# **Online-Only Supplemental Material**

**eMethod1.** Markov chains

**eMethod2.** The application of Weighted Association Rules Mining

**eMethod3.** Automatic search of important associations from the results of the Markov chain analysis

**eMethod4.** Discussion of validity

**eTable1.** Distribution of the total patients and multimorbidity patients by year

**eTable2.** Selection of 103 chronic diseases with their prevalence in Intego

**eTable3.** The important associations from the Markov Chain results

**eFigure1**. Subgroup analysis – Markov chain analysis

**eFigure2**. Subgroup analysis – WARM

**Reference**

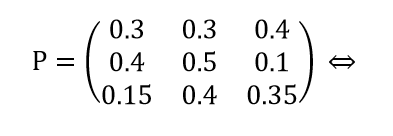
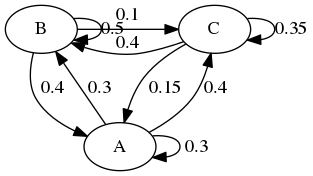
# **eMethod1. Markov chains**

A Markov chain is a stochastic model which describes a sequence of possible events whose probability only depends on the state attained in the previous event.1 The “memorylessness” property of the Markov chain allows us to predict the next state of the process only based on the current state and the accuracy of the prediction will be as good as the prediction based on the full history. This stochastic model is widely used in various areas such as control system, finance, queueing problems, game theory and so on.2 It is also applied in the area of healthcare and medical decision making, for example, to estimate transition between different status after surgery.3 It is a useful tool to conduct cost effectiveness analysis and help decision makers learn more about possible transitions between different statuses.

Markov chains are often described by the transition matrix from time *n* to time *n+1*. The transition matrix describes the probabilities of particular transitions, each element of the matrix is the transition probability from state *i* to state *j* in one time step. A sequence of directed graphs are often used to represent the Markov chains and transition matrices, where the edges of graph *n* are labeled by the probability of going from one state at time *n* to the other states at time *n+1*.

The figure below shows an example of the transition matrix of a simple Markov chain and the visualization of the transition matrix. This can be regarded as estimated probability matrix of 3 diseases. The probability of the development from disease A to disease B is 0.3 (the number in the blue square in the matrix and plot) while the probability of the development from disease B to disease A is 0.4 (the number in the red square). Other probabilities can be interpreted in the same way. This is a good method for the practice-oriented researchers to study the pattern of disease evolution quantitatively.

An example to demonstrate how to implement the Markov chain to study multimorbidity is uploaded to Github (<https://github.com/XiShi6/Multimorbidity>).

***Figure eMethod1.1*** *An Example of a Markov Chain Transition Matrix and Visualization*

# **eMethod2. The application of Weighted Association Rules Mining**

Association Rule Mining (ARM) is a method to uncover the combinations of items that occur together frequently, discovering interesting relations. It is often applied in the studies of supermarket sales and customer behaviors and is gaining attention in the clinical research. There are already some studies applying ARM to discover the frequent combination of disease occurrence.4,5 However, as stated in Introduction of the paper, ARM doesn’t take the sequence into account. Besides, association rules among less frequent diseases will turn out to be strong if diseases with a large difference in prevalence are involved, because with fewer observations, the weight of each observation will take up a larger percentage. If two low-prevalence diseases with limited observations coincidentally co-occur for several times, it will result in strong association rules when using traditional Association Rule Mining, while these cases are of noninterest. Weighted Association Rule Mining (WARM) is a good solution to overcome the two limitations of traditional Association Rule Mining.

One limitation of the first-order Markov chain has been fully discussed in Discussion. Therefore, WARM was applied as a supplementary method of the Markov chain. The results of WARM are rules derived from a large amount of transactions and can be written as , *X* and *Y* are two different sets of items, known as itemsets. The itemsets can include either one or multiple diseases and they can differ from the combinations appeared in the transactions, as these itemsets are the summarized results instead of original data. As shown in Table 2, it was possible to derive a rule with multiple diseases in the itemset, . In this case, with a combination application of the Markov chain and WARM, we could example the patterns of diseases without the limitation of number of diseases.

WARM introduces weights to make the co-occurrence with the same items different from each other if the sequence is different. In our study, the transaction and itemset weights were calculated using the HITS algorithm,6,7 and the frequent itemsets were found based on the Eclat algorithm.8 The weighted support of an itemset was a generalization of support taking the weights of the transactions into consideration and the itemset was frequent if the weighted support was higher than the threshold. These weights were determined by the global link structure of the database, making it more reasonable than counting-based measurement. The weighted confidence was calculated from the weighted support of the itemset. These weighted measures can be different if the sequence changes. Although *X→Y* and *Y→X* have the same itemsets, the rule with better weighted measures means the sequence of this rule is more often observed. The weighted association rules were created from the weighted frequent itemsets following the efficient algorithm proposed by Pei et.al.9

An example to demonstrate how to implement WARM to study multimorbidity is uploaded to Github (<https://github.com/XiShi6/Multimorbidity>).

