



**HAL**  
open science

# DISEASES DIAGNOSIS AND CLASSIFICATION AS PRIME NUMBER REPRESENTATIONS

Pengfei Zhang

► **To cite this version:**

Pengfei Zhang. DISEASES DIAGNOSIS AND CLASSIFICATION AS PRIME NUMBER REPRESENTATIONS. 2022. hal-03723482

**HAL Id: hal-03723482**

**<https://hal.science/hal-03723482v1>**

Preprint submitted on 14 Jul 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1           **DISEASES DIAGNOSIS AND CLASSIFICATION AS PRIME**  
2                           **NUMBER REPRESENTATIONS\***

3   PENGFEI ZHANG<sup>†</sup>

4       **Abstract.** Classification and diagnosis of diseases by criteria are prevalent in clinical medicine.  
5 We demonstrate that prime number representation of these classifications can be used for automated  
6 diagnosis generations as well as analyses of structural questions in disease classifications.

7       **Key words.** Disease Classifications, Applied Number Theory

8       **MSC codes.** 11Z05, 00A69

9       **1. Introduction.** Criteria based classification of diseases is becoming the norm  
10 in medicine. The classic example is the Diagnostic and Statistical Manual of Men-  
11 tal Disorders (DSM), a criteria based guideline used in psychiatry for diagnostic  
12 purposes.[1] Similarly, the International Classification of Headache Disorders (ICHD),  
13 was conceived by neurologists in 1988, providing clinicians with an authoritative cod-  
14 ified diagnostic paradigm for headache disorders. [5] In both basic research as well  
15 as clinical practice, both Diagnostic and Statistical Manual of Mental Disorders, 5th  
16 edition (DSM5) and the International Classification of Headache Disorders 3rd edition  
17 (ICHD3) are widely used as the gold standard approach in diagnosis. Diagnosis by  
18 criteria therefore fundamentally influence both fields.

19 These are not the only examples, other criteria such as International Classifica-  
20 tion of Vestibular Disorders and International classification of Orofacial Pain seek to  
21 comprehensively classified diseases in a domain. Criteria also exists for fibromyalgia,  
22 restlessless syndrome, and lupus, to name a few. [4] [6]

23 The benefits of codifying diagnosis is clear; diagnostic criteria allow clinicians to be  
24 precise in diagnosis and therefore treatments and research of diseases. However, as  
25 in any classification, problems of internal structural consistencies rise to the fore-  
26 front when classifications are applied widely. This is evident in scholarly debates: In  
27 headache, for example, whether post traumatic headaches should be classified by its  
28 phenotype or by the onset and duration after trauma is a hotly debated topic.[2] The  
29 reason, is that phenotypically, chronic post traumatic headache, by the ICHD3 defi-  
30 nition, can be made to conform to a diagnosis of chronic migraine or chronic tension  
31 type headache. [5] In other words, post traumatic headaches suffer from a classifica-  
32 tion problem on "uniqueness": Does our current classification system for headaches  
33 offers unique diseases with a unique classification?

34 On the other end of the spectrum, classifications often struggle with the "missing  
35 diagnosis" problem. Consider, for example, that the first version of International  
36 classification of headache disorders contains only 13 categories of diagnoses; its most  
37 recent embodiment, the ICHD3, contains 14 categories of disorders with an appendix.  
38 Therefore we are diagnosing diseases today not diagnosable in ICHD's first version.  
39 A similar phenomenon occurs in psychiatry; in the original DSM, a number of disor-  
40 ders we now thought of as canonical were not included. These are what is considered  
41 missing diagnosis questions: What disease entity are we missing in our current clas-  
42 sifications? In other words, is our classification "complete"?

---

\*Submitted to the editors June 29th, 2020.

**Funding:** This work was not funded.

<sup>†</sup>Rutgers Robert Wood Johnson Hospital, New Brunswick, NJ (pz124@rwjms.rutgers.edu).

43 Of course, "missing" diagnosis in a criteria is inevitable; it is a sign of scientific ad-  
 44 vancements. Yet it is troublesome that there are missing diagnosis which seems to  
 45 arise purely as downstream effects of "holes" in the Classification system. An ex-  
 46 ample of a well known "gap" in headaches occurs in a category of headaches called  
 47 the "trigeminal autonomic cephalalgias" (TAC): There is a diagnostic "gap" between  
 48 the longest of the TACs - hemicrania continua - and the second longest - "cluster  
 49 headaches"; the former being constant, the latter lasting maximum of only three  
 50 hours. TAC which are longer than 3 hours but not constant certainly exists but are  
 51 often times unclassifiable under the current schema. [7] Therefore aside from the lim-  
 52 itations of scientific advancement, what are these headache disorders with recognized  
 53 phenotypes and yet is undiagnosable in the current diagnostic criteria?

54 In this paper we construct a mathematical tool to answer these classification ques-  
 55 tions. Specifically, we first propose that prime number factorization can be used to  
 56 encode disease classifications as long as the latter can be represented by propositional  
 57 logic. This method allows for automated diagnosis of diseases. From this ability we  
 58 can then seek to provide a foundation for answering these structural questions of com-  
 59 pleteness and uniqueness that arises from disease classification.

60 This paper is organized in the following fashion: Since the technique is novel, we will  
 61 demonstrate this technique by applying it to International Classification of Headache  
 62 Disorders (ICHD3). We will then supply the reader with the formal justification  
 63 mathematically. Finally, in the last two sections we will sketch out how to apply our  
 64 methods to uniqueness and completeness problems.

65

66 **2. Demonstration of Method by Application to ICHD3.** ICHD3 is a  
 67 widely used and canonical way of diagnosing headache disorders and follows a criteria  
 68 system; headache disorders are classified by groupings of phenotypes. For example,  
 69 the criteria for migraine without aura are presented below, and the disorder is defined  
 70 by the existence of a specific set of headache phenotypes.

71

72 DEFINITION 2.1. (from [5]) *Migraine without aura diagnostic criteria:*

73

74 A. *At least five attacks fulfilling criteria B–D*

75 B. *Headache attacks lasting 4–72 hours (when untreated or unsuccessfully treated)*

76 C. *Headache has at least two of the following four characteristics:*

77 1. *unilateral location*

78 2. *pulsating quality*

79 3. *moderate or severe pain intensity*

80 4. *aggravation by or causing avoidance of routine physical activity (e.g. walking  
 81 or climbing stairs)*

82 D. *During headache at least one of the following:*

83 1. *nausea and/or vomiting*

84 2. *photophobia and phonophobia*

85 E. *Not better accounted for by another ICHD-3d diagnosis*

86

87 One can translate this criteria directly into propositional logic statements as follows:  
 88 (We use the alphanumeric designation of the criteria as short-hand for each pheno-  
 89 type.)

90 DEFINITION 2.2 (propositional translation of ICHD3). *Migraine without aura is*

91 *defined as the following:*  $A \wedge B \wedge [(C1 \wedge C2) \vee (C1 \wedge C3) \vee (C1 \wedge C4) \vee (C2 \wedge C3) \vee$   
 92  $(C2 \wedge C4) \vee (C3 \wedge C4)] \wedge (D1 \vee D2) \wedge E$

This definition can be expanded, of course, into its disjunctive normal form:

$$(A \wedge B \wedge C1 \wedge C2 \wedge D1 \wedge E) \vee (A \wedge B \wedge C1 \wedge C2 \wedge D2 \wedge E) \vee (A \wedge B \wedge C1 \wedge C3 \wedge D1 \wedge E) \dots$$

93 In clinical practice, given a set of headache phenotype, such as a patient profile,  
 94 headache disorders can be diagnosed based on satisfactions of the above criteria. This  
 95 is therefore equivalent to diagnosis based on propositional logic statements.

96 Of course, migraine without aura is not the only possible headache diagnoses in  
 97 ICHD3. Indeed more than 200 headache disorders exist. Each of these headache  
 98 disorders are also defined by a diagnostic criteria similar to the one presented above.  
 99 We will refer the reader to the actual ICHD3 for more detail. [5]

100 **3. Translations of propositional logic criteria into encoding.** Since every  
 101 ICHD3 criterion can be translated to propositional logic statements in its disjunctive  
 102 normal form, one can translate the classification into prime number encoding using  
 103 the following algorithm:

104

105 Step 1: All headache phenotypes in a diagnosis paradigm are assigned a unique  
 106 prime number. (Appendix 1) Notice that negations crucial to diagnostic criteria are  
 107 also assigned a prime number.

