

Impact of imperfect diagnosis in ME/CFS association studies

João Malato^{1,2,*}, Luís Graça¹ and Nuno Sepúlveda^{2,3}

¹*Instituto de Medicina Molecular João Lobo Antunes, Faculty of Medicine, University of Lisbon, Lisbon, Portugal*

²*CEAUL—Centro de Estatística e Aplicações da Universidade de Lisboa, University of Lisbon, Lisbon, Portugal*

³*Faculty of Mathematics and Information Science, Warsaw University of Technology, Warsaw, Poland*

Correspondence*:

Corresponding Author

jmalato@medicina.ulisboa.pt

2 ABSTRACT

3 The absence of an objective disease biomarker puts studies of Myalgic Encephalomyelitis/Chronic
4 Fatigue Syndrome (ME/CFS) under the curse of imperfect diagnosis. This problem leads to
5 frequent reports that fail to reproduce prior published studies. To address the impact of imperfect
6 diagnosis on the robustness of studies' conclusions, we conducted a simulation study to quantify
7 the statistical power to detect a disease association with a hypothetical binary factor in the
8 presence of imperfect diagnosis. Using the classical case-control design, studies with sample
9 sizes of less than 500 individuals per group could not reach the target power of at least 80% to
10 detect realistic disease associations. We then recreated serological association studies in which
11 the chance of imperfect diagnosis was combined with the probability of misclassifying a binary
12 factor, as it happens in a typical serological association study. In this case, the target power
13 of 80% could only be achieved for studies with more than 1000 individuals per group. Given
14 the current sample sizes of ME/CFS studies, our results suggest that most studies are likely to
15 be underpowered due to imperfect diagnosis alone. To increase reproducibility across studies,
16 we provided some practical recommendations, such as the use of standard case definitions
17 together with multi-centric study designs, and routine reporting of power calculations under a
18 non-negligible chance of misdiagnosis. Our results can also inform the design of future studies
19 under the assumption of misdiagnosis.

20 **Keywords:** misdiagnosis, misclassification, association studies, simulation, power studies, myalgic encephalomyelitis, chronic fatigue
21 syndrome

1 INTRODUCTION

22 Myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS) is a complex disease whose patients
23 manifest unexplained long-lasting fatigue or post-exertional malaise upon minimal physical or mental
24 effort (1, 2). The disease has a sex bias towards women, with a ratio of about two women to one man
25 among cases (3). The real burden of the disease remains unknown, but current prevalence estimates range
26 from 0.1% to 2.2% (4).

27 Given the lack of a biomarker for the disease, the diagnosis of suspected cases is made by a set of key
28 symptoms together with the exclusion of known diseases, such as multiple sclerosis and diabetes, which

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

29 could explain fatigue (5). To make the disease diagnosis even more cumbersome, there are multiple case
30 definitions proposed in the literature (6). A possible solution to this problem is to follow expert consensus
31 recommendations for clinical and research practices. In this regard, the European Network on ME/CFS
32 (EUROMENE) suggests the use of the Centre for Disease Control (Fukuda/CDC) criterion (1) and the
33 Canadian Consensus Criteria (CCC) (2) for research purposes (7). In a clinical setting, the ME/CFS
34 diagnosis can also contemplate the Institute of Medicine (IOM) criterion due to its simplicity (8, 9).

35 The use of different case definitions by clinicians and researchers opens the door to not identifying the
36 same population of patients. As a consequence, it is difficult to ascertain the epidemiology and the real
37 economic costs of the disease (4, 7). Such a discrepancy in case identification has been demonstrated in a
38 study where less than 75% of patients had the same diagnostic outcome across Fukuda/CDC, CCC, and IOM
39 criterion (10). In the same study, data from a symptom assessment questionnaire alone could not completely
40 separate patients diagnosed with ME/CFS from healthy controls and patients with multiple sclerosis. Similar
41 symptoms overlap was found between ME/CFS, multiple sclerosis (11, 12), and rheumatoid arthritis (13);
42 the exclusion of these two autoimmune diseases is, however, part of current ME/CFS diagnostic criteria. In
43 this scenario, current ME/CFS diagnostic criteria are likely to have different sensitivities and are far from
44 100%.

45 In general, imperfect diagnostic tests imply the presence of estimation bias in subsequent statistical
46 inference. It also reduces the statistical power to detect true disease associations or the efficiency of
47 estimating the true prevalence (14, 15). These implications are possible explanations for the lack of
48 scientific reproducibility in ME/CFS research. In this regard, Nacul et al. (16) suggested missing important
49 disease associations due to the problem of imperfect ME/CFS diagnosis (i.e., misdiagnosis). However, this
50 study did not include a systematic analysis of the potential impact of misdiagnosis on ME/CFS-related
51 associations. In this scenario, we aimed to investigate the statistical power of hypothetical case-control
52 association studies of ME/CFS in the presence of misdiagnosis. For this purpose, we performed an
53 extensive simulation study under different sample sizes, levels of misdiagnosis, and strength of disease
54 associations. We also investigated the additional impact of sensitivity and specificity associated with
55 serological evaluations that are often performed in ME/CFS (17, 18, 19). Finally, we extended our study to
56 discuss data from two publications (20, 21).

2 STATISTICAL METHODOLOGY

57 2.1 Statistical formulation of the problem

58 To study the impact of misclassification, let us assume a classical case-control association study in which
59 diagnosed ME/CFS patients and healthy controls were matched for possible confounders such as age,
60 gender, or body mass index (BMI). The main objective of this study is to investigate a potential association
61 between a candidate causal factor (e.g., a genetic factor or the occurrence of a given infection) and ME/CFS.
62 For simplicity, let us assume that this factor has only two possible outcomes: present versus absent. This is
63 the classical epidemiological situation of a putative risk factor in which individuals can be divided into
64 exposed and non-exposed. The hypothetical data can be summarised by a 2×2 contingency table, whose
65 sampling distribution is the product of two independent Binomial distributions, one Binomial distribution
66 per group,

$$f(x_i|n_i; \theta_i) = \prod_{i=0,1} \binom{n_i}{x_i} \theta_i^{x_i} (1 - \theta_i)^{n_i - x_i},$$

67 where x_0 and x_1 are the observed frequencies of healthy controls and suspected cases with presence of the
68 candidate causal factor, respectively, n_0 and n_1 are the corresponding sample sizes of each group, and θ_0
69 and θ_1 are the probabilities for the presence of the candidate causal factor in healthy controls and suspected
70 cases, respectively.

71 To investigate whether there is evidence for an association between this candidate factor, one can perform
72 a hypothesis test in which the null and alternative hypotheses, H_0 and H_1 , are

$$H_0 : \theta_0 = \theta_1 \text{ versus } H_1 : \theta_0 \neq \theta_1 .$$

73 The same hypotheses can be rewritten in terms of the odds ratio as follows

$$H_0 : \frac{\theta_1(1 - \theta_0)}{\theta_0(1 - \theta_1)} = 1 \text{ versus } H_1 : \frac{\theta_1(1 - \theta_0)}{\theta_0(1 - \theta_1)} \neq 1 .$$

74 To test these hypotheses, one can use the Pearson's χ^2 test for independence in 2-way contingency tables.
75 Using a significance level of 0.05, p-values < 0.05 suggest a rejection of the null hypothesis.

76 To study the impact of misdiagnosis on the power of above Pearson's χ^2 test, we considered seven
77 simplifying assumptions for the statistical formulation of the problem:

- 78 I. Suspected cases are composed of both apparent and true patients;
- 79 II. The presence of the candidate causal factor is only associated with true patients;
- 80 III. Apparent cases are similar to healthy controls in what the association with the candidate causal factor
81 is concerned;
- 82 IV. The chance of making a ME/CFS misdiagnosis is only dependent on the true clinical status of the
83 cases and not on the confounders;
- 84 V. The association to be detected is independent of disease duration and disease triggers, among other
85 factors occurring during the disease course;
- 86 VI. There is no misdiagnosis (or misclassification) of the true health status of the controls;
- 87 VII. The presence or absence of the candidate causal factor is determined accurately in each individual.

88 Assumption VI is invoked to exclude the (unlikely) situation in which healthy controls could include true
89 undiagnosed ME/CFS patients. Assumption VII is appropriate for genetic association studies in which the
90 genotypic error rate is quantified and typically low (22).

91 The above assumptions allow, then, to rewrite the probability of the causal factor being present in
92 suspected cases as

$$\theta_1 = \gamma\theta_0 + (1 - \gamma)\theta_1^* , \tag{1}$$

93 where γ is the misdiagnosis (or misclassification) probability, and θ_1^* is the probability of the candidate
94 causal factor being present in true ME/CFS cases. The direct consequence of the above formula is a dilution
95 effect of the association operated by the possibility of misdiagnosis. If misdiagnosis could be directly
96 observed, the classical 2×2 contingency table could be augmented as in Table 1, where the diagnosed
97 cases are divided into apparent and true cases in accordance with assumption I.

98 A more-complex statistical situation is to consider the chance of misdiagnosis together with the
99 impossibility of determining the presence of the candidate causal factor accurately. This situation is
100 the curse of all serological studies on ME/CFS in which the objective is to evaluate the association between

Table 1. Augmented version of the observed 2×2 contingency table in presence of misdiagnosis of ME/CFS cases for a classical case-control association study.

Causal factor	Controls	ME/CFS-diagnosed cases	
		(Apparent)	(True)
Present	θ_0	$\gamma\theta_0$	$(1 - \gamma)\theta_1^*$
Absent	$1 - \theta_0$	$\gamma(1 - \theta_0)$	$(1 - \gamma)(1 - \theta_1^*)$

101 the disease and the presence of antibodies against a given infectious agent (e.g., Epstein-Barr virus (23)).
 102 In these studies, individuals are deemed seropositive or seronegative based on a threshold in the antibody
 103 distribution. When the estimated antibody distributions for these two populations overlap, the respective
 104 serological classification is subject to uncertainty ruled by the underlying sensitivity and specificity of the
 105 statistical method used (24). The description below follows as if the data at hand were generated from this
 106 type of study.

107 To model this more complex situation, previous assumption VII is replaced with the two additional
 108 assumptions:

- 109 VII. There are only two possible serological outcomes for each individual: seronegative or seropositive;
 110 VIII. The sensitivity and specificity of the serological classification are shared across all the individuals.

111 The revised assumption VII aims to exclude the situation where the serological classification can
 112 contemplate an indeterminate status due to the laboratory protocol (21) or the presence of multiple
 113 serological populations (25). Similarly to assumption V for misdiagnosis, the new assumption VIII intends
 114 to disregard the effect of confounders (i.e., age or gender) and disease-related factors (i.e., disease duration
 115 or disease severity) on the performance of the serological classification.

116 Under the assumptions I–VIII the probability of recording the presence of the candidate causal factor in a
 117 ME/CFS patient can be generalised to

$$\theta_1 = \pi_{se}\gamma\theta_0 + (1 - \pi_{sp})\gamma(1 - \theta_0) + \pi_{se}(1 - \gamma)\theta_1^* + (1 - \pi_{sp})(1 - \gamma)(1 - \theta_1^*), \quad (2)$$

118 where π_{se} and π_{sp} are sensitivity and specificity for the serological classification, respectively. Note that,
 119 when $\pi_{se} = \pi_{sp} = 1$ (perfect serological testing), the above formula converts to equation (1). Table 1 can
 120 thus be further augmented by including the true serological status of the individuals (Table 2).

Table 2. Augmented version of the observable 2×2 frequency Table in the case-control association study with possible misdiagnosis of ME/CFS cases and misclassification of the true serological status (seropositive, S^+ , and seronegative, S^-).

Estimated Serological status	True serological status	Controls	ME/CFS-diagnosed cases	
			(Apparent)	(True)
S^+	S^+	$\pi_{se}\theta_0$	$\pi_{se}\gamma\theta_0$	$\pi_{se}(1 - \gamma)\theta_1^*$
	S^-	$(1 - \pi_{sp})(1 - \theta_0)$	$(1 - \pi_{sp})\gamma(1 - \theta_0)$	$(1 - \pi_{sp})(1 - \gamma)(1 - \theta_1^*)$
S^-	S^+	$(1 - \pi_{se})\theta_0$	$(1 - \pi_{se})\gamma\theta_0$	$(1 - \pi_{se})(1 - \gamma)\theta_1^*$
	S^-	$\pi_{sp}(1 - \theta_0)$	$\pi_{sp}\gamma(1 - \theta_0)$	$\pi_{sp}(1 - \gamma)(1 - \theta_1^*)$

121 2.2 Simulation study

122 To investigate the impact of both misdiagnosis scenarios on detecting an association between a candidate
123 causal factor and ME/CFS, we performed a comprehensive simulation study assuming matched sample
124 sizes for ME/CFS patients and healthy controls ($n_0 = n_1$). We considered the following sample sizes per
125 study group: 100, 250, 500, 1000, 2500, and 5000. This choice of sample size is probably optimistic for
126 ME/CFS research, given the sample sizes (mean = 86.2, range = 16–300 for ME/CFS group) reported in
127 immune marker studies conducted in Europe (26).

128 To parameterise the simulation study, we first specified the association between the candidate causal
129 factor and true ME/CFS patients by the odds ratio (hereafter denoted as Δ_T) and the probability of the
130 presence of the causal candidate factor in healthy controls (θ_0). We considered $\Delta_T \in \{1.25, 1.5, 2, 3, 5, 10\}$,
131 whose values ranged from weak to strong associations. Given the lack of disease-specific biomarkers, it is
132 expected that the putative associations to be detected should not be too strong. However, the simulation
133 of strong associations is important to assess what to expect under an optimistic scenario. In addition, we
134 specified $\theta_0 \in \{0.05, 0.1, 0.25, 0.5\}$. In the case of candidate gene scenarios, θ_0 could represent the minor
135 allele frequency of a given single nucleotide variant in the healthy population. Combining the values of θ_0
136 together with the ones specified for Δ_T , we could study the impact of misdiagnosis on detecting whether
137 ME/CFS follows the “common disease, rare variant hypothesis” or the “common disease, common variant
138 hypothesis”. For these two contrasting hypotheses, a common disease could be the result of either multiple
139 rare genetic factors with a high penetrance or many common genetic variants with low penetrance (27, 28).
140 Having θ_0 and Δ_T fixed in the respective values, we can determine the value of θ_1^* as follows

$$\theta_1^* = \frac{\theta_0 \Delta_T}{1 + \theta_0 (\Delta_T - 1)}. \quad (3)$$

141 With respect to the probability of misdiagnosis, γ , we allowed this parameter to vary from 0 (all cases were
142 correctly diagnosed) to 1 (all cases were incorrectly diagnosed) with a lag of 0.01.

143 To simulate data from the second misdiagnosis scenario, we considered fixed parameters $\Delta_T = 3$ and
144 $\theta_0 = 0.25$. For parameters π_{se} and π_{sp} , we considered all possible combinations of 0.80, 0.90, 0.925, 0.975,
145 and 1.0, where $\pi_{se} = \pi_{sp} = 1$ corresponded to the first scenario.

146 For each misdiagnosis scenario, parameter set, and sample size, we simulated 10,000 datasets to estimate
147 the power of rejecting H_0 when the possibility of misdiagnosis was ignored. A detailed description of
148 the simulation procedure can be found elsewhere (10). In each data set, we rejected H_0 if the p-value of
149 the Pearson’s χ^2 test was less than the usual level of significance, $\alpha = 0.05$. The power of the study was
150 estimated as the overall proportion of simulated data sets in which H_0 was rejected. To better understand
151 the practical implications of the simulation results, we specified a target power of at least 80% for the
152 probability of rejecting the null hypothesis in favour of the alternative hypothesis, given that the latter was
153 true. The term used for this target power is generally denoted as $1 - \beta$, where β represents the probability
154 of incorrectly failing to reject H_0 (Type II error). A target value of 80% suggests that a true association is
155 very likely to be reproducible in a follow-up study.

156 2.3 Application to two ME/CFS studies

157 We also studied the impact of misdiagnosis on available data from a candidate gene association study and
158 an immunological evaluation study. The first study recruited 201 self-reported healthy controls and 305

159 ME/CFS patients whose symptoms complied with CCC (20). All participants were genotyped for 5 SNPs
160 previously associated with autoimmune diseases:

- 161 • rs2476601, tyrosine phosphatase non-receptor type 22 (*PTPN22*);
- 162 • rs3087243, cytotoxic T-lymphocyte-associated protein 4 (*CTLA4*);
- 163 • rs1800629 and rs1799724, tumor necrosis factor (*TNF*);
- 164 • rs3807306, interferon regulatory factor 5 (*IRF5*).

165 The study found significant associations of rs2476601 (*PTPN22*) and rs3087243 (*CTLA4*) with ME/CFS
166 patients potentially triggered by an acute infection.

167 The second study encompassed data of 251 patients and 107 healthy controls from the UK ME/CFS
168 biobank (21). Patients complied with either Fukuda/CDC or CCC. The original objective of this study
169 was to perform a detailed comparison between these two groups in terms of immune-cell data. Here,
170 we focused on data of antibody positivity related to six different herpesviruses: human cytomegalovirus
171 (CMV), Epstein-Barr virus (EBV), herpes simplex virus 1 and 2 (HSV1 and HSV2), varicella-zoster virus
172 (VZV), and human herpesvirus (HHV6). Note that, for each virus, antibody positivity was determined by a
173 cut-off value as recommended by the manufacturers of the serological kits. However, another re-analysis of
174 the same data provided evidence for an overlap between the antibody distribution of hypothetical antibody-
175 negative and antibody-positive populations (24). Therefore, besides the potential ME/CFS misdiagnosis,
176 the subsequent antibody classification could also be affected by misclassification.

177 In both studies, we estimated the power of detecting an association as a function of misdiagnosis
178 probability, γ , using simulated data generated from the reported associations, as explained later.

3 RESULTS

179 3.1 Simulation study: impact of ME/CFS misdiagnosis

180 The power to detect an association with ME/CFS decreased with the misdiagnosis probability (Figure 1
181 and Figure 2). The maximum power was achieved when all the patients were assumed to be true cases
182 ($\gamma = 0$), that is, when there was no dilution effect of the true association to be detected. In the other
183 extreme, when all the patients were assumed to be apparent cases ($\gamma = 1$), the corresponding power was
184 equal to the significance level set up for the analysis ($\alpha = 0.05$). This result was a direct consequence of
185 assumption III, in which the misdiagnosed cases were deemed identical to healthy controls in what the
186 association with the candidate causal factor was concerned.

