboratory assays conducted with nonsense-mediated decay inhibition. Whenever abnormal transcripts are identified at similar levels in controls they will be considered to be naturally occurring isoforms and not mRNA aberrations
- Variants reported to occur in a specific ethinic group at allele frequency ≥1%, (MAF ≥ 0.01, tested in ≥ 160 individuals) and that have not yet been excluded as known founder pathogenic sequence variants ("founder mutations")
- Variants demonstrating:
  - ✓ variant-specific proficient function in protein and mRNA based lab assays (see MMR functional assay supplementary material)
    - OR - co-occurrence *in trans* with a known pathogenic sequence variant in the same gene in a patient with colorectal cancer after age 45 (or other LS cancer above the median age of onset for that cancer in LS***), and who has no previous or current evidence of clinical manifestations of CMMRD
  - ✓ **plus one of the following:**
    - present in control reference groups at allele frequency 0.01-1% (MAF 0.0001-0.01, tested in ≥ 160 individuals)
    - OR - lack of co-segregation with disease consistent with a dominant high-risk pathogenic sequence variant in pedigrees (LR≤ 0.01)
    - OR - estimated risk <1.7, with upper bound 95% confidence limit <4, as determined by large well-designed case-control studies that consider size, geography/ethnicity and quality control measures
    - OR - ≥3 colorectal tumors with MSS **and/or** no loss of MMR protein expression **and/or** loss of MMR protein(s) in LS spectrum tumors, that is inconsistent with the gene demonstrating genetic variation
- Variants demonstrating any **three** of the following:
  - present in control reference groups at allele frequency 0.01-1% (MAF 0.0001-0.01, tested in ≥ 160 individuals)

  - **OR -** lack of co-segregation with disease consistent with a dominant high-risk pathogenic sequence variant in pedigrees (LR≤ 0.01)
  - OR - estimated risk <1.7, with upper bound 95% confidence limit <4, as determined by large well-designed case-control studies that consider size, geography/ethnicity and quality control measures
  - OR - ≥3 colorectal tumors with MSS **and/or** no loss of MMR protein expression **and/or** loss of MMR protein(s) in LS spectrum tumors, that is inconsistent with the gene demonstrating genetic variation

Class1 – Not pathogenic/low clinical significance

- Variants with probability of pathogenicity <0.001 using a multifactorial likelihood model
- Variants reported to occur in control reference groups at allele frequency ≥1% (MAF ≥ 0.01, tested in ≥ 160 individuals) and excluded as founder pathogenic sequence variants
- Variants demonstrating:
  - ✓ Variant-specific proficient function in protein and mRNA based lab assays (see MMR functional assay supplementary material)
    - OR - co-occurrence *in trans* with a known pathogenic sequence variant in the same gene in a patient with colorectal cancer after age 45 (or other LS cancer above the median age of onset for that cancer in LS***), and who has no previous or current evidence of clinical manifestations of CMMRD
  - ✓ **Plus any two of the following:**
    - present in control reference groups at allele frequency 0.01-1% (MAF 0.0001-0.01, tested in ≥ 160 individuals)
    - OR - lack of co-segregation with disease consistent with a dominant high-risk pathogenic sequence variant in pedigrees (LR≤ 0.01)
    - OR - estimated risk <1.5, with upper bound 95% confidence limit <4, as determined by large well-designed case-control studies that consider size, geography/ethnicity and quality control measures
    - OR - ≥3 colorectal tumors with MSS **and/or** no loss of MMR protein expression **and/or** loss of MMR protein(s) in LS spectrum tumors, that is inconsistent with the gene demonstrating genetic variation
- Variants demonstrating all of the following:
  Present in control reference groups at allele frequency 0.01-1% (MAF 0.0001-0.01, tested in ≥ 160 individuals)
  - ✓ Lack of co-segregation with disease consistent with a dominant high-risk pathogenic sequence variant in pedigrees (LR≤ 0.01)
  - ✓ Estimated risk <1.5, with upper bound 95% confidence limit <4, as determined by large well-designed case-control studies that consider size, geography/ethnicity and quality control measures
  - ✓ ≥3 colorectal tumors with MSS **and/or** no loss of MMR protein expression **and/or** loss of MMR protein(s) in LS spectrum tumors, that is inconsistent with the gene demonstrating genetic variation

Figure 1. Flowchart used to assist in interpretation of available functional assay data. For variants that had normal/inconclusive/intermediate MMR activity in 2 independent assays, but deficient protein function in 2 independent assays, abrogated function was assigned. NMD – nonsense mediated decay.



Footnotes to Mismatch Repair Gene Variant Classification Guidelines:

The upper 95% confidence limit of frequency for an allele observed once in 160 individuals (320 chromosomes) is <0.01, the allele frequency considered sufficient for classification as class 1 (not pathogenic/low clinical significance). Characteristics of a pathogenic variant were selected to achieve a combined likelihood ratio of >100:1 in conjunction with a prior of 0.5 (estimated prior irrespective of *in silico* predictions). Namely, the co-segregation and family history descriptions were selected to estimate segregation odds of 10:1, results from functional assays were assumed to have likelihood ratio 5:1 and MSI/IHC data was conservatively assumed to carry a likelihood ratio of 5:1. Tumor data was based on two tumors for MSI and/**or** IHC to allow for the possibility that pathogenic missense alterations may not all demonstrate loss of MMR protein expression. Where both MSI and IHC information are available, MSI-H results will take precedence over normal immunohistochemical results, since MSI is more specific than MMR immunohistochemistry in colorectal cancers[10,11]. For variants that demonstrate tumor MSS and loss of MMR protein expression, technical or other explanations[10,12-14] for the discrepancy should be investigated. The need for multiple results was implemented to minimize the chance of a "sporadic" MSI-H or negative immunohistochemical result for MLH1 methylated tumors (0.15x0.15 = 0.0225). Independent tumors can include multiple primary tumors from a single individual.