108 Step 2: If AND is used between two phenotypes in the criteria, then the  
 109 corresponding prime number for those two phenotypes are multiplied together. If  
 110 OR is used between two phenotypes, then the prime numbers for the corresponding  
 111 headache phenotypes are put in a list.

112 Step 3: Compositions of Step 2 operations are then used throughout the whole  
 113 encoding procedure.

114 For the purpose of demonstration, one example of the encoding migraine without aura  
 115 is shown below.[8]

116 [4075291, 9333731, 9596653, 26304151, 27045113, 61942033, 28934661, 297496243,  
 117 450154441, 681362363, 1030998881, 1060041103, 1920203023, 2905542301,  
 118 2987388563, 6842083483, 21122233253, 31960965311, 32861274193, 75262918313,  
 119 212104587979, 2333150467703]

120 A patient profile, thought of as a collection of headache phenotypes, can be ex-  
 121 pressed using only logic conjunction. Therefore, a patient profile can be expressed  
 122 as one composite integer. For example, a patient who has the phenotype of five  
 123 headaches, each lasting 4 to 72 hours, unilateral, pulsating, with nausea, photopho-  
 124 bia and no other diagnosis can be represented by  $19*17*71*73*37*67 = 4150116211$ .  
 125 (We do not encode criteria E as it is in every single ICHD3 diagnoses. However, this  
 126 modification can be make and the algorithm would still remain intact.)

127

128 We will show below that a patient profile, represented as a composite number, di-  
 129 vides at least one number in its corresponding diagnosis' encoding. In other words,  
 130 a patient who has migraine without aura must have an encoding which divide by at  
 131 least one number in the migraine without aura encoding. This observation forms the  
 132 basis for automated diagnosis.

133

134 For example, 61942033 divides 4150116211 in example above, therefore migraine with-  
 135 out aura is diagnostic. Since the above process need not limit itself to headache clas-  
 136 sification, this result can be applied to any disease classification following a similar

137 criteria paradigm.

138 Finally, we must note two peculiarities to our formulation: 1) the encodings are square  
 139 free since one cannot count a criteria twice; 2) although theoretically possible, in prac-  
 140 tice logical conflicts cannot exist in a patient nor a criteria encoding. (For example,  
 141 it makes no sense to be both photophobic and not photophobic at the same time in  
 142 either the criteria or user encoding.) Both of these peculiarities will be exploited in  
 143 our analysis of "uniqueness" and "completeness" below.

144 **4. Mathematical Justification of the Technique.** We will now construct a  
 145 formal definition of "encoding" for a given collection of diagnostic criteria. We will  
 146 first construct the following abstraction:

DEFINITION 4.1. *Let  $S$  be the following sequence:*

$$S = \{S_1, S_2, S_3 \dots S_i \dots S_{m-1}, S_m\}$$

147 where  $S_i$  represents individual statements (i.e. phenotype) in a classification.

148 We now define a function that "encodes" the propositional values for each indi-  
 149 vidual statements based on whether the users answers *True* or *False*.

DEFINITION 4.2. *Define the following:*

$$\varphi(S_i) = \begin{cases} P_i^1 & \text{if } S_i = \text{True} \\ P_i^0 & \text{if } S_i = \text{False} \end{cases}$$

150 where  $P_i$  is the  $i^{\text{th}}$  prime. Here "True" or "False" assigned by the user for each  $S_i$   
 151 statement.