187 As expected, the most optimistic scenarios were associated with $\Delta_T = 5$ or 10 (i.e., strong associations
188 between the candidate causal factor and true ME/CFS cases). In these scenarios, it was possible to find
189 a maximum misdiagnosis probability in which the target power of 80% was achieved irrespective of
190 the sample size and the probability of the causal factor being present in healthy controls (Table 3). For
191 $\Delta_T = 10$, a misdiagnosis probability of 0.53 was sufficient to ensure the desired power for sample sizes
192 greater of equal to 100 individuals per study group ($n_i \geq 100$), irrespective of θ_0 . This minimum probability
193 reduced to 0.24 for $\Delta_T = 5$. That is, when the association between the candidate causal factor and true
194 ME/CFS cases was strong, the chance of missing its diagnosis could only occur if the chance of ME/CFS
195 misdiagnosis was simultaneously high. Therefore, if the “common disease, rare variant” hypothesis holds
196 true in ME/CFS, it is very likely to be detected by genome-wide association studies.

197 Similar optimistic scenarios were observed for sample sizes of 2500 and 5000 individuals per study
198 group with the exception of the case of lowest $\Delta_T = 1.25$. Combining these large sample sizes with strong
199 associations between the candidate causal factor and the true ME/CFS cases, the target power could only
200 not be achieved when almost all the cases were misdiagnosed (with misdiagnosis probability greater than
201 or equal to 0.88).

202 The most pessimist scenarios were associated with either $\Delta_T = 1.25, 1.5$ or $n_0 = n_1 = 100$. When
203 $\Delta_T = 1.25$, the sample size had to increase to 2500 or 5000 individuals per group in order to achieve the
204 desired target power. Therefore, for this weak association, the likelihood of finding reproducible results
205 was very low, even under the assumption of a perfect diagnosis. As a consequence, testing the “common
206 disease, common variant hypothesis” in ME/CFS is likely to fail in future genetic associations. Finally,
207 the case of $n_0 = n_1 = 100$ was particularly problematic given that it was not possible to find any value
208 misdiagnosis probability in which the desired power could be achieved for $\Delta_T \leq 2$ (Figure 1).

209 3.2 Simulation study: impact of ME/CFS misdiagnosis and misclassification on the 210 candidate causal factor

211 We then simulated data of a hypothetical association study in which there were both imperfect diagnoses
212 and misclassification of the candidate causal factor (Figure 2). As explained earlier in this paper, this
213 scenario is the curse of any serological association study in ME/CFS, given the estimation of seropositivity
214 of each individual could be affected by the sensitivity and specificity associated with the classification rule
215 used. At this point, it was clear that for values of $\Delta_T = 1.25, 1.5$, and 2, the desired power was not often
216 achieved for sample sizes smaller than 500 individuals per group in the case of perfect classification of the
217 causal factor. Therefore, the additional assumption of imperfect classification of the candidate causal factor
218 would make previously estimated power even worse. Because of that, we only performed our simulation
219 study on the more optimistic scenario in which $\Delta_T = 3$ (Table 4).

220 3.3 Application to data from two ME/CFS studies

221 We illustrated the problem of misdiagnosis in data from two ME/CFS studies (20, 21). We started with
222 data from a candidate gene association study (20). In this study, some genetic associations were only found
223 significant when comparing healthy controls to ME/CFS patients with an infectious disease trigger onset
224 (Table 5). The estimated allele-related odds ratios varied from 0.84 [95%CI = (0.56, 1.27)] (rs1799724,
225 *TNF2*) to 1.63 [95%CI = (1.04, 2.55)] (rs2476601, *PTPN22*). In our re-analysis, we investigated the impact
226 of misdiagnosis if a replication study were conducted in the same population. In line with the original
227 study, no genotyping errors were assumed for the genetic data. The reported odds ratios were assumed to
228 be the true ones for the population and data were simulated with the same number of reported alleles as
229 in the original study. A description of the simulation procedure can be found elsewhere (10). Similarly
230 to the previous sections, we estimated the probability of rejecting H_0 as a function of the misdiagnosis
231 probability.

232 Again, the estimated probability of rejecting H_0 decreased with the misdiagnosis probability (Figure 3A).
233 More importantly, when the chance of misdiagnosis was low ($\gamma < 0.09$), it was possible to obtain the
234 target power of 80% for the allele association reported for rs3087243 in *CTLA4*. Therefore, the target
235 power cannot be ensured for $\gamma > 0.09$. For the remaining SNPs, the target power was never achieved
236 irrespective of the misdiagnosis probability. This is particularly problematic for rs2476601 in *PTPN22*
237 whose association was reported to be significant at a 0.05 significance level. For this SNP, a chance of

Table 3. Maximum values of misdiagnosis probability γ that still ensured the power of at least 80% to reject the null hypothesis of lack of association at different values of true odds ratio, Δ_T , risk factor probability on healthy controls and false positive cases, θ_0 , and sample sizes, n_i , $i = (0, 1)$. Cells with no value indicate that the desired power could not be reached using the respective parameter combination, even when considering a perfect diagnostic scenario ($\gamma = 0$).

$\Delta_T \backslash \theta_0$	0.05	0.1	0.25	0.5	n_i
10	0.59	0.65	0.64	0.53	100
5	0.24	0.43	0.50	0.42	
3	—	0.02	0.25	0.23	
2	—	—	—	—	
1.5	—	—	—	—	
1.25	—	—	—	—	
10	0.77	0.79	0.77	0.70	250
5	0.56	0.66	0.69	0.63	
3	0.20	0.41	0.53	0.50	
2	—	—	0.23	0.26	
1.5	—	—	—	—	
1.25	—	—	—	—	
10	0.84	0.86	0.84	0.78	500
5	0.70	0.76	0.78	0.73	
3	0.47	0.60	0.67	0.65	
2	—	0.27	0.46	0.47	
1.5	—	—	0.04	0.13	
1.25	—	—	—	—	
10	0.89	0.90	0.89	0.84	1000
5	0.80	0.84	0.85	0.81	
3	0.64	0.72	0.77	0.75	
2	0.32	0.50	0.62	0.62	
1.5	—	0.05	0.32	0.38	
1.25	—	—	—	—	
10	0.93	0.94	0.93	0.90	2500
5	0.88	0.90	0.90	0.88	
3	0.78	0.83	0.85	0.84	
2	0.58	0.69	0.76	0.76	
1.5	0.18	0.42	0.58	0.59	
1.25	—	—	0.20	0.28	
10	0.95	0.95	0.95	0.93	5000
5	0.91	0.93	0.93	0.91	
3	0.84	0.88	0.90	0.88	
2	0.71	0.78	0.83	0.83	
1.5	0.44	0.59	0.70	0.72	
1.25	—	0.20	0.44	0.49	

238 approximately 10% misdiagnosis deemed a power of about 50%, conceptually making the replication of
 239 this result equivalent to flipping a coin.

240 The second study referred to putative associations of 6 herpes virus infections with ME/CFS using
 241 serological data (21). In these data, all individuals were classified as seronegative or seropositive for each
 242 antibody used. Under the assumption of perfect serological classification and diagnosis, the associations of
 243 these serological data with severely affected ME/CFS patients ranged from 0.65 [95%CI = (0.21, 1.97)] to

Table 4. Maximum values of misdiagnosis probability γ that still ensures a power of rejecting the null hypotheses of at least 80% for $\Delta_T = 3$ and $\theta_0 = 0.25$, where π_{se} and π_{sp} represent sensitivity and specificity associated with the classification of the candidate, respectively. See Table 3 for more information.

$\pi_{sp} \backslash \pi_{se}$	1	0.975	0.925	0.9	0.8	n
1	0.25	0.23	0.20	0.19	0.11	100
0.975	0.22	0.20	0.17	0.15	0.06	
0.925	0.17	0.14	0.09	0.08	–	
0.9	0.13	0.11	0.07	0.04	–	
0.8	0.03	–	–	–	–	
1	0.53	0.52	0.51	0.50	0.45	250
0.975	0.51	0.50	0.48	0.47	0.42	
0.925	0.47	0.46	0.43	0.42	0.36	
0.9	0.45	0.43	0.41	0.39	0.32	
0.8	0.38	0.36	0.31	0.29	0.18	
1	0.67	0.67	0.66	0.65	0.62	500
0.975	0.66	0.65	0.64	0.63	0.59	
0.925	0.63	0.62	0.60	0.59	0.55	
0.9	0.61	0.61	0.59	0.57	0.52	
0.8	0.56	0.54	0.51	0.50	0.42	
1	0.77	0.77	0.76	0.75	0.73	1000
0.975	0.76	0.75	0.74	0.74	0.72	
0.925	0.74	0.73	0.72	0.71	0.68	
0.9	0.73	0.72	0.71	0.70	0.67	
0.8	0.68	0.67	0.65	0.64	0.59	
1	0.85	0.85	0.85	0.84	0.83	2500
0.975	0.85	0.85	0.84	0.84	0.82	
0.925	0.84	0.83	0.82	0.82	0.80	
0.9	0.83	0.82	0.81	0.81	0.79	
0.8	0.80	0.79	0.78	0.78	0.74	
1	0.90	0.90	0.89	0.89	0.88	5000
0.975	0.89	0.89	0.89	0.88	0.87	
0.925	0.88	0.88	0.87	0.87	0.86	
0.9	0.88	0.87	0.87	0.87	0.85	
0.8	0.86	0.85	0.84	0.84	0.81	

244 1.60 [95%CI = (0.83, 3.09)] for Epstein-Barr virus (EBV) and Herpes simplex virus 1 (HSV1), respectively
 245 (Table 6). For this case, none of the associations was deemed significant at the usual significance level of
 246 5% according to the original study (p-values ≥ 0.16).