***** Likelihood ratios for segregation can be derived by Bayes factor analysis adapted from the method of Thompson et al[15], as described previously[6]. Penetrance estimates for *MLH1* and *MSH2* variants to be derived from Quehenberger et al[16] and those for *MSH6* and *PMS2* variants to be derived from Baglietto et al[17] and Senter et al[18].

****** Standard MSI markers panel: BAT25, BAT26, BAT40, BAT34, D5S346, D17S250, ACTC, D18S55, D10S197, MYCL [19]; D2S123, D18S69 [20]; NR21, NR24, NR27 [21]

******* Lynch syndrome tumors include: colorectal/colon/rectal, endometrial, ovarian, small bowel/small intestine, renal pelvis, ureter, and stomach/gastric carcinomas, sebaceous skin tumors (adenomas and carcinomas), gliomas.

1. Plon, S.E. *et al.* Sequence variant classification and reporting: recommendations for improving the interpretation of cancer susceptibility genetic test results. *Hum Mutat* **29**, 1282-91 (2008).
2. Goldgar, D.E. *et al.* Integrated evaluation of DNA sequence variants of unknown clinical significance: application to BRCA1 and BRCA2. *Am J Hum Genet* **75**, 535-44 (2004).
3. Easton, D.F. *et al.* A systematic genetic assessment of 1,433 sequence variants of unknown clinical significance in the BRCA1 and BRCA2 breast cancer-predisposition genes. *Am J Hum Genet* **81**, 873-83 (2007).
4. Goldgar, D.E. *et al.* Genetic evidence and integration of various data sources for classifying uncertain variants into a single model. *Hum Mutat* **29**, 1265-72 (2008).
5. Tavtigian, S.V., Greenblatt, M.S., Lesueur, F. & Byrnes, G.B. In silico analysis of missense substitutions using sequence-alignment based methods. *Hum Mutat* **29**, 1327-36 (2008).

6. Arnold, S. *et al.* Classifying MLH1 and MSH2 variants using bioinformatic prediction, splicing assays, segregation, and tumor characteristics. *Hum Mutat* **30**, 757-70 (2009).
7. Spurdle, A.B. Clinical relevance of rare germline sequence variants in cancer genes: evolution and application of classification models. *Curr Opin Genet Dev* **20**, 315-23 (2010).
8. Spurdle, A.B., Couch, F.J., Hogervorst, F.B., Radice, P. & Sinilnikova, O.M. Prediction and assessment of splicing alterations: implications for clinical testing. *Hum Mutat* **29**, 1304-13 (2008).
9. Vasen, H.F., Watson, P., Mecklin, J.P. & Lynch, H.T. New clinical criteria for hereditary nonpolyposis colorectal cancer (HNPCC, Lynch syndrome) proposed by the International Collaborative group on HNPCC. *Gastroenterology* **116**, 1453-6 (1999).
10. Lucci-Cordisco, E., Boccuto, L., Neri, G. & Genuardi, M. The use of microsatellite instability, immunohistochemistry and other variables in determining the clinical significance of MLH1 and MSH2 unclassified variants in Lynch syndrome. *Cancer Biomark* **2**, 11-27 (2006).
11. Mangold, E. *et al.* Tumours from MSH2 mutation carriers show loss of MSH2 expression but many tumours from MLH1 mutation carriers exhibit weak positive MLH1 staining. *J Pathol* **207**, 385-95 (2005).
12. Pastrello, C. *et al.* Stability of BAT26 in tumours of hereditary nonpolyposis colorectal cancer patients with MSH2 intragenic deletion. *Eur J Hum Genet* **14**, 63-8 (2006).
13. Samowitz, W.S. & Slattery, M.L. Missense mismatch repair gene alterations, microsatellite instability, and hereditary nonpolyposis colorectal cancer. *J Clin Oncol* **20**, 3178; author reply 3178-9 (2002).
14. Xie, J. *et al.* An MLH1 Mutation Links BACH1/FANCJ to Colon Cancer, Signaling, and Insight toward Directed Therapy. *Cancer Prev Res (Phila)* **3**, 1409-1416 (2010).
15. Thompson, D., Easton, D.F. & Goldgar, D.E. A full-likelihood method for the evaluation of causality of sequence variants from family data. *Am J Hum Genet* **73**, 652-5 (2003).
16. Quehenberger, F., Vasen, H.F. & van Houwelingen, H.C. Risk of colorectal and endometrial cancer for carriers of mutations of the hMLH1 and hMSH2 gene: correction for ascertainment. *J Med Genet* **42**, 491-6 (2005).
17. Baglietto, L. *et al.* Risks of Lynch syndrome cancers for MSH6 mutation carriers. *J Natl Cancer Inst* **102**, 193-201 (2010).
18. Senter, L. *et al.* The clinical phenotype of Lynch syndrome due to germ-line PMS2 mutations. *Gastroenterology* **135**, 419-28 (2008).
19. Lindor, N.M. *et al.* Immunohistochemistry versus microsatellite instability testing in phenotyping colorectal tumors. *J Clin Oncol* **20**, 1043-8 (2002).
20. Muller, W. *et al.* The reliability of immunohistochemistry as a prescreening method for the diagnosis of hereditary nonpolyposis colorectal cancer (HNPCC)--results of an international collaborative study. *Fam Cancer* **1**, 87-92 (2001).
21. Buhard, O. *et al.* Multipopulation analysis of polymorphisms in five mononucleotide repeats used to determine the microsatellite instability status of human tumors. *J Clin Oncol* **24**, 241-51 (2006).