# **eMethod3. Automatic search of important associations from the results of the Markov chain analysis**

As it may be difficult to go through the whole heatmap to find all interesting associations, we conducted an automatic search to detect the important associations by selecting the top three diseases with the highest probability in each column and selecting the diseases with probability higher than 0.05 in each row (eTable 3). The total sum of probability in each row equals to 1, meaning that larger probability takes up a larger share compared to all other diseases. It is possible that one disease has relatively equal probabilities around 0.0097 with all other diseases, in this case, it cannot be regarded as strong association even though the probability is the highest among the whole row. Thus, the search by row was determined by the absolute values, indicating the most likely follow-up diseases after the diagnosis of one chronic disease. In contrast, high absolute values are not very meaningful when searching by column, as columns for diseases with high prevalence have high absolute values in almost the whole column, such as the vertical line for Hypertension and Diabetes in Figure 1. Therefore, the top three diseases with the highest probability in the column were selected as the most likely antecedent diseases.

There is also a user-friendly way of interacting with a heatmap: that is, it is in principle possible to click on cells or regions in the heatmap to generate a more detailed picture of the pattern of relationships of the specific diseases involved; another strategy is to depict the heatmap into a graph and then visual bits and pieces of the graph. We do not present such features in this paper, but this can be done in principle for practice-oriented health care professionals or practice-oriented researchers when it is of interest to focus on specific combinations of diseases or specific clusters of diseases.

# **eMethod4. Discussion of validity**

The results of a Markov chain analysis may be questioned because the majority of the probabilities are very small, lower than 0.1. Although the values are small, they are significantly larger than the average level if we take the total number into consideration. If the probability of all other conditions diseases becoming the subsequent of one specific condition (e.g. alcohol abuse) had been equal, the average probability would have been 0.0098 (1/102). The gap is indubitable when we compare the average probability with the actual probability of developing depression (p=0.102), hypertension (p=0.060) and diabetes (p=0.044) after having diagnosis of alcohol abuse, as the actual probabilities of diabetes were 4.5 times higher than the average probability and that of depression were as large as more than 10 times. Besides, some conditions had probabilities much lower than the average probability or even close to 0. Under the principle of closer relationship leading to higher probability and the opposite with lower probability, we can observe a clear division to indicate common multimorbidity. In this case, although the values seem low, the methodology provides significant contrast in the results to distinguish likely or unlikely sequence of conditions.

It is the same in the case of WARM where the support of the selected rules might be interpreted at first sight as very low. The average probability of certain combinations is trivial when the total number is 103. The total number of possible combinations of selecting 2 out of 103 is 5253 and it is 176851 when selecting 3 out of 103. WARM does not have restrictions on how many conditions to include in one rule, leading to a huge total possible number (1.014E+31). Therefore, selected rules can be considered relevant when the selection criteria includes support > 0.001.

# **eTable1. Distribution of the total patients and multimorbidity patients by year**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Total Patients** | | | **Multimorbidity Patients** | | |
| **Year** | **Number Of the First Diagnosis** | **Number Of the Last Diagnosis** | **Number of Patients** | **Number Of the First Diagnosis** | **Number Of the Last Diagnosis** | **Number of Patients** |
| 1991 | 6704 | 1163 | 6704 | 5751 | 210 | 5751 |
| 1992 | 5430 | 1295 | 10971 | 4526 | 391 | 10067 |
| 1993 | 4537 | 1201 | 14213 | 3757 | 421 | 13433 |
| 1994 | 4634 | 1437 | 17646 | 3801 | 604 | 16813 |
| 1995 | 4338 | 1601 | 20547 | 3420 | 683 | 19629 |
| 1996 | 4163 | 1762 | 23109 | 3242 | 841 | 22188 |
| 1997 | 3748 | 1709 | 25095 | 2952 | 913 | 24299 |
| 1998 | 4034 | 2048 | 27420 | 3086 | 1100 | 26472 |
| 1999 | 3825 | 2315 | 29197 | 2816 | 1306 | 28188 |
| 2000 | 4952 | 2949 | 31834 | 3683 | 1680 | 30565 |
| 2001 | 3716 | 2663 | 32601 | 2623 | 1570 | 31508 |
| 2002 | 4159 | 3096 | 34097 | 2878 | 1815 | 32816 |
| 2003 | 4576 | 3655 | 35577 | 3119 | 2198 | 34120 |
| 2004 | 4295 | 4042 | 36217 | 2819 | 2566 | 34741 |
| 2005 | 4003 | 4101 | 36178 | 2551 | 2649 | