**S1 Table.** Final 141 features selected for inclusion in the machine learning model to predict acute hepatic porphyria. Features are scored by number of occurrances in an individual patient medical record, and then normalized.

|  |  |  |  |
| --- | --- | --- | --- |
| **INDEX** | **FEATURE** | **SOURCE DOCUMENTS** | **DESCRIPTION** |
| 1 | ABDOMINAL\_PAIN\_DX\_NAME | Encounter Diagnosis, Patient Problem List | Text description of diagnosis code (ICD9) |
| 2 | ABDOMINAL\_PAIN\_UNSPECIFIED\_SITE\_DX\_NAME | Encounter Diagnosis, Patient Problem List | Text description of diagnosis code (ICD9) |
| 3 | ALTERNATIVE\_THERAPY\_-\_PINEAL\_HORMONE\_AGENTS\_PHARM\_SUBCLASS\_NAME | Concomittent Medications, Administered Medications, Medications Ordered | Text description of drug subclass |
| 4 | ANALGESIC\_OPIOID\_OXYCODONE\_COMBINATIONS\_PHARM\_SUBCLASS\_NAME | Concomittent Medications, Administered Medications, Medications Ordered | Text description of drug subclass |
| 5 | ANTI-ANXIETY\_-\_BENZODIAZEPINES\_PHARM\_CLASS\_NAME | Concomittent Medications, Administered Medications, Medications Ordered | Text description of drug class |
| 6 | ANTICONVULSANT\_-\_GABA\_ANALOGS\_PHARM\_SUBCLASS\_NAME | Concomittent Medications, Administered Medications, Medications Ordered | Text description of drug subclass |
| 7 | ANTIEMETIC\_-\_PHENOTHIAZINES\_PHARM\_SUBCLASS\_NAME | Concomittent Medications, Administered Medications, Medications Ordered | Text description of drug subclass |
| 8 | ANTIHISTAMINE\_-\_1ST\_GENERATION\_-\_ETHANOLAMINES\_PHARM\_SUBCLASS\_NAME | Concomittent Medications, Administered Medications, Medications Ordered | Text description of drug subclass |
| 9 | ANTIHISTAMINE\_-\_1ST\_GENERATION\_-\_PHENOTHIAZINES\_PHARM\_SUBCLASS\_NAME | Concomittent Medications, Administered Medications, Medications Ordered | Text description of drug subclass |
| 10 | BASO\_#\_COMPONENT\_NAME | Lab Results | Percent Basophils performed |
| 11 | CALCIUM\_REPLACEMENT\_PHARM\_CLASS\_NAME | Concomittent Medications, Administered Medications, Medications Ordered | Text description of drug class |
| 12 | CBC\_WITH\_DIFFERENTIAL\_PROC\_NAME | Procedures Ordered | CBC with diff order present |
| 13 | CNSLT0031\_PROC\_CODE | Procedures Ordered | Code for consult to Gastroenterology |
| 14 | CONSULT\_TO\_GASTROENTEROLOGY\_PROC\_NAME | Procedures Ordered | Consult to Gastoenterology ordered |
| 15 | COPD\_(CHRONIC\_OBSTRUCTIVE\_PULMONARY\_DISEASE)\_(HCC)\_DX\_NAME | Encounter Diagnosis, Patient Problem List | Text description of diagnosis code (ICD9) |
| 16 | CREATININE\_URINE\_CONCENTRATION\_COMPONENT\_NAME | Lab Results | lab result component present |
| 17 | CREATININEUR(REFERRAL)\_COMPONENT\_NAME | Lab Results | lab result component present |
| 18 | DIFFERENTIAL\_PROC\_NAME | Procedures Ordered | blood differential order present |
| 19 | DIPHENHYDRAMINE\_HCL\_GENERIC\_NAME\_1 | Concomittent Medication, Medications Ordered | Generic name of medication |
| 20 | ELEVATED\_WHITE\_BLOOD\_CELL\_COUNT\_UNSPECIFIED\_DX\_ICD10\_NAME | Encounter Diagnosis, Patient Problem List | Text description of diagnosis code (ICD10) |
| 21 | EOS\_#\_COMPONENT\_NAME | Lab Results | eosinaphil count lab result present |
| 22 | ESSENTIAL\_(PRIMARY)\_HYPERTENSION\_DX\_ICD10\_NAME | Encounter Diagnosis, Patient Problem List | Text description of diagnosis code (ICD10) |
| 23 | FERRITIN\_SERUM\_PROC\_NAME | Procedures Ordered | serum ferritin order present |
| 24 | HYDROMORPHONE\_HCL\_GENERIC\_NAME\_1 | Concomittent Medication, Medications Ordered | Generic name of medication |
| 25 | LAB00047\_PROC\_CODE | Procedures Ordered | Plasma lipase procedure ordered |
| 26 | LAB00364\_PROC\_CODE | Procedures Ordered | Microscopic urine exam ordered |
| 27 | LAB00681\_PROC\_CODE | Procedures Ordered | CBC with differential ordered |
| 28 | LAB100107\_PROC\_CODE | Procedures Ordered | Blood differential ordered |
| 29 | LAB100227\_PROC\_CODE | Procedures Ordered | Urine volume measurement ordered |
| 30 | LAB100882\_PROC\_CODE | Procedures Ordered | Multi-tube blood draw ordered |
| 31 | LIPASE\_\_(LAB)\_COMPONENT\_NAME | Lab Results | plasma lipase result component present |
| 32 | LIPASE\_PLASMA\_PROC\_NAME | Procedures Ordered | plasma lipase order present |
| 33 | LYMPHOCYTE\_#\_COMPONENT\_NAME | Lab Results | blood lymphocyte count results present |
| 34 | MAGNESIUM\_SALTS\_REPLACEMENT\_PHARM\_CLASS\_NAME | Concomittent Medications, Administered Medications, Medications Ordered | Text description of drug class |
| 35 | MELATONIN\_GENERIC\_NAME\_1 | Concomittent Medication, Medications Ordered | Generic name of medication |
| 36 | MINERALS\_AND\_ELECTROLYTES\_-\_CALCIUM\_REPLACEMENT/VITAMIN\_D\_COMBINATIONS\_PHARM\_SUBCLASS\_NAME | Concomittent Medications, Administered Medications, Medications Ordered | Text description of drug subclass |
| 37 | MISC\_REF\_TEST\_NAME\_COMPONENT\_NAME | Lab Results | Special test given with name of test in RESULT\_TEXT |
| 38 | MISC\_REF\_TEST\_RESULT\_COMPONENT\_NAME | Lab Results | Result of special test present |
| 39 | MONOCYTE\_#\_COMPONENT\_NAME | Lab Results | blood monocyte count results present |
| 40 | NAUSEA\_WITH\_VOMITING\_UNSPECIFIED\_DX\_ICD10\_NAME | Encounter Diagnosis, Patient Problem List | Text description of diagnosis code (ICD10) |
| 41 | NEUTROPHIL\_#\_COMPONENT\_NAME | Lab Results | blood neutrophil count results present |
| 42 | NGRAM\_0^pramipexole | Notes | Bigram of [token]^[token] found in free text. |
| 43 | NGRAM\_0^tablet | Notes | Bigram of [token]^[token] found in free text. |
| 44 | NGRAM\_10^olanzapine | Notes | Bigram of [token]^[token] found in free text. |
| 45 | NGRAM\_10^tablet | Notes | Bigram of [token]^[token] found in free text. |
| 46 | NGRAM\_100^sodium | Notes | Bigram of [token]^[token] found in free text. |
| 47 | NGRAM\_4^mg | Notes | Bigram of [token]^[token] found in free text. |
| 48 | NGRAM\_4^odt | Notes | Bigram of [token]^[token] found in free text. |
| 49 | NGRAM\_90^albuterol | Notes | Bigram of [token]^[token] found in free text. |
| 50 | NGRAM\_abdominal | Notes | Unigram of [token] found in free text. |
| 51 | NGRAM\_abdominal^pain | Notes | Bigram of [token]^[token] found in free text. |
| 52 | NGRAM\_acute | Notes | Unigram of [token] found in free text. |
| 53 | NGRAM\_acute^distress | Notes | Bigram of [token]^[token] found in free text. |
| 54 | NGRAM\_ambulatory | Notes | Unigram of [token] found in free text. |
| 55 | NGRAM\_antibiotics | Notes | Unigram of [token] found in free text. |
| 56 | NGRAM\_antibiotics^sulfonamide | Notes | Bigram of [token]^[token] found in free text. |
| 57 | NGRAM\_atraumatic | Notes | Unigram of [token] found in free text. |
| 58 | NGRAM\_bipolar | Notes | Unigram of [token] found in free text. |
| 59 | NGRAM\_cigarettes | Notes | Unigram of [token] found in free text. |
| 60 | NGRAM\_compazine | Notes | Unigram of [token] found in free text. |
| 61 | NGRAM\_control^pain | Notes | Bigram of [token]^[token] found in free text. |
| 62 | NGRAM\_depakote | Notes | Unigram of [token] found in free text. |
| 63 | NGRAM\_dilaudid | Notes | Unigram of [token] found in free text. |
| 64 | NGRAM\_discharged | Notes | Unigram of [token] found in free text. |
| 65 | NGRAM\_disintegrating | Notes | Unigram of [token] found in free text. |
| 66 | NGRAM\_docusate | Notes | Unigram of [token] found in free text. |
| 67 | NGRAM\_docusate^sodium | Notes | Bigram of [token]^[token] found in free text. |
| 68 | NGRAM\_dose^oral | Notes | Bigram of [token]^[token] found in free text. |
| 69 | NGRAM\_duloxetine | Notes | Unigram of [token] found in free text. |
| 70 | NGRAM\_ed | Notes | Unigram of [token] found in free text. |
| 71 | NGRAM\_edisylate] | Notes | Unigram of [token] found in free text. |
| 72 | NGRAM\_extended^tablet | Notes | Bigram of [token]^[token] found in free text. |
| 73 | NGRAM\_fibromyalgia | Notes | Unigram of [token] found in free text. |
| 74 | NGRAM\_flare | Notes | Unigram of [token] found in free text. |
| 75 | NGRAM\_flares | Notes | Unigram of [token] found in free text. |
| 76 | NGRAM\_focal | Notes | Unigram of [token] found in free text. |
| 77 | NGRAM\_gallops | Notes | Unigram of [token] found in free text. |
| 78 | NGRAM\_genitourinary | Notes | Unigram of [token] found in free text. |
| 79 | NGRAM\_glycol | Notes | Unigram of [token] found in free text. |
| 80 | NGRAM\_glycol^polyethylene | Notes | Bigram of [token]^[token] found in free text. |
| 81 | NGRAM\_gram | Notes | Unigram of [token] found in free text. |
| 82 | NGRAM\_hydromorphone | Notes | Unigram of [token] found in free text. |
| 83 | NGRAM\_instructed | Notes | Unigram of [token] found in free text. |
| 84 | NGRAM\_iv | Notes | Unigram of [token] found in free text. |
| 85 | NGRAM\_latex | Notes | Unigram of [token] found in free text. |
| 86 | NGRAM\_magnesium | Notes | Unigram of [token] found in free text. |
| 87 | NGRAM\_melatonin | Notes | Unigram of [token] found in free text. |
| 88 | NGRAM\_miralax | Notes | Unigram of [token] found in free text. |
| 89 | NGRAM\_mouth^needed | Notes | Bigram of [token]^[token] found in free text. |
| 90 | NGRAM\_mouth^twelve | Notes | Bigram of [token]^[token] found in free text. |
| 91 | NGRAM\_nausea | Notes | Unigram of [token] found in free text. |
| 92 | NGRAM\_nausea^vomiting | Notes | Bigram of [token]^[token] found in free text. |
| 93 | NGRAM\_odt | Notes | Unigram of [token] found in free text. |
| 94 | NGRAM\_odt^ondansetron | Notes | Bigram of [token]^[token] found in free text. |
| 95 | NGRAM\_olanzapine | Notes | Unigram of [token] found in free text. |
| 96 | NGRAM\_oncology | Notes | Unigram of [token] found in free text. |
| 97 | NGRAM\_ondansetron | Notes | Unigram of [token] found in free text. |
| 98 | NGRAM\_oral^powder | Notes | Bigram of [token]^[token] found in free text. |
| 99 | NGRAM\_oxycodone | Notes | Unigram of [token] found in free text. |
| 100 | NGRAM\_pain^severe | Notes | Bigram of [token]^[token] found in free text. |
| 101 | NGRAM\_pathology | Notes | Unigram of [token] found in free text. |
| 102 | NGRAM\_penicillins | Notes | Unigram of [token] found in free text. |
| 103 | NGRAM\_phenergan | Notes | Unigram of [token] found in free text. |
| 104 | NGRAM\_polyethylene | Notes | Unigram of [token] found in free text. |
| 105 | NGRAM\_powder | Notes | Unigram of [token] found in free text. |
| 106 | NGRAM\_pramipexole | Notes | Unigram of [token] found in free text. |