DEFINITION 4.3. *The "encoding" for a set of user assigned phenotype is the fol-  
 lowing:*

$$Q(S) = \prod_i^m \varphi(S_i)$$

152 Notice that the "space" in which these encodings exist is simply the square free inte-  
 153 gers up to the  $m^{\text{th}}$  prime. We will call this  $I_m$ , representing all possible values of a  
 154 user's questionnaire:

155

DEFINITION 4.4.

$$I_m = \{i | i \in P_1^{n_1} P_2^{n_2} P_3^{n_3} \dots P_m^{n_m}, n_m \leq 1, m \in \mathbb{Z}, p \in \mathbb{P}\}$$

156 We will first prove that prime factorization and propositional values in the user's  
 157 answers are connected for logic conjunction:

158 **THEOREM 4.5.** *Given  $a \in I_m$ , if  $P_i$  divides  $a$ , then  $S_i = \text{True}$ .*

*Proof.* By fundamental theorem of arithmetic:

$$a = P_1^{n_1} P_2^{n_2} P_3^{n_3} \dots P_m^{n_m}$$

159 Since  $P_i$  divides  $a$ , then  $n_i \geq 1$ . Now  $I_m$  is square free and  $a \in I_m$ , therefore  $n_i = 1$ .  
 160 Suppose to the contrary that  $S_i = \text{False}$ , then  $\varphi(S_i) = P_i^0$ . This implies that  $n_i = 0$ ,  
 161 reaching a contradiction.  $\square$

162 Now we will prove the following relationship between two encoding:

163 THEOREM 4.6. Let  $\iota, \pi \in I_m$ , where  $\pi = P'_{a_1} P'_{a_2} \dots P'_{a_j}$  and the user assigns  
 164  $S_1, S_2 \dots S_j$  such that  $\iota$  is the encoding by definition 4.3. If  $\pi$  divides  $\iota$ , then  $S_{a_1} =$   
 165  $True, S_{a_2} = True \dots S_{a_j} = True$

166 (We use the subscript  $a_1, a_2, \dots a_j$  to differentiate arbitrary collections of prime from  
 167 the  $j^{th}$  prime.)

168 *Proof.* Given that  $\pi$  divides  $\iota$ , then  $P'_{a_1}$  divides  $\iota$ ,  $P'_{a_2}$  divides  $\iota$ ,  $P'_{a_3}$  divides  $\iota \dots P'_{a_j}$   
 169 divides  $\iota$

170 By Theorem 4.5,  $S_{a_1} = True, S_{a_2} = True \dots S_{a_j} = True$   $\square$

171 We now need to define the notion of diagnostic criteria and what it means to  
 172 diagnose a disorder.

173 DEFINITION 4.7. A diagnostic statement,  $\sigma$ , for a specific disorder is the assign-  
 174 ments of *True* to a subsequence of  $S$  and the assignments of *False* otherwise.

175 DEFINITION 4.8. A diagnostic criteria for a specific disorder is a collection of  
 176 diagnostic statements for that disorder.

177 For example, a migraine diagnostic statement is assigning each of  $S_1, S_2, S_3, S_4, S_5$   
 178 to *True* if  $S_1$  is "at least five headaches",  $S_2$  is "each lasting 4 to 72 hours",  $S_3$   
 179 is "unilateral",  $S_4$  is "pulsating", and  $S_5$  is "nausea". This diagnostic statement  
 180 describes one of a number of conditions that satisfies migraine without aura according  
 181 to ICHD3. A diagnostic criteria is the collection of all such statements which satisfies  
 182 migraine without aura. We are able to define a diagnostic criteria as a collection  
 183 of diagnostic statements due to the fact that all propositional statements can be  
 184 translated into a disjunctive normal form.

185 DEFINITION 4.9. We say that  $i'$ , an assignment of *True* and *False* to elements  
 186 of  $S$ , is considered a "diagnosis" of a diagnostic criteria,  $M' = [\sigma_1, \sigma_2, \dots \sigma_j \dots, \sigma_k]$ , if  
 187  $\sigma_j^T$  is a subsequence of  $i'^T$  for some  $j$ , where  $\sigma_j^T$  the subsequent of  $\sigma_j$  that is assigned  
 188 as *True* and  $i'^T$  is the subsequent of  $i'$  that is assigned as *True*.

189 This should be intuitive as any assignment that conforms to a diagnostic criteria  
 190 is the diagnosis. The  $i'$  here is really just the representation for a set of phenotype,  
 191 such as a patient profile. (This set needs to be ordered in order to be assigned a  
 192 prime number by  $\varphi$ , therefore sequence is used and not set in our definitions.) The  
 193 definition simply suggests that if there is a diagnostic statement,  $\sigma_j$ , which matches  
 194 a subsequence of *True* as  $i'$ , then it is diagnostic. Notice that *False* is not taken in  
 195 to consideration here, specifically due to the fact that negations are assigned a prime  
 196 number and encoded directly.

197  
 198 We can pin down these definitions mathematically by directly considering encod-  
 199 ings:

DEFINITION 4.10. A diagnosis set,  $M$ , for a diagnostic criteria

$$M' = [\sigma_1, \sigma_2, \dots \sigma_k]$$

is defined as the following:

$$M = [Q_1(S), Q_2(S), Q_3(S) \dots Q_n(S)]$$

200 where  $\sigma_1$  assigns *True* or *False* to elements of  $S$  in  $Q_1(S)$  in accordance to Definition  
 201 4.3. The same applies to  $\sigma_2$  to  $Q_2(S)$  ... etc.

202 This is the mathematical encoding for a diagnostic criteria where every element in the  
 203 set  $M$  is a composite number encoding a diagnostic statement. A concrete example  
 204 of this is shown in section 3.

205

206 Now we find prove the crux of our argument:

207 **THEOREM 4.11.** *Let  $M$  be a diagnostic set for diagnostic criteria  $M'$ ,  $i$  an encod-*  
 208 *ing for  $i'$ , and let  $\pi \in M$  where  $\pi$  encodes  $\sigma \in M'$ . Then if  $\pi$  divides  $i$ , then  $i'$  is*  
 209 *diagnostic of  $M'$ .*

210 *Proof.* Since  $\pi \in M$ ,  $\pi \in I_m$ . Similarly, since  $i$  is an encoding,  $i \in I_m$  also.  
 211 Let  $\pi = P'_{a1}P'_{a2}\dots P'_{aj}$ . By theorem 4.6, since  $\pi$  divides  $i$ , then  $S_{a1} = True, S_{a2} =$   
 212  $True\dots S_{aj} = True$  for  $i'$ . Furthermore,  $S_{a1} = True, S_{a2} = True\dots S_{aj} = True$  for  
 213  $\sigma$  by theorem 4.6, since  $\pi$  divides itself. Since  $\sigma \in M$ , and the above suggests that  
 214 whichever element of  $S$  is *True* for  $\sigma$  is *True* also for  $i'$ , therefore the former is a  
 215 subsequence of the latter. By definition 4.9  $i'$  is diagnostic of  $M'$ .  $\square$

216 **THEOREM 4.12.** *Let  $i'$  be diagnostic of  $M'$ . Then there exists  $\pi \in M$  where  $\pi$*   
 217 *encodes  $\sigma \in M'$  such that  $\pi$  divides  $i$ .*

218 *Proof.* Let  $i$  be an encoding of  $i'$  and  $i'$  be a diagnostic of  $M'$ . Then there exists  
 219  $\sigma_j$  such that  $\sigma_j^T$  is a subsequence of  $i'^T$  by definition 4.9. Applying definition 4.3 to  
 220 this  $\sigma_j$  yield a  $\pi$  such that  $\pi = Q_j(S)$ . Therefore  $\pi \in M$  by definition 4.10. Now we  
 221 need to show that this  $\pi$  divides  $i$ : Since  $\sigma_j^T$  is a subsequence of  $i'^T$ , so  $Q(\sigma_j)$  divides  
 222  $Q(i')$ .  $\square$

223 The above two theorems therefore justify the following claim:

224 **THEOREM 4.13.** *Let  $M$  be the diagnostic set of  $M'$ ,  $i$  be an encoding of  $i'$ , and*  
 225  *$\pi \in M$  where  $\pi$  encodes  $\sigma$ . Then  $\pi$  divides  $i$  if and only if  $i'$  is diagnostic of  $M'$ .*

226 **5. Sketch of Applications to The Question of "Uniqueness".** The prob-  
 227 lem of dual diagnosis, or codiagnoses, is essentially a problem of uniqueness. In other  
 228 words, the question is whether one disease with a specific phenotype can receive two  
 229 different diagnosis under the same classification schema. In order to tackle the prob-  
 230 lem it is useful to generate all possible combinations of duo diagnosis directly by  
 231 utilizing the encodings above in the following fashion:

232 Let  $i$  be diagnostic of  $M'$  and  $N'$  where  $M' = [\sigma_1, \sigma_2, \sigma_3\dots\sigma_k]$  and  $N' = [\tau_1, \tau_2, \tau_3\dots\tau_l]$ .  
 233 Then by theorem 4.13, there exists both  $\sigma \in M'$  such that  $\sigma$  divides  $i$  and  $\tau \in N'$   
 234 such that  $\tau$  divides  $i$ . Therefore,  $\sigma$  multiplies  $\tau$  divides  $i$ . In other words, the set  
 235  $C = [\sigma\tau | \sigma \in M', \tau \in N', \forall \sigma, \tau]$  yield the condition for all possible codiagnoses.

236 In practice, however, this is an overcounting of duo diagnoses, since we allow both a  
 237 phenotype as well as its negation to occupy unique prime numbers. For example, in  
 238 Appendix 1, "no photophobia" (prime encoding 43) and "photophobia" (prime 67).  
 239 Therefore we simply need to remove from  $C$  all elements that is divisible by the en-  
 240 codings of conflicting phenotypes. (That is,  $43*67$ , in our example above.)

241

242 **6. Sketch of Applications to The Question of "Completeness".** The  
 243 problem of missing diagnosis is essentially a problem of completeness. In other words  
 244 we want to characterize which sets of disease phenotypes cannot have a diagnosis un-  
 245 der a classification schema. One way to answer this question is to screen through all  
 246 possible combinations of disease phenotype and see which one cannot be diagnosed .  
 247 However, even in our examples above , screening through all possible combinations  
 248 would involve screening through  $2^{27}$  combinations. In a large set of criteria, such as

249 ICHD3, the number of combinations becomes astronomical and this method become  
250 impractical.

251 However, for large enough numbers we can estimate the number of "gaps" in the  
252 classification by applications of theorem on square free numbers:

253 Since  $I_m$  is the set of square free integers, the question of "impossible" diagnoses is  
254 equivalent in asking which of square free integers are not divisible by specific sets of  
255 prime. For example, the number of ways in which a headache phenotype which is  
256 "photophobic" has no ICHD3 diagnosis is equivalent to the following heuristic: (the  
257 number of integers in  $I_m$  that is divisible by 67) - (the number of all square free integer  
258 divisible by prime factorization of each ICHD3 dx 1, etc) - (all square free integer  
259 divisible by prime factorization by any of the illogical combinations).

260 All variables in the above heuristics can be estimated either considering the natural  
261 density of square free numbers through a method developed by Brown in 2021. The  
262 main result is reproduced below:[3]

263

*THEOREM 6.1. Given two sets of disjoint primes  $P$  and  $T$ , where  $T$  is finite, Then the porportion of all numbers which are sqaure-free and divisible by all of the primes in  $T$  an dby non of teh primes in  $P$  is:*

$$\frac{6}{\pi^2} \prod_{p \in T} \frac{1}{1+p} \prod_{p \in P} \frac{p}{1+p}$$

264 Our approach therefore offers a theoretical solution for the open question of "impos-  
265 sible" diagnosis.

266 **7. Discussion:** This paper proposes a methods of prime number encoding for  
267 disease diagnosis and classification. We first present a demonstration using the  
268 ICHD3. Following this demonstration, we then introduce the language of prime num-  
269 ber encoding mathematically. Then we applied it to disease diagnosis and proved  
270 that diagnosis can be interpreted as a prime number divisibility problem. Finally, we  
271 show how the encoding can be applied to the uniqueness problem directly. We also  
272 show how the properties of square free integers can be used to tackle the completeness  
273 problem.

274 An important limitation of our construction is the inherent logical inconsistency of  
275 various disease diagnosis paradigms. For example, in ICHD3, The last criteria of  
276 every single diagnostic criteria requires an exclusion of other diagnosis. This is, of  
277 course, not logically consistent . In fact, the ICHD3 itself notes this inconsistency  
278 in the section on diagnosis of New Daily Persistent Headaches; In this section, the  
279 criteria allows for duo diagnosis of multiple different disorders to satisfy this logical  
280 inconsistency due to exclusion. [5] We can, of course, indirectly incorporate the last  
281 criteria of exclusion into our diagnostic prime number paradigm. However, if we do  
282 so, then the burden of deciding whether this exclusion criteria applies will lie with the  
283 user. In other words, no mathematical construction or model can correct for internal  
284 logical inconsistencies of a diagnostic guide.