247 The original serological classification was based on a cut-off in the antibody levels determined by the
 248 2σ -rule. In general, this cut-off is defined by the mean plus twice the standard deviation of a known
 249 seronegative population. Under the assumption of a normal distribution for the seronegative population, the
 250 expected specificity of the serological specificity is approximately 0.975 (24). We assumed this value for
 251 π_{sp} . For simplicity, we assumed $\pi_{se} = \pi_{sp}$. Again, we simulated data from this scenario as the original
 252 study and estimated the probability of rejecting H_0 as a function of the misdiagnosis probability.

253 In this study, it was not possible to reach the desired power of 80% to reject H_0 for any of the antibodies
 254 (Figure 3B). The best case was observed for the data of antibodies against HSV1. In this case, the maximum
 255 power was approximately 0.50 in the absence of misdiagnosis. This power dropped to 0.30 when $\gamma = 0.25$.

Table 5. Reported associations of a candidate gene association study (20) where $\hat{\theta}_0$ represents the frequencies of the non-reference allele for healthy controls and $\hat{\Delta}_T$ is the odds ratio of these allele frequencies when comparing ME/CFS patients with an infectious disease trigger to healthy controls. P-values are associated with the Pearson's χ^2 test for 2×2 contingency tables.

SNP	Gene	$\hat{\theta}_0$	$\hat{\Delta}_T$	CI ($\hat{\Delta}_T$)	p-value
rs3087243	<i>CTLA4</i>	0.56	1.54	(1.17, 2.03)	0.002
rs2476601	<i>PTPN22</i>	0.08	1.63	(1.04, 2.55)	0.033
rs1799724	<i>TNF₂</i>	0.13	0.84	(0.56, 1.27)	0.409
rs1800629	<i>TNF₁</i>	0.16	0.89	(0.61, 1.30)	0.551
rs3807306	<i>IRF5</i>	0.51	0.94	(0.72, 1.22)	0.637

Table 6. Summary of serological findings from (21), where $\hat{\theta}_0$ represents the seroprevalence of healthy controls and $\hat{\Delta}_T$ refers to the odds ratio for being seropositive when comparing severely-affected ME/CFS patients with healthy controls. The 95% CI ($\hat{\Delta}_T$) P-values are associated with the Pearson's χ^2 test for 2×2 contingency tables.

Herpes virus	$\hat{\theta}_0$	$\hat{\Delta}_T$	95% CI ($\hat{\Delta}_T$)	p-value
Herpes simplex virus 1	0.42	1.60	(0.83, 3.09)	0.163
Herpes simplex virus 2	0.34	1.36	(0.69, 2.66)	0.377
Epstein-Barr virus	0.93	0.65	(0.21, 1.97)	0.442
Cytomegalovirus	0.37	0.84	(0.42, 1.67)	0.613
Varicella-Zoster virus	0.97	0.75	(0.12, 4.63)	0.757
Human Herpesvirus-6	0.95	1.27	(0.24, 6.79)	0.776

256 For the remaining cases, the power was almost exclusively less than 0.20. This could partially be explained
 257 by the fact that θ_0 is higher than 0.93 for antibody data related to EBV, HHV6, and VZV. Therefore, the
 258 chance of finding a significant association is meagre if the study were to be repeated in the same population
 259 and under the same conditions.

4 DISCUSSION

260 This study showed the impact of disease misdiagnosis on the reproducibility of associations from a
 261 given ME/CFS study. This impact encompasses non-trivial combinations of sample size, effect size (or
 262 strength of association), and misdiagnosis probability, in general, plus specificity/sensitivity in the case of
 263 serological association studies. In this regard, our simulation study showed that strong associations with
 264 ME/CFS can be detected with reasonable power even when the probability of misdiagnosis is not negligible.
 265 However, strong associations are unlikely to be detected in ME/CFS, given the persistent difficulty in
 266 finding reproducible biomarkers (26) and the lack of consistently significant results reported by different
 267 genetic association studies of the disease (29, 30, 31, 32). In addition, even if these strong associations
 268 could be detected, the respective estimates should be taken as an underestimation of the true ones due to
 269 the presence of apparent cases. In other words, the assumption of misdiagnosis implies that the association
 270 estimates are intrinsically biased and, therefore, should be corrected, as done in the estimation of disease
 271 prevalence in the presence of imperfect diagnostic tests (15). In general, the chance of misdiagnosis (or
 272 misclassification) brings a situation of overparameterization of the respective statistical analysis, where
 273 the misdiagnosis probability cannot be estimated from the data directly. To deal with this situation, there

274 are non-trivial solutions in the statistical literature (33, 34). However, the discussion about these advanced
275 solutions was out of the scope of this study.

276 It was also clear that sufficiently large sample sizes can compensate for the reduction in power due to
277 misdiagnosis and misclassification of a potential causal factor. In this regard, a minimum sample size of
278 500 individuals per group minimises the negative impact of misdiagnosis alone on power. This minimum
279 sample size required an increase towards 1000 individuals per group when the misdiagnosis probability is
280 combined with the misclassification probabilities associated with the potential causal factor. In general,
281 the sampling of such large numbers of individuals is challenging, but is becoming feasible in common,
282 well-recognised, and highly-funded diseases, such as cancer, cardiovascular diseases (35), and autoimmune
283 disorders (36, 37). However, for ME/CFS, where the research is affected by limited funding and poor
284 societal recognition (38), the recruitment of large numbers of patients might be problematic. A possible
285 solution to this problem is to use a collection of samples available from existing biobanks. In this regard,
286 the UK ME/CFS Biobank is fully dedicated to the advancement and acceleration of ME/CFS research. It
287 comprises biological samples of more than 500 individuals, including ME/CFS patients, healthy controls,
288 and patients with multiple sclerosis as an additional control group (39). The strength of this biobank is
289 a detailed characterisation of all the participants and the use of current standards for ME/CFS diagnosis.
290 However, this biobank only contemplates around 250 ME/CFS patients. As an alternative, the UK Biobank
291 has been reported to include 1890 to 2105 ME/CFS cases (30, 32). However, it is unclear the accuracy
292 of the respective data given that ME/CFS diagnosis was self-reported by participants. In this scenario,
293 the most obvious solution to increase the sample size is to conduct multi-centric studies, as suggested
294 by the European ME/CFS Biomarker Landscape project (26). Such a research strategy has already been
295 adopted by a recent genetic association study that contemplated Norwegian, Danish, and British cohorts at
296 different stages of the analysis (32). Notwithstanding their attractiveness and potential impact on biomarker
297 discovery, multi-centric studies require considerable funding, robust research protocols and ethics, and
298 strong compliance among the participating sites. In this scenario, the major challenge of these studies
299 consists in the fact that ME/CFS research is currently underfunded (40, 41). However, there is hope
300 emerging from initiatives such as EUROMENE and the US Clinician Coalition, whose objective is to
301 promote and foster collaboration and shared resources among their members. In addition, the similarity
302 between long-Covid and ME/CFS highlighted the need to increase research funding on these conditions
303 to mitigate their economic costs in the future (42). Therefore, large multi-centric studies in ME/CFS are
304 expected to become a routine research practice in the future.

305 The potential misclassification of a binary causal factor is often neglected in the analysis of both genetic
306 and serological association studies of ME/CFS. In previous studies, this neglect does not seem problematic
307 because estimated genotype error rates are below 1% (43). However, higher genotype error rates can be
308 obtained for rare genetic markers (44), and genetic markers with minor allele frequencies less than 0.01 or
309 0.05 are typically excluded from genetic association studies of ME/CFS (22, 32, 45). The parameterization
310 of our simulation study (i.e., $\theta_0 > 0.05$) is then in line with this research practice and, thus, the respective
311 findings are directly applicable for understanding the reproducibility of current genetic association studies
312 of ME/CFS. Within this scenario, Dibble et al. (31) advocated for conducting well-powered studies with
313 large cohort sizes. In the present context, such a suggestion is not only important to detect common
314 variants with low genetic effects, but also to minimise the impact of misdiagnosis on the underlying
315 power function, as demonstrated by the present study. In the case of serological association studies, the
316 serological status of each individual can only be determined statistically. Generally, when the antibody
317 distributions of the underlying seronegative or seropositive populations do not overlap with each other, the
318 serological status of each individual can be determined almost perfectly. In practice, serological data of

319 herpesviruses from the UK ME/CFS Biobank showed that the distributions of these serological populations
320 can overlap with each other, as demonstrated by data related to herpesviruses serology from the UK
321 ME/CFS Biobank (19). However, current serological association studies do not report the sensitivity
322 and specificity of the corresponding serological classification rule and, therefore, it is unclear how these
323 misclassification parameters vary from one study to another and how they affect the power of detecting
324 underlying associations. The situation is problematic given that these studies are also affected by the high
325 seroprevalence of many viruses in the population, the possibility of batch and laboratory effects, differences
326 in the antigens used, interactions between serological outcomes, and typical confounders, such as gender
327 and age, and the use of different criteria for seropositivity estimation (19, 46, 47). Therefore, it is imperative
328 the reporting of the uncertainty associated with the serological classification rule in future studies.