| 107 | NGRAM\_propranolol | Notes | Unigram of [token] found in free text. |
| 108 | NGRAM\_protocol | Notes | Unigram of [token] found in free text. |
| 109 | NGRAM\_psychosis | Notes | Unigram of [token] found in free text. |
| 110 | NGRAM\_risperidone | Notes | Unigram of [token] found in free text. |
| 111 | NGRAM\_rubs | Notes | Unigram of [token] found in free text. |
| 112 | NGRAM\_scoliosis | Notes | Unigram of [token] found in free text. |
| 113 | NGRAM\_seroquel | Notes | Unigram of [token] found in free text. |
| 114 | NGRAM\_severe | Notes | Unigram of [token] found in free text. |
| 115 | NGRAM\_stomach | Notes | Unigram of [token] found in free text. |
| 116 | NGRAM\_sulfa | Notes | Unigram of [token] found in free text. |
| 117 | NGRAM\_sulfonamide | Notes | Unigram of [token] found in free text. |
| 118 | NGRAM\_urine | Notes | Unigram of [token] found in free text. |
| 119 | NGRAM\_vicodin | Notes | Unigram of [token] found in free text. |
| 120 | NGRAM\_zofran | Notes | Unigram of [token] found in free text. |
| 121 | NORMAL\_RANGE\_COMPONENT\_NAME | Lab Results | Lab test result within normal ranges |
| 122 | OBSTRUCTIVE\_SLEEP\_APNEA\_(ADULT)\_(PEDIATRIC)\_DX\_ICD10\_NAME | Encounter Diagnosis, Patient Problem List | Text description of diagnosis code (ICD10) |
| 123 | OBSTRUCTIVE\_SLEEP\_APNEA\_DX\_NAME | Encounter Diagnosis, Patient Problem List | Text description of diagnosis code (ICD9) |
| 124 | ONDANSETRON\_HCL\_GENERIC\_NAME\_1 | Concomittent Medication, Medications Ordered | Generic name of medication |
| 125 | OXYCODONE\_HCL/ACETAMINOPHEN\_GENERIC\_NAME\_1 | Concomittent Medication, Medications Ordered | Generic name of medication |
| 126 | PATHOLOGY\_PROC\_NAME | Procedures Ordered | Transcribed pathology report present |
| 127 | PELVIC\_AND\_PERINEAL\_PAIN\_DX\_ICD10\_NAME | Encounter Diagnosis, Patient Problem List | Text description of diagnosis code (ICD10) |
| 128 | PINEAL\_HORMONE\_AGENTS\_PHARM\_CLASS\_NAME | Concomittent Medications, Administered Medications, Medications Ordered | Text description of drug subclass |
| 129 | PROCHLORPERAZINE\_EDISYLATE\_GENERIC\_NAME\_1 | Concomittent Medication, Medications Ordered | Generic name of medication |
| 130 | PROMETHAZINE\_HCL\_GENERIC\_NAME\_1 | Concomittent Medication, Medications Ordered | Generic name of medication |
| 131 | RADIOLOGY\_PROC\_NAME | Procedures Ordered | Transcribed radiology report present |
| 132 | RAINBOW\_HOLD\_TUBE\_-\_BLUE\_TOP\_PROC\_NAME | Procedures Ordered | Multi-tube blood draw ordered |
| 133 | RESTLESS\_LEGS\_SYNDROME\_DX\_ICD10\_NAME | Encounter Diagnosis, Patient Problem List | Text description of diagnosis code (ICD10) |
| 134 | TOBACCO\_ABUSE\_DX\_NAME | Encounter Diagnosis, Patient Problem List | Text description of diagnosis code (ICD9) |
| 135 | TRIPLE\_P04\_CRYSTALS\_COMPONENT\_NAME | Lab Results | Component of result of lab test |
| 136 | TRNS00039\_PROC\_CODE | Procedures Ordered | Transcribed pathology report present |
| 137 | TRNS00040\_PROC\_CODE | Procedures Ordered | Transcribed imaging report present |
| 138 | UNSPECIFIED\_ABDOMINAL\_PAIN\_DX\_ICD10\_NAME | Encounter Diagnosis, Patient Problem List | Text description of diagnosis code (ICD10) |
| 139 | UNSPECIFIED\_ABDOMINAL\_PAIN\_DX\_ICD10\_NAME | Encounter Diagnosis, Patient Problem List | Text description of diagnosis code (ICD10) |
| 140 | URINE\_MICROSCOPIC\_EXAM\_PROC\_NAME | Lab Results | Name of lab test procedure |
| 141 | VOL(URINE)\_PROC\_NAME | Lab Results | Name of lab test procedure |