285 Future directions of this research involves direct application of our theory in real life  
286 settings. Fortunately, unlike other kinds of data, diagnostic criteria are readily avail-  
287 able and free to most clinicians. Therefore prime number encodings of diagnostic  
288 criteria should be an easy undertaking. A corollary to our project, is that prime  
289 number encoding can be used for automated diagnosis of diseases. In a world where  
290 psychiatric and headache access is limited financially in various parts of the world,



291 this may provide hey importantly diagnostic tool for clinicians.  
292 Finally, we must note that the assumptions and indeed the goal of this project is a  
293 Kantian one. Indeed, what we are investigating is not so much true disease patho-  
294 physiology but the condition of possibility of our thinking, as a scientific community,  
295 on headaches and psychiatric disorders. In other words, we are investigating what  
296 Foucault would called the "Order of Things" in medicine whenever criteria based  
297 categorization of disease is applied.

298 **8. Conclusions.** Prime number encoding of disease classification in criteria al-  
299 lows for automated diagnosis . This automated diagnosis allows us to investigate  
300 the uniqueness and completeness of disease classifications. This paper provides the  
301 theoretical language necessary for such an investigation.

302 Appendix 1:

303

304

Statement Number	Description of Criteria
2	1 to 14 episodes per month
3	10 episodes but less than 1 per month
5	15 days per month
7	30 min to 7 days in duration
11	aggravated by physical activity
13	bilateral location
17	duration between 4 to 72 hours
19	greater than 5 episodes
23	hours to days or unremitting
29	mild to moderate pain
31	moderate to severe
37	nausea/vomiting
41	no nausea/vomiting
43	no phonophobia
47	no photophobia
53	not pulsating
59	not aggravated by activity
61	Phonophobic
67	Photophobic
71	pulsating
73	unilateral

305

## REFERENCES

- 306 [1] AMERICAN PSYCHIATRIC ASSOCIATION, *Diagnostic and statistical manual of mental disorders*,  
307 American Psychiatric Association, 5th ed., 2013.
- 308 [2] ASHINA H, EIGENBRODT AK, ET AL. , *Post-traumatic headache attributed to traumatic brain*  
309 *injury: classification, clinical characteristics, and treatment*, *Lancet Neurol*, 20 (2021),  
310 pp. 460–469, [https://doi.org/10.1016/S1474-4422\(21\)00094-6](https://doi.org/10.1016/S1474-4422(21)00094-6).
- 311 [3] R. BROWN, *The natural density of some sets of square-free numbers*, *Integers*, 21 (2021), <http://math.colgate.edu/~integers/v81/v81.pdf>.
- 312 [4] COMMITTEE FOR THE CLASSIFICATION OF VESTIBULAR DISORDERS OF THE BÁRÁNY SOCIETY,  
313 *International classification of vestibular disorders*, (2018), <https://jvr-web.org/icvd>.
- 314 [5] INTERNATIONAL HEADACHE SOCIETY, *The international classification of headache disorders, 3rd*  
315 *edition*, *Cephalalgia*, 38 (2018), pp. 1–211, <https://doi.org/10.1177/0333102413485658>.
- 316 [6] INTERNATIONAL HEADACHE SOCIETY, *International classification of orofacial pain, 1st edition*,  
317 *Cephalalgia*, 40 (2022), pp. 129–221, <https://doi.org/10.1177/0333102419893823>.
- 318 [7] T. ROZEN, *Lash: A syndrome of long-lasting autonomic symptoms with hemicrania (a new*  
319 *indomethacin- responsive headache)*, *Headache: The Journal of Head and Face Pain*, 40  
320 (2000), pp. 483–486, <https://doi.org/10.1046/j.1526-4610.2000.00073.x>.
- 321 [8] P. ZHANG, *American headache society 64th annual scientific meeting june 9–12*, *Headache: The*  
322 *Journal of Head and Face Pain*, 62 (2022), pp. 1–170, [https://doi.org/doi/10.1111/head.](https://doi.org/doi/10.1111/head.14317)  
323 14317.  
324