329 In this study, we assumed independence between disease association and possible confounding factors.
330 In practice, this assumption might not always be true, given the evidence that age, gender, or exposure to
331 a given infectious agent can influence the respective findings (19, 48). Hence, the above assumption is
332 more appropriate for interpreting studies in which individuals are sampled from a specific combination of
333 confounding factors. Such a sampling strategy is particularly adopted in DNA methylation studies in which
334 adult women with matched age and BMI are the target population (49).

335 Furthermore, we assumed no misdiagnosis of healthy controls. This assumption deserves a comment in
336 light of current data and research practices. On the one hand, ME/CFS patients and some healthy controls
337 share the same symptoms and level of fatigue (10, 50). On the other hand, studies such as Loebel et al.
338 (17) and Szklarski et al. (48) recruited healthy individuals solely based on self-reporting. Such recruiting
339 strategy, although reducing research costs, increases the chance of sampling asymptomatic but unhealthy
340 individuals. A similar problem can be expected in Kaushik (51), Johnston et al. (52), and Lande et al.
341 (53), using control samples from blood donors. In Herrera et al. (29), Blauensteiner et al. (54), de Vega
342 et al. (55), and Trivedi et al. (56), both asymptomatic non-obese and obese individuals were recruited for
343 subsequent matching with ME/CFS patients. However, some of these individuals can be seen as cases of
344 undiagnosed metabolic syndrome (57, 58, 59). Recruiting these individuals is particularly problematic
345 given the association of metabolic syndrome with ME/CFS (60). With all of the evidence, healthy controls
346 could also be divided into true and apparent controls, thus further reducing the statistical power to detect a
347 putative disease association. Therefore, the most general assumption is to consider the possibility in which
348 undiagnosed patients of ME/CFS or any other disease can be mistakenly recruited as healthy controls.
349 To avoid this situation, the health status of putative controls should also be routinely confirmed with lab
350 testing.

351 We formulated the statistical problem in terms of ME/CFS misdiagnosis, considering the suspected
352 ME/CFS cohort to be composed of misdiagnosed and true patients. From a modelling standpoint, our
353 formulation can be reframed as a related problem in which the ME/CFS group can be divided into two
354 subgroups of patients, and the association to be detected is only present in one of these subgroups. This
355 alternative formulation follows the idea that ME/CFS encompasses multiple disease subtypes (61). However,
356 we should be cautious when interpreting our results in this alternative scenario. First, the assumption of
357 only two disease subgroups seems overly simplistic, given that it was suggested more than seven genomic
358 subtypes for the disease (62, 63). Second, these potential subtypes might overlap with each other, given that
359 no clear stratification of patients could be identified using symptoms' data (10). Hence, the membership of
360 each patient to a given disease subtype might be probabilistic in nature. However, this situation is more
361 complex and was not captured by the presented mathematical formulation.

362 In summary, this study discusses the reproducibility of ME/CFS studies in terms of disease misdiagnosis.
363 Given the absence of an objective biomarker for disease diagnosis, the most obvious recommendations
364 are: (i) to use recommended case definitions for ME/CFS diagnosis; (ii) to increase cohort sizes using
365 multi-centric studies; and (iii) to confirm the health status of putative controls. These recommendations
366 combined are likely to minimise the impact of misdiagnosis of both patients and controls on the disease
367 associations to be detected. A final recommendation is to perform post-hoc power calculations as a function
368 of misdiagnosis probability in order to quantify the likelihood of a reproducible finding in the presence of
369 imperfect diagnosis of ME/CFS.

CONFLICT OF INTEREST STATEMENT

370 The authors declare that the research was conducted in the absence of any commercial or financial
371 relationships that could be construed as a potential conflict of interest.

AUTHOR CONTRIBUTIONS

372 JM and NS conceived and designed the analysis. JM performed the computer simulations, statistical
373 analysis, and wrote the initial manuscripts. All authors participated actively in the discussion of the results.
374 All authors have read, revised, and approved the final version of the manuscript.

FUNDING

375 JM acknowledges funding from Fundação para a Ciência e Tecnologia, Portugal (grant ref.
376 SFRH/BD/149758/2019). NS acknowledges funding from the Fundação para a Ciência e Tecnologia,
377 Portugal (ref. UIDB/00006/2020), and the Polish National Agency for Academic Exchange, Poland (ref.
378 PPN/ULM/2020/1/00069/U/00001).

ACKNOWLEDGMENTS

379 JM would like to thank Prof. Przemysław Biecek and the MI² DataLab group at Faculty of Mathematics
380 and Information Science, Warsaw University of Technology, Warsaw, Poland, for their support during part
381 of the analysis and discussion of this article.

DATA AVAILABILITY STATEMENT

382 All simulations and analyses were done using R statistical software, version 4.1.0 (64).
383 The source code and datasets generated can be found in the GitHub public repository
384 (<https://github.com/jtmalato/misclassification-simulations>).

REFERENCES

- 385 1. Fukuda K, Strausm SE, Hickie I, Sharpe MC, Dobbins JG, Komaroff A. The chronic fatigue syndrome:
386 A comprehensive approach to its definition and study. *Annals of Internal Medicine* **121** (1994) 953.
387 doi:10.7326/0003-4819-121-12-199412150-00009.
- 388 2. Carruthers BM, Jain AK, Meirleir KLD, Peterson DL, Klimas NG, Lerner AM, et al. Myalgic
389 encephalomyelitis/chronic fatigue syndrome. *Journal of Chronic Fatigue Syndrome* **11** (2003) 7–115.
390 doi:10.1300/j092v11n01_02.

- 391 3. Lim EJ, Ahn YC, Jang ES, Lee SW, Lee SH, Son CG. Systematic review and meta-analysis of the
392 prevalence of chronic fatigue syndrome/myalgic encephalomyelitis (CFS/ME). *Journal of Translational*
393 *Medicine* **18** (2020). doi:10.1186/s12967-020-02269-0.
- 394 4. Estévez-López F, Mudie K, Wang-Steverding X, Bakken IJ, Ivanovs A, Castro-Marrero J, et al.
395 Systematic review of the epidemiological burden of myalgic encephalomyelitis/chronic fatigue
396 syndrome across Europe: Current evidence and EUROMENE research recommendations for
397 epidemiology. *Journal of Clinical Medicine* **9** (2020) 1557. doi:10.3390/jcm9051557.
- 398 5. Smith MEB, Nelson HD, Haney E, Pappas M, Daeges M, Wasson N, et al. Diagnosis and treatment of
399 myalgic encephalomyelitis/chronic fatigue syndrome. Tech. rep. (2014). doi:10.23970/ahrqepcerta219.
- 400 6. Brurberg KG, Fønhus MS, Larun L, Flottorp S, Malterud K. Case definitions for chronic fatigue
401 syndrome/myalgic encephalomyelitis (CFS/ME): a systematic review. *BMJ Open* **4** (2014) e003973.
402 doi:10.1136/bmjopen-2013-003973.
- 403 7. Pheby DF, Araja D, Berkis U, Brenna E, Cullinan J, de Korwin JD, et al. The development of a
404 consistent Europe-wide approach to investigating the economic impact of myalgic encephalomyelitis
405 (ME/CFS): A report from the European network on ME/CFS (EUROMENE). *Healthcare* **8** (2020) 88.
406 doi:10.3390/healthcare8020088.
- 407 8. Institute of Medicine. *Beyond Myalgic Encephalomyelitis/Chronic Fatigue Syndrome: Redefining an*
408 *Illness* (Washington, DC: The National Academies Press) (2015). doi:10.17226/19012.
- 409 9. Nacul L, Authier FJ, Scheibenbogen C, Lorusso L, Helland IB, Martin JA, et al. European network
410 on myalgic encephalomyelitis/chronic fatigue syndrome (EUROMENE): Expert consensus on the
411 diagnosis, service provision, and care of people with ME/CFS in europe. *Medicina* **57** (2021) 510.
412 doi:10.3390/medicina57050510.
- 413 10. Malato J, Graça L, Nacul L, Lacerda E, Sepúlveda N. Statistical challenges of investigating a disease
414 with a complex diagnosis. *medRxiv* (2021). doi:10.1101/2021.03.19.21253905.
- 415 11. Morris G, Maes M. Myalgic encephalomyelitis/chronic fatigue syndrome and encephalomyelitis
416 disseminata/multiple sclerosis show remarkable levels of similarity in phenomenology and
417 neuroimmune characteristics. *BMC Medicine* **11** (2013). doi:10.1186/1741-7015-11-205.
- 418 12. Gaber TAZK, Oo WW, Ringrose H. Multiple sclerosis/chronic fatigue syndrome overlap: When two
419 common disorders collide. *NeuroRehabilitation* **35** (2014) 529–534. doi:10.3233/NRE-141146.
- 420 13. Ali S, Matcham F, Irving K, Chalder T. Fatigue and psychosocial variables in autoimmune rheumatic
421 disease and chronic fatigue syndrome: A cross-sectional comparison. *Journal of Psychosomatic*
422 *Research* **92** (2017) 1–8. doi:https://doi.org/10.1016/j.jpsychores.2016.11.002.
- 423 14. Greenland S. Basic methods for sensitivity analysis of biases. *International Journal of Epidemiology*
424 **25** (1996) 1107–1116. doi:https://doi.org/10.1093/ije/25.6.1107-a.
- 425 15. Rogan WJ, Gladen B. Estimating prevalence from the results of a screening test. *American Journal of*
426 *Epidemiology* **107** (1978) 71–76. doi:10.1093/oxfordjournals.aje.a112510.
- 427 16. Nacul L, Lacerda EM, Kingdon CC, Curran H, Bowman EW. How have selection bias and disease
428 misclassification undermined the validity of myalgic encephalomyelitis/chronic fatigue syndrome
429 studies? *Journal of Health Psychology* **24** (2017) 1765–1769. doi:10.1177/1359105317695803.
- 430 17. Loebel M, Eckey M, Sotzny F, Hahn E, Bauer S, Grabowski P, et al. Serological profiling of the EBV
431 immune response in chronic fatigue syndrome using a peptide microarray. *PLOS ONE* **12** (2017)
432 e0179124. doi:10.1371/journal.pone.0179124.
- 433 18. Shikova E, Reshkova V, ntoniya Kumanova, Raleva S, Alexandrova D, Capo N, et al.
434 Cytomegalovirus, epstein-barr virus, and human herpesvirus-6 infections in patients with myalgic

