## Supplementary material

## Table S1. The ICD-9-CM (International Classification of Disease, 9th Revision, Clinical Modification) codes of Influenza-like illness (ILI)

| ICD-9 code | Description |
|---|---|
| **079.99** | Unspecified virus infection in condition classified elsewhere and of unspecified site |
| **307.81** | Tension headache |
| **372.3** | Other and unspecified conjunctivitis |
| **382.9** | Unspecified otitis media |
| **460** | Acute nasopharyngitis [common cold] |
| **461.1** | Acute sinusitis frontal |
| **461.9** | Acute sinusitis, unspecified |
| **462** | Acute pharyngitis |
| **463** | Acute tonsillitis |
| **464** | Acute laryngitis and tracheitis |
| **464.4** | Croup |
| **465** | Acute upper respiratory infections of multiple or unspecified site |
| **465.8** | Acute upper respiratory infections of other multiple sites |
| **465.9** | Acute upper respiratory infections of unspecified site |
| **466** | Acute bronchitis and bronchiolitis |
| **466.0** | Acute bronchitis |
| **466.19** | Acute bronchiolitis due to other infectious organisms |
| **472** | Chronic pharyngitis and nasopharyngitis |
| **477.9** | Allergic rhinitis cause unspecified |
| **482.31** | Pneumonia due to Streptococcus, Group A |
| **482.32** | Pneumonia due to Streptococcus, Group B |
| **482.39** | Pneumonia due to other Streptococcus |
| **482.4** | Pneumonia due to Staphylococcus |
| **482.41** | Pneumonia due to Staphylococcus aureus |
| **482.49** | Pneumonia due to other Staphylococcus |
| **482.81** | Pneumonia due to Anaerobes |
| **482.82** | Pneumonia due to Escherichia coli (E. coli) |
| **482.83** | Pneumonia due to other gram-negative bacteria |
| **482.84** | Legionnaires' disease |
| **482.89** | Pneumonia due to other specified bacteria |
| **483** | Pneumonia due to other specified organism |
| **483.1** | Pneumonia due to Chlamydia |
| **484.1** | Pneumonia in cytomegalic inclusion disease |
| **484.3** | Pneumonia in whooping cough |

| ICD-9 code | Description |
|---|---|
| **484.6** | Pneumonia in aspergillosis |
| **484.7** | Pneumonia in other systemic mycoses |
| **486** | Pneumonia, organism unspecified |
| **487** | Influenza |
| **487.0** | Influenza with pneumonia |
| **487.1** | Influenza with other respiratory manifestations |
| **487.8** | Influenza with other manifestations |
| **490** | Bronchitis, not specified as acute or chronic |
| **496** | Chronic airways obstruction, not elsewhere classified |
| **780.6** | Fever |
| **780.79** | Other malaise and fatigue |
| **782.5** | Cyanosis |
| **784** | Symptoms involving head and neck |
| **785.6** | Enlargement of lymph nodes |
| **786.07** | Wheezing |
| **786.09** | Other dyspnea and respiratory abnormalities |
| **786.2** | Cough |
| **786.5** | Chest pain |
| **787.02** | Nausea alone |
| **787.03** | Vomiting alone |

1

**Table S2. The ICD-9-CM (International Classification of Disease, 9th Revision, Clinical Modification) codes for the 19 comorbidities incorporated in this study.**

| No. | Comorbidity* | ICD-9-CM codes | Catastrophic illness codes # |
|---|---|---|---|
| 1 | Heart disease | 394.x, 395.x, 396.x, 398.91, 402.0, 402.11, 402.91, 404.01, 404.11, 404.13, 404.91, 404.93, 410.x, 412, 411.x, 413.x, 414.x, 424.0, 424.1, 424.2, 424.3, 427.0, 427.2, 427.3, 427.31,427.32, 427.4x, 427.5, 427.6x, 427.8x, 427.9x, 428.x | |
| 2 | PVD | 440.2x, 440.3x, 443.x, 444.2x, 447.1, 785.4, V43.4 | |
| 3 | Hypertension | 401.x, 402.x, 403.x, 404.x, 405.x, 437.2 | |
| 4 | CVA | 342.x, 430, 431, 432.x, 433.x, 434.x, 435.x, 436, 437.x, 438.x, Procedure 38.48 | |
| 5 | Neurological disease | Dementia: 290.x, 331.0, 331.2; Epilepsy: 345.x | |
| 6 | Pulmonary disease | COPD: 490, 491.x, 492.x, 494.x, 496, 500, 501, 502, 503, 504, 505, 506.4; Asthma: 493.x | |
| 7 | Autoimmune disease | 340.x, 694.4, 694.5, 696.0, 696.1, 696.8, 710.x, 714.x, 720.x, 724.x | 05 (Autoimmune disease); 21 (Multiple sclerosis) |
| 8 | Severe liver disease | 070.22, 070.23, 070.32, 070.33, 070.54, 456.0, 456.1, 456.20, 456.21, 571.x, 573.x, 572.2-572.8 | 25 (Liver Cirrhosis ) |
| 9 | Diabetes | 250.0x, 250.1x, 250.2x, 250.3x, 250.8x, 250.9x, 250.4x, 250.5x, 250.6x, 250.7x | |
| 10 | Hyperthyroidism | 242.x | |
| 11 | Hypothyroidism | 243.x, 244.x | |
| 12 | Renal disease | 403.11, 403.91, 404.12, 404.13, 404.92, 404.93, 582.x, 583.x, 585, 586, 588.x | |
| 13 | Cancer without metastasis | 140.x-195.x<br>✻Patients without code 196.x-199.x and 200.xx-208.xx | 01 (Cancer) |
| 14 | Metastatic cancer | 196.x-199.x | 01 (Cancer) |
| 15 | Leukemia and lymphoma | 200.xx-208.xx<br>✻Patients with code 200.xx-208.xx, even though they have code 140.x-195.x and 196.x-199.x. | 01 (Cancer) |
| 16 | AIDS | 042 | |
| 17 | Tuberculosis | 010.xx-018.xx | |
| 18 | Mental illness | 295.x-299.x | 06 (Chronic mental illness) |
| 19 | Pregnancy/ postpartum | 630-679<br>Hospital Order Code: 57114C, 57115C, 81004C, 81005B, 81005C, 81010C, 81017C, 81024C, 81025C, 81026C, 81028C, 81029C, 81034C, 97001A, 97001K, 97002A, 97003B, 97004C, 97005D, 97006K, 97007A, 97008B, 97009C, 97014C, 97931K, 97932A, 97933B, 97934C, 98001K, 98002A, 98003B, 98004C, P3503C, P3504C | |

*Abbreviations: AIDS (acquired immune deficiency syndrome); COPD (chronic obstructive pulmonary disease); CVA (cerebrovascular accident); PVD (peripheral vascular disease).

#Patients with these comorbidities were confirmed with records on catastrophic illness card through National Health Insurance in Taiwan, and the specific code number represented its own category of catastrophic Illness and subsidies.

**Table S3. Parameter settings for the DT and DNN models**

| Model | Package | Parameter Setting |
|---|---|---|
| **DT** | **R Package rpart** | split = Information |
| | | minbuck = 100 |
| | | xval = 10 |
| | | prior = 0.35 ~ 0.85 with 0.01 resolution |
| | | cp = 0.001~0.008 with0.001 resolution |
| **Multilayer DNN** | **Python Keras** | Input neurons: 6 or 18 corresponding to 6 or 18 features. |
| | | hidden layers: 3 or 4 |
| | | hidden neurons: 8,16, 24, 32. |
| | | 1 output neuron |

**Table S4. Characteristics of the three DT models incorporated in the proposed DT-based multivariate analysis**

| | Model with 85.50% sensitivity | Model with 90.82% sensitivity | Model with 95.95% sensitivity |
|---|---|---|---|
| Parameter values | prior=0.51.<br>cp=0.0040 | prior=0.60<br>cp=0.0007 | prior=0.81<br>cp=0.0008 |
| Features | 1. Age<br>2. Gender<br>3. Heart Disease<br>4. Metastatic Cancer<br>5. CVA<br>6. Diabetes | 1.Age<br>2.Gender<br>3.Heart Disease<br>4.Metastatic Cancer<br>5.CVA<br>6.Diabetes<br>7.Cancer without Metastasis<br>8.Renal Disease<br>9.Neurological Disease | 1.Age<br>2.Gender<br>3.CVA<br>4.Diabetes<br>5.Metastatic Cancer<br>6.Heart Disease<br>7.Autoimmune Disease<br>8.Neurological Disease<br>9.Pregnancy/ Postpartum<br>10. Pulmonary Disease |

Note:

1. The following six features were present in all three DT models with different levels of sensitivity: age, gender, heart disease, metastatic cancer, cerebrovascular accident (CVA), and diabetes.

**Table S5. Definitions of performance metrics**

| metric | Definition |
|---|---|
| TP | True positive |
| TN | True negative |
| FP | False positive |
| FN | False negative |
| Accuracy | (TP+TN)/(TP+FP+TN+FN) |
| Sensitivity | TP/(TP+FN) |
| Specificity | TN/(TN+FP) |
| PPV | TP/(TP+FP) |
| NPV | TN/(TN+FN) |
| F1 score | 2TP/(2TP+FP+FN) |
| MCC | Matthews correlation coefficient |

Abbreviations: PPV (positive predictive value); NPV (negative predictive value).

**Table S6(a). Detailed performance data of the DT prediction models**

| Model | Feature set | Performance metric | Model with 85% sensitivity | Model with 90% sensitivity | Model with 95% sensitivity |
|---|---|---|---|---|---|
| **DT** | **6 features (by the proposed DT-based method)** | Accuracy | 61.564% (61.534% - 61.593%) | 58.048% (57.996% - 58.100%) | 50.498% (50.389% - 50.608%) |
| | | PPV | 30.141% (30.128% - 30.155%) | 28.778% (28.758% - 28.798%) | 26.085% (26.044% - 26.126%) |
| | | Sensitivity | 86.007% (85.981% - 86.034%) | 90.073% (90.026% - 90.119%) | 95.300% (95.274% - 95.325%) |
| | | Specificity | 56.192% (56.150% - 56.234%) | 51.010% (50.937% - 51.083%) | 40.653% (40.515% - 40.791%) |
| | | F1 | 0.446 (0.446 - 0.447) | 0.436 (0.436 - 0.436) | 0.410 (0.409 - 0.410) |
| | | MCC | 0.324 (0.324 - 0.325) | 0.318 (0.318 - 0.319) | 0.291 (0.291 - 0.292) |
| | **6 features (by mRMRe)** | Accuracy | 60.493% (60.477% - 60.510%) | 56.144% (56.138% - 56.150%) | 50.343% (50.325% - 50.361%) |
| | | PPV | 29.428% (29.421% - 29.435%) | 27.951% (27.948% - 27.953%) | 26.025% (26.019% - 26.031%) |
| | | Sensitivity | 85.314% (85.288% - 85.340%) | 90.901% (90.887% - 90.915%) | 95.315% (95.307% - 95.324%) |
| | | Specificity | 55.038% (55.014% - 55.063%) | 48.506% (48.497% - 48.514%) | 40.459% (40.436% - 40.483%) |
| | | F1 | 0.438 (0.438 - 0.438) | 0.428 (0.428 - 0.428) | 0.409 (0.409 - 0.409) |
| | | MCC | 0.310 (0.310 - 0.311) | 0.307 (0.307 - 0.308) | 0.290 (0.290 - 0.290) |
| | **6 features (by LASSO)** | Accuracy | 60.151% (60.114% - 60.187%) | 56.349% (56.298% - 56.399%) | 48.138% (48.129% - 48.146%) |
| | | PPV | 29.311% (29.299% - 29.324%) | 27.990% (27.971% - 28.009%) | 25.255% (25.253% - 25.257%) |
| | | Sensitivity | 85.840% (85.776% - 85.903%) | 90.467% (90.427% - 90.508%) | 95.863% (95.850% - 95.877%) |
| | | Specificity | 54.505% (54.447% - 54.563%) | 48.850% (48.780% - 48.921%) | 37.649% (37.636% - 37.662%) |
| | | F1 | 0.437 (0.437 - 0.437) | 0.428 (0.427 - 0.428) | 0.400 (0.400 - 0.400) |
| | | MCC | 0.311 (0.310 - 0.311) | 0.306 (0.306 - 0.307) | 0.277 (0.277 - 0.277) |
| | **18 features** | Accuracy | 61.088% (61.072% - 61.104%) | 57.858% (57.852% - 57.863%) | 50.775% (50.717% - 50.834%) |
| | | PPV | 29.789% (29.785% - 29.793%) | 28.780% (28.778% - 28.782%) | 26.279% (26.257% - 26.301%) |
| | | Sensitivity | 85.468% (85.429% - 85.508%) | 90.804% (90.796% - 90.812%) | 95.944% (95.932% - 95.956%) |
| | | Specificity | 55.730% (55.702% - 55.758%) | 50.617% (50.609% - 50.625%) | 40.849% (40.776% - 40.922%) |
| | | F1 | 0.442 (0.442 - 0.442) | 0.437 (0.437 - 0.437) | 0.413 (0.412 - 0.413) |
| | | MCC | 0.317 (0.317 - 0.317) | 0.321 (0.321 - 0.321) | 0.298 (0.298 - 0.298) |

Note: Please refer to supplementary Table S5 for the definitions of performance metrics.

**Table S6(b). Detailed performance data of the LR prediction models.**

| Model | Feature set | Performance metric | Model with 85% sensitivity | Model with 90% sensitivity | Model with 95% sensitivity |
|---|---|---|---|---|---|
| **LR** | **6 features (by the proposed DT-based method)** | Accuracy | 61.070% (61.039% - 61.101%) | 56.996% (56.965% - 57.028%) | 48.538% (48.487% - 48.589%) |
| | | PPV | 29.712% (29.694% - 29.730%) | 28.241% (28.226% - 28.256%) | 25.291% (25.272% - 25.310%) |
| | | Sensitivity | 84.996% (84.994% - 84.997%) | 89.999% (89.997% - 90.001%) | 94.999% (94.998% - 95.000%) |
| | | Specificity | 55.812% (55.774% - 55.850%) | 49.744% (49.706% - 49.782%) | 38.328% (38.265% - 38.390%) |
| | | F1 | 0.440 (0.440 - 0.441) | 0.430 (0.430 - 0.430) | 0.399 (0.399 - 0.400) |
| | | MCC | 0.314 (0.314 - 0.314) | 0.309 (0.309 - 0.309) | 0.274 (0.273 - 0.274) |
| | **6 features (by mRMRe)** | Accuracy | 60.176% (60.157% - 60.195%) | 55.831% (55.797% - 55.866%) | 48.309% (48.262% - 48.357%) |
| | | PPV | 29.205% (29.194% - 29.216%) | 27.680% (27.664% - 27.696%) | 25.206% (25.188% - 25.223%) |
| | | Sensitivity | 84.994% (84.990% - 84.998%) | 90.004% (90.000% - 90.008%) | 94.999% (94.998% - 95.000%) |
| | | Specificity | 54.722% (54.698% - 54.745%) | 48.321% (48.279% - 48.364%) | 38.049% (37.991% - 38.107%) |
| | | F1 | 0.435 (0.435 - 0.435) | 0.423 (0.423 - 0.424) | 0.398 (0.398 - 0.399) |
| | | MCC | 0.306 (0.305 - 0.306) | 0.299 (0.299 - 0.299) | 0.272 (0.272 - 0.272) |
| | **6 features (by LASSO)** | Accuracy | 59.835% (59.802% - 59.869%) | 55.426% (55.391% - 55.462%) | 48.017% (47.982% - 48.052%) |
| | | PPV | 29.017% (28.998% - 29.035%) | 27.490% (27.474% - 27.507%) | 25.098% (25.085% - 25.111%) |
| | | Sensitivity | 84.994% (84.992% - 84.996%) | 90.005% (90.002% - 90.008%) | 95.001% (94.998% - 95.004%) |
| | | Specificity | 54.306% (54.265% - 54.348%) | 47.827% (47.783% - 47.871%) | 37.692% (37.649% - 37.735%) |
| | | F1 | 0.433 (0.432 - 0.433) | 0.421 (0.421 - 0.421) | 0.397 (0.397 - 0.397) |
| | | MCC | 0.303 (0.302 - 0.303) | 0.296 (0.295 - 0.296) | 0.270 (0.270 - 0.270) |
| | **18 features** | Accuracy | 62.006% (61.985% - 62.026%) | 57.671% (57.641% - 57.702%) | 49.901% (49.842% - 49.959%) |
| | | PPV | 30.261% (30.249% - 30.274%) | 28.577% (28.561% - 28.592%) | 25.810% (25.788% - 25.833%) |
| | | Sensitivity | 84.996% (84.995% - 84.997%) | 89.997% (89.996% - 89.998%) | 94.998% (94.997% - 94.999%) |
| | | Specificity | 56.953% (56.928% - 56.978%) | 50.567% (50.530% - 50.604%) | 39.990% (39.919% - 40.061%) |
| | | F1 | 0.446 (0.446 - 0.446) | 0.434 (0.434 - 0.434) | 0.406 (0.406 - 0.406) |
| | | MCC | 0.322 (0.322 - 0.323) | 0.315 (0.314 - 0.315) | 0.285 (0.284 - 0.285) |

Note: Please refer to supplementary Table S5 for the definitions of performance metrics.

**Table S6(c). Detailed performance data of the DNN prediction models.**

| Model | Feature set | Performance metrics | Model with 85% sensitivity | Model with 90% sensitivity | Model with 95% sensitivity |
|---|---|---|---|---|---|
| **DNN** | **6 features (by the proposed DT-based method)** | Accuracy | 62.090% (62.065% - 62.115%) | 58.168% (58.138% - 58.197%) | 50.449% (50.376% - 50.523%) |
| | | PPV | 30.312% (30.297% - 30.327%) | 28.829% (28.814% - 28.844%) | 26.026% (25.997% - 26.055%) |
| | | Sensitivity | 84.996% (84.994% - 84.999%) | 90.000% (89.998% - 90.003%) | 94.998% (94.996% - 95.000%) |
| | | Specificity | 57.055% (57.025% - 57.086%) | 51.172% (51.135% - 51.208%) | 40.659% (40.570% - 40.749%) |
| | | F1 | 0.447 (0.447 - 0.447) | 0.437 (0.437 - 0.437) | 0.409 (0.408 - 0.409) |
| | | MCC | 0.323 (0.323 - 0.323) | 0.319 (0.319 - 0.319) | 0.289 (0.288 - 0.289) |
| | **6 features (by mRMRe)** | Accuracy | 60.755% (60.731% - 60.779%) | 57.117% (57.088% - 57.147%) | 50.733% (50.678% - 50.788%) |
| | | PPV | 29.532% (29.518% - 29.546%) | 28.302% (28.287% - 28.316%) | 26.139% (26.117% - 26.160%) |
| | | Sensitivity | 84.998% (84.995% - 85.001%) | 90.006% (90.003% - 90.009%) | 94.999% (94.997% - 95.002%) |
| | | Specificity | 55.427% (55.398% - 55.457%) | 49.890% (49.854% - 49.926%) | 41.005% (40.938% - 41.072%) |
| | | F1 | 0.438 (0.438 - 0.438) | 0.431 (0.430 - 0.431) | 0.410 (0.410 - 0.410) |
| | | MCC | 0.311 (0.311 - 0.311) | 0.310 (0.310 - 0.310) | 0.291 (0.291 - 0.291) |
| | **6 features (by LASSO)** | Accuracy | 60.866% (60.827% - 60.906%) | 56.950% (56.913% - 56.988%) | 49.935% (49.874% - 49.996%) |
| | | PPV | 29.595% (29.572% - 29.618%) | 28.219% (28.201% - 28.237%) | 25.824% (25.800% - 25.847%) |
| | | Sensitivity | 84.995% (84.992% - 84.999%) | 90.005% (90.002% - 90.007%) | 94.999% (94.996% - 95.003%) |
| | | Specificity | 55.563% (55.515% - 55.612%) | 49.686% (49.641% - 49.732%) | 40.031% (39.956% - 40.105%) |
| | | F1 | 0.439 (0.439 - 0.439) | 0.430 (0.429 - 0.430) | 0.406 (0.406 - 0.406) |
| | | MCC | 0.312 (0.312 - 0.312) | 0.309 (0.308 - 0.309) | 0.285 (0.284 - 0.285) |
| | **18 features** | Accuracy | 62.800% (62.741% - 62.860%) | 59.349% (59.297% - 59.400%) | 53.034% (52.988% - 53.081%) |
| | | PPV | 30.744% (30.708% - 30.781%) | 29.447% (29.420% - 29.475%) | 27.091% (27.071% - 27.110%) |
| | | Sensitivity | 84.995% (84.995% - 84.995%) | 89.997% (89.996% - 89.998%) | 94.999% (94.998% - 95.000%) |
| | | Specificity | 57.923% (57.850% - 57.995%) | 52.613% (52.551% - 52.676%) | 43.812% (43.756% - 43.868%) |
| | | F1 | 0.452 (0.451 - 0.452) | 0.444 (0.443 - 0.444) | 0.422 (0.421 - 0.422) |
| | | MCC | 0.330 (0.329 - 0.330) | 0.329 (0.329 - 0.330) | 0.309 (0.309 - 0.310) |

Note: Please refer to supplementary Table S5 for the definitions of performance metrics.