- 435 encephalomyelitis/chronic fatigue syndrome. *Journal of Medical Virology* (2020). doi:10.1002/jmv.
436 25744.
- 437 **19.** Domingues TD, Grabowska AD, Lee JS, Ameijeiras-Alonso J, Westermeier F, Scheibenbogen C,
438 et al. Herpesviruses serology distinguishes different subgroups of patients from the United Kingdom
439 Myalgic Encephalomyelitis/Chronic Fatigue Syndrome Biobank. *Frontiers in Medicine* **8** (2021).
440 doi:10.3389/fmed.2021.686736.
- 441 **20.** Steiner S, Becker SC, Hartwig J, Sotzny F, Lorenz S, Bauer S, et al. Autoimmunity-related risk variants
442 in PTPN22 and CTLA4 are associated with ME/CFS with infectious onset. *Frontiers in Immunology*
443 **11** (2020). doi:10.3389/fimmu.2020.00578.
- 444 **21.** Cliff JM, King EC, Lee JS, Sepúlveda N, Wolf AS, Kingdon C, et al. Cellular immune function in
445 myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS). *Frontiers in Immunology* **10** (2019).
446 doi:10.3389/fimmu.2019.00796.
- 447 **22.** Grabowska AD, Lacerda EM, Nacul L, Sepúlveda N. Review of the quality control checks performed
448 by current genome-wide and targeted-genome association studies on myalgic encephalomyelitis/chronic
449 fatigue syndrome. *Frontiers in Pediatrics* **8** (2020). doi:10.3389/fped.2020.00293.
- 450 **23.** Ruiz-Pablos M, Paiva B, Montero-Mateo R, Garcia N, Zabaleta A. Epstein-Barr virus and the origin
451 of myalgic encephalomyelitis or chronic fatigue syndrome. *Frontiers in Immunology* **12** (2021).
452 doi:10.3389/fimmu.2021.656797.
- 453 **24.** Domingues TD, Mouriño H, Sepúlveda N. Analysis of antibody data using finite mixture models based
454 on scale mixtures of skew-normal distributions. *medRxiv* (2021). doi:10.1101/2021.03.08.21252807.
- 455 **25.** Sepúlveda N, Stresman G, White MT, Drakeley CJ. Current mathematical models for analyzing
456 anti-malarial antibody data with an eye to malaria elimination and eradication. *Journal of Immunology*
457 *Research* **2015** (2015) 1–21. doi:10.1155/2015/738030.
- 458 **26.** Scheibenbogen C, Freitag H, Blanco J, Capelli E, Lacerda E, Authier J, et al. The European ME/CFS
459 Biomarker Landscape project: an initiative of the European network EUROMENE. *Journal of*
460 *translational medicine* **15** (2017) 162. doi:https://doi.org/10.1186/s12967-017-1263-z.
- 461 **27.** Schork NJ, Murray SS, Frazer KA, Topol EJ. Common vs. rare allele hypotheses for complex diseases.
462 *Current Opinion in Genetics & Development* **19** (2009) 212–219. doi:10.1016/j.gde.2009.04.010.
- 463 **28.** Gibson G. Rare and common variants: twenty arguments. *Nature Reviews Genetics* **13** (2012) 135–145.
464 doi:10.1038/nrg3118.
- 465 **29.** Herrera S, de Vega WC, Ashbrook D, Vernon SD, McGowan PO. Genome-epigenome interactions
466 associated with myalgic encephalomyelitis/chronic fatigue syndrome. *Epigenetics* **13** (2018) 1174–1190.
467 doi:10.1080/15592294.2018.1549769.
- 468 **30.** Tanigawa Y, Li J, Justesen JM, Horn H, Aguirre M, DeBoever C, et al. Components of genetic
469 associations across 2,138 phenotypes in the UK Biobank highlight adipocyte biology. *Nature*
470 *Communications* **10** (2019). doi:10.1038/s41467-019-11953-9.
- 471 **31.** Dibble JJ, McGrath SJ, Ponting CP. Genetic risk factors of ME/CFS: a critical review. *Human*
472 *Molecular Genetics* **29** (2020) R117–R124. doi:10.1093/hmg/ddaa169.
- 473 **32.** Hajdarevic R, Lande A, Mehlsen J, Rydland A, Sosa DD, Strand EB, et al. Genetic association study
474 in myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS) identifies several potential risk loci.
475 *Brain, Behavior, and Immunity* **102** (2022) 362–369. doi:10.1016/j.bbi.2022.03.010.
- 476 **33.** Paulino CD, Soares P, Neuhaus J. Binomial regression with misclassification. *Biometrics* **59** (2003)
477 670–675. doi:10.1111/1541-0420.00077.

- 478 **34.** Luo S, Chan W, Detry MA, Massman PJ, Doody RS. Binomial regression with a misclassified
479 covariate and outcome. *Statistical Methods in Medical Research* **25** (2012) 101–117. doi:10.1177/
480 0962280212441965.
- 481 **35.** Giri A, Hellwege JN, Keaton JM, Park J, Qiu C, Warren HR, et al. Trans-ethnic association study
482 of blood pressure determinants in over 750,000 individuals. *Nature Genetics* **51** (2018) 51–62.
483 doi:10.1038/s41588-018-0303-9.
- 484 **36.** International Multiple Sclerosis Genetics Consortium (IMSGC). Analysis of immune-related loci
485 identifies 48 new susceptibility variants for multiple sclerosis. *Nature Genetics* **45** (2013) 1353–1360.
486 doi:10.1038/ng.2770.
- 487 **37.** Bjornevik K, Cortese M, Healy BC, Kuhle J, Mina MJ, Leng Y, et al. Longitudinal analysis reveals
488 high prevalence of Epstein-Barr virus associated with multiple sclerosis. *Science* **375** (2022) 296–301.
489 doi:10.1126/science.abj8222.
- 490 **38.** Pheby DFH, Araja D, Berkis U, Brenna E, Cullinan J, de Korwin JD, et al. A literature review of GP
491 knowledge and understanding of ME/CFS: A report from the socioeconomic working group of the
492 European network on ME/CFS (EUROMENE). *Medicina* **57** (2020) 7. doi:10.3390/medicina57010007.
- 493 **39.** Lacerda EM, Mudie K, Kingdon CC, Butterworth JD, O'Boyle S, Nacul L. The UK ME/CFS Biobank:
494 A disease-specific biobank for advancing clinical research into myalgic encephalomyelitis/chronic
495 fatigue syndrome. *Frontiers in Neurology* **9** (2018). doi:10.3389/fneur.2018.01026.
- 496 **40.** Dimmock ME, Mirin AA, Jason LA. Estimating the disease burden of ME/CFS in the United States
497 and its relation to research funding. *Journal of Medicine and Therapeutics* **1** (2016). doi:10.15761/jmt.
498 1000102.
- 499 **41.** Mirin AA, Dimmock ME, Jason LA. Research update: The relation between ME/CFS disease burden
500 and research funding in the USA. *Work* **66** (2020) 277–282. doi:10.3233/WOR-203173.
- 501 **42.** Mirin AA, Dimmock ME, Jason LA. Updated ME/CFS prevalence estimates reflecting post-COVID
502 increases and associated economic costs and funding implications. *Fatigue: Biomedicine, Health &*
503 *Behavior* (2022) 1–11. doi:10.1080/21641846.2022.2062169.
- 504 **43.** Smith AK, Fang H, Whistler T, Unger ER, Rajeevan MS. Convergent genomic studies identify
505 association of GRIK2 and NPAS2 with chronic fatigue syndrome. *Neuropsychobiology* **64** (2011)
506 183–194. doi:10.1159/000326692.
- 507 **44.** Hunter-Zinck H, Shi Y, Li M, Gorman BR, Ji SG, Sun N, et al. Genotyping array design and data
508 quality control in the Million Veteran Program. *The American Journal of Human Genetics* **106** (2020)
509 535–548. doi:10.1016/j.ajhg.2020.03.004.
- 510 **45.** Hajdarevic R, Lande A, Rekeland I, Rydland A, Strand EB, Sosa DD, et al. Fine mapping of the major
511 histocompatibility complex (MHC) in myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS)
512 suggests involvement of both HLA class I and class II loci. *Brain, Behavior, and Immunity* **98** (2021)
513 101–109. doi:10.1016/j.bbi.2021.08.219.
- 514 **46.** Ariza ME. Commentary: Antibodies to human herpesviruses in myalgic encephalomyelitis/chronic
515 fatigue syndrome patients. *Frontiers in Immunology* **11** (2020). doi:10.3389/fimmu.2020.01400.
- 516 **47.** Sepúlveda N, Malato JT, Sotzny F, Grabowska AD, Fonseca A, Cordeiro C, et al. Revisiting IgG
517 antibody reactivity to Epstein-Barr virus in myalgic encephalomyelitis/chronic fatigue syndrome and
518 its potential application to disease diagnosis. *medRxiv* (2022). doi:10.1101/2022.04.20.22273990.
- 519 **48.** Szklarski M, Freitag H, Lorenz S, Becker SC, Sotzny F, Bauer S, et al. Delineating the association
520 between soluble CD26 and autoantibodies against G-protein coupled receptors, immunological and
521 cardiovascular parameters identifies distinct patterns in post-infectious vs. non-infection-triggered

- 522 myalgic encephalomyelitis/chronic fatigue syndrome. *Frontiers in Immunology* **12** (2021). doi:10.
523 3389/fimmu.2021.644548.
- 524 **49.** Malato J, Sotzny F, Bauer S, Freitag H, Fonseca A, Grabowska A, et al. The SARS-CoV-2 receptor
525 angiotensin-converting enzyme 2 (ACE2) in myalgic encephalomyelitis/chronic fatigue syndrome:
526 A meta-analysis of public DNA methylation and gene expression data. *Heliyon* **7** (2021) e07665.
527 doi:10.1016/j.heliyon.2021.e07665.
- 528 **50.** Cella M, Chalder T. Measuring fatigue in clinical and community settings. *Journal of Psychosomatic*
529 *Research* **69** (2010) 17–22. doi:10.1016/j.jpsychores.2009.10.007.
- 530 **51.** Kaushik N. Gene expression in peripheral blood mononuclear cells from patients with chronic fatigue
531 syndrome. *Journal of Clinical Pathology* **58** (2005) 826–832. doi:10.1136/jcp.2005.025718.
- 532 **52.** Johnston S, Staines D, Klein A, Marshall-Gradisnik S. A targeted genome association study examining
533 transient receptor potential ion channels, acetylcholine receptors, and adrenergic receptors in chronic
534 fatigue syndrome/myalgic encephalomyelitis. *BMC Medical Genetics* **17** (2016). doi:10.1186/
535 s12881-016-0342-y.
- 536 **53.** Lande A, Fluge Ø, Strand EB, Flåm ST, Sosa DD, Mella O, et al. Human leukocyte antigen alleles
537 associated with myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS). *Scientific Reports* **10**
538 (2020). doi:10.1038/s41598-020-62157-x.
- 539 **54.** Blauensteiner J, Bertinat R, León LE, Riederer M, Sepúlveda N, Westermeier F. Altered endothelial
540 dysfunction-related miRs in plasma from ME/CFS patients. *Scientific Reports* **11** (2021). doi:10.1038/
541 s41598-021-89834-9.
- 542 **55.** de Vega WC, Vernon SD, McGowan PO. DNA methylation modifications associated with chronic
543 fatigue syndrome. *PLoS ONE* **9** (2014) e104757. doi:10.1371/journal.pone.0104757.
- 544 **56.** Trivedi MS, Oltra E, Sarria L, Rose N, Beljanski V, Fletcher MA, et al. Identification of myalgic
545 encephalomyelitis/chronic fatigue syndrome-associated DNA methylation patterns. *PLOS ONE* **13**
546 (2018) e0201066. doi:10.1371/journal.pone.0201066.
- 547 **57.** Bo S, Ciccone G, Pearce N, Merletti F, Gentile L, Cassader M, et al. Prevalence of undiagnosed
548 metabolic syndrome in a population of adult asymptomatic subjects. *Diabetes Research and Clinical*
549 *Practice* **75** (2007) 362–365. doi:10.1016/j.diabres.2006.06.031.
- 550 **58.** Cabré JJ, Martín F, Costa B, Piñol JL, Llor JL, Ortega Y, et al. Metabolic syndrome as a cardiovascular
551 disease risk factor: Patients evaluated in primary care. *BMC Public Health* **8** (2008). doi:10.1186/
552 1471-2458-8-251.
- 553 **59.** Eftekhazadeh A, Asghari G, Serahati S, Hosseinpanah F, Azizi A, Barzin M, et al. Predictors of
554 incident obesity phenotype in nonobese healthy adults. *European Journal of Clinical Investigation* **47**
555 (2017) 357–365. doi:10.1111/eci.12743.
- 556 **60.** Maloney EM, Boneva RS, Lin JMS, Reeves WC. Chronic fatigue syndrome is associated with
557 metabolic syndrome: results from a case-control study in Georgia. *Metabolism* **59** (2010) 1351–1357.
558 doi:10.1016/j.metabol.2009.12.019.
- 559 **61.** Jason LA, Corradi K, Torres-Harding S, Taylor RR, King C. Chronic fatigue syndrome: The need for
560 subtypes. *Neuropsychology Review* **15** (2005) 29–58. doi:10.1007/s11065-005-3588-2.
- 561 **62.** Kerr JR, Petty R, Burke B, Gough J, Fear D, Sinclair LI, et al. Gene expression subtypes in patients
562 with chronic fatigue syndrome/myalgic encephalomyelitis. *The Journal of Infectious Diseases* **197**
563 (2008) 1171–1184. doi:10.1086/533453.
- 564 **63.** Zhang L, Gough J, Christmas D, Matthey DL, Richards SCM, Main J, et al. Microbial infections in
565 eight genomic subtypes of chronic fatigue syndrome/myalgic encephalomyelitis. *Journal of Clinical*
566 *Pathology* **63** (2009) 156–164. doi:10.1136/jcp.2009.072561.

567 **64.** R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical
568 Computing, Vienna, Austria (2020).

FIGURE CAPTIONS

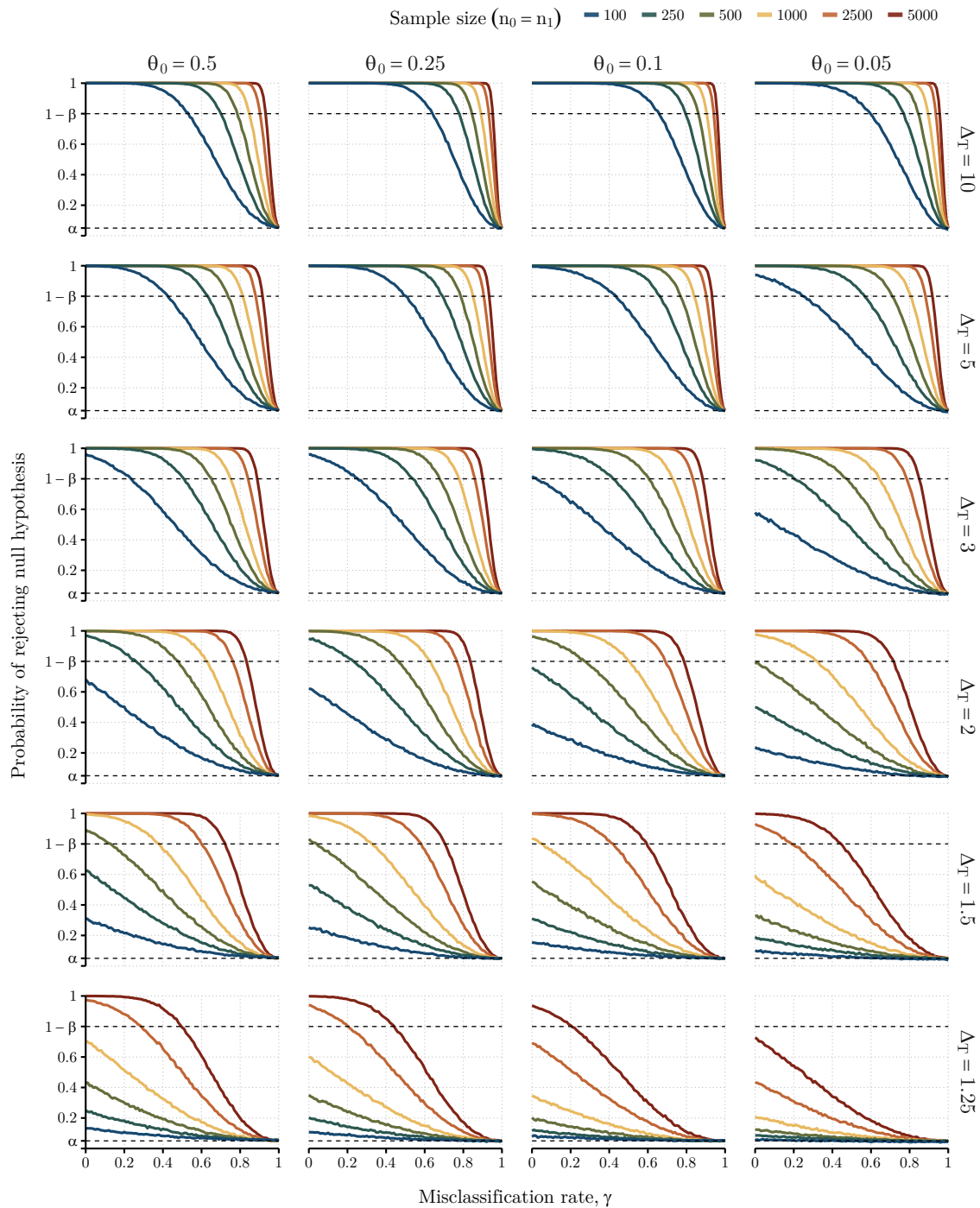


Figure 1. Probabilities of rejecting the null hypothesis, i.e., absence of association between the two populations, as function of the misclassification rate. Each column represents the values attributed to the risk allele frequency found in matched healthy controls and false positive ME/CFS cases ($\theta_0 \in \{0.05, 0.1, 0.25, 0.5\}$). Each row varies the true odds ratio for the association between risk allele frequency assessed between true positive cases and healthy controls ($\Delta_T \in \{1.25, 1.5, 2, 3, 5, 10\}$). Power analysis was estimated for different sample sizes of 100, 250, 500, 1000, 2500, and 5000 ($n_0 = n_1$), represented by the different coloured lines on each scenario. Upper dashed line indicates the target power, where the probability of rejecting the null hypothesis is $1 - \beta = 0.80$. Lower dashed line indicates the significance level used, $\alpha = 0.05$.

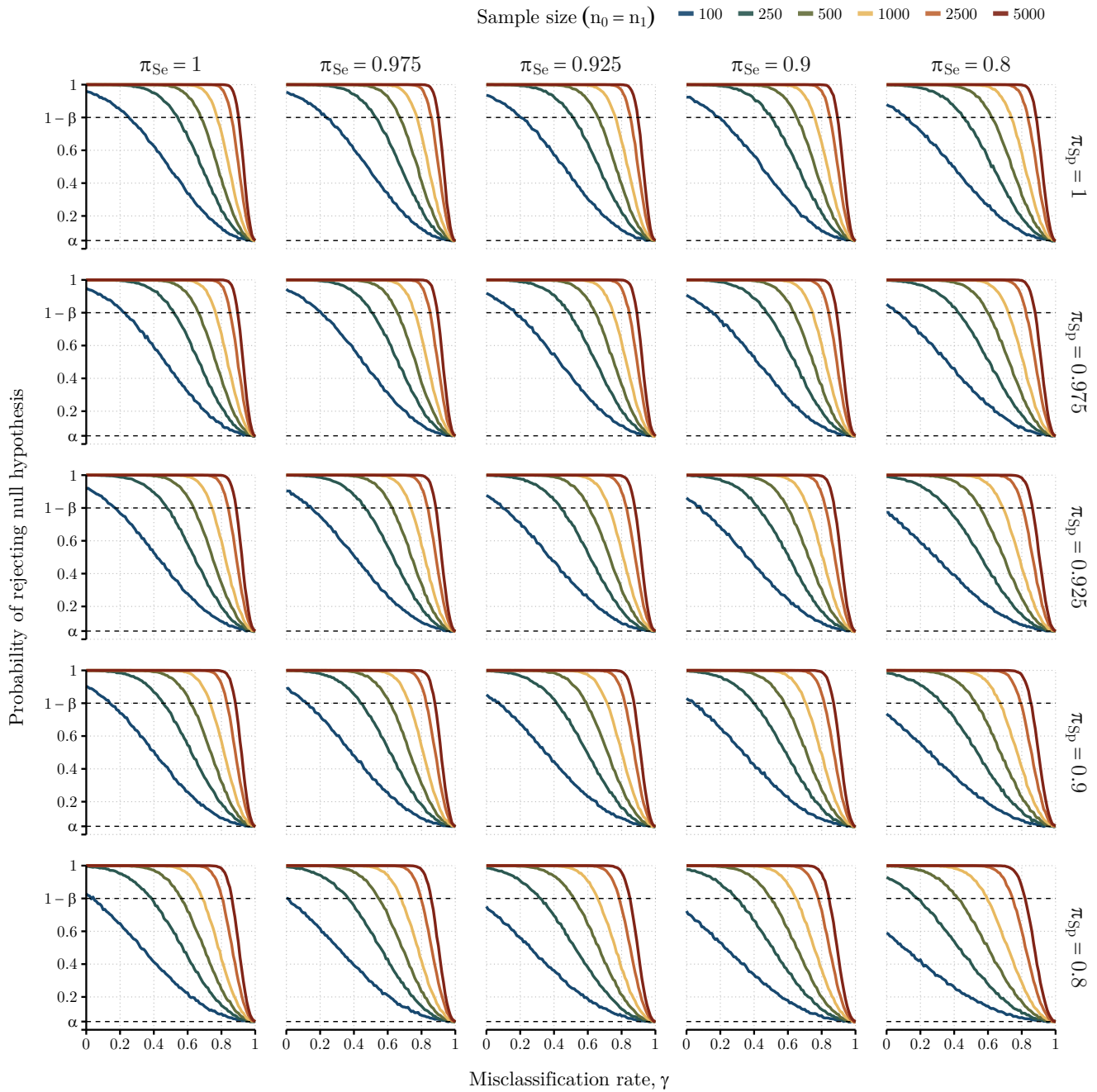


Figure 2. Probabilities of rejecting H_0 as function of the misclassification rate. Each scenario represents simulated results with combination of serology test's sensitivity, π_{se} , and specificity, π_{sp} , for columns and rows, respectively. Power analysis was estimated for different sample sizes of 100, 250, 500, 1000, 2500, and 5000 ($n_0 = n_1$), represented by the different coloured lines on each scenario, with probability of exposure in healthy controls fixed as $\theta_0 = 0.25$ and true odds ratio $\Delta_T = 3$. Upper dashed line indicates the target power, where the probability of rejecting the null hypothesis is $1 - \beta = 0.80$. Lower dashed line indicates the significance level used, $\alpha = 0.05$.

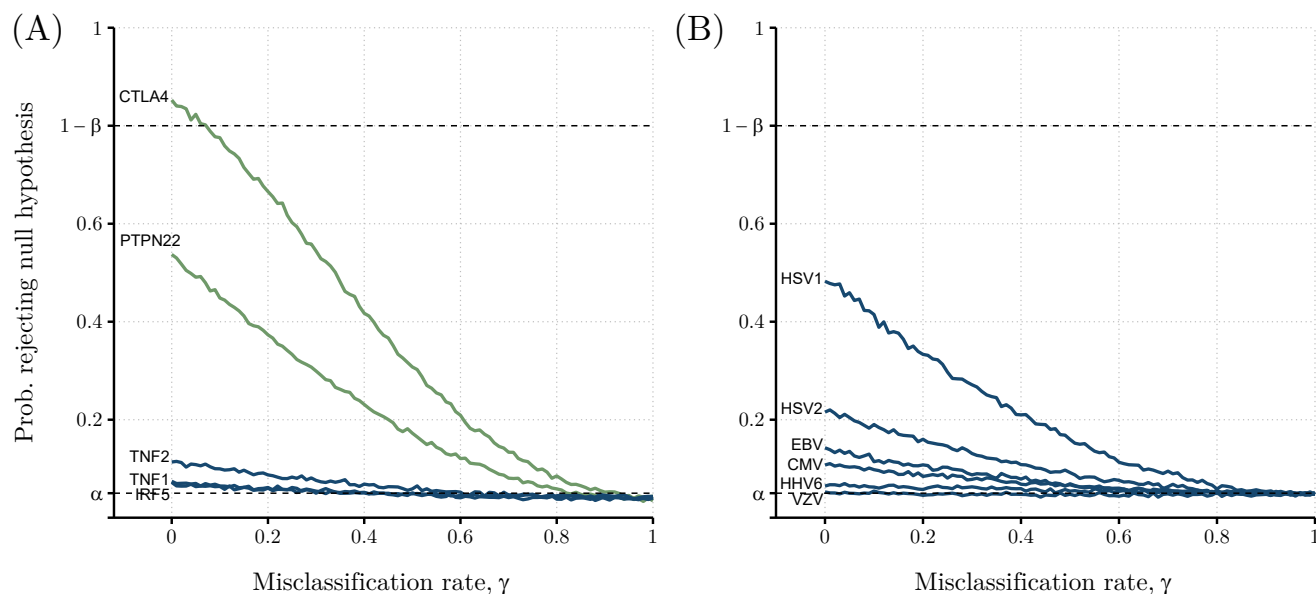


Figure 3. Probabilities of rejecting the H_0 considering real-world data from (A) five different SNPs studied (genes *PTPN22*, *CTLA4*, *TNF* (1 and 2), and *IRF5*); and (B) six different human herpes viruses (CMV, EBV, HSV1 and HSV2, VZV, and HHV6), as function of misclassification rate. For each study, risk allele frequencies or probability of exposure and true odds ratio were determined by Steiner et al. (20) ($n_0 = 201$; $n_1 = 305$) and Cliff et al. (21) ($n_0 = 107$; $n_1 = 251$; $\pi_{se} = \pi_{sp} = 0.975$), with determined values shown in Tables 5 and 6, respectively. Green coloured lines indicate candidate risk factors where a significant association with the disease was found in the original study. Blue coloured lines show non-significant ME/CFS risk factors. Upper dashed line indicates the target power, where the probability of rejecting the null hypothesis is $1 - \beta = 0.80$. Lower dashed line indicates the significance level used, $\alpha = 0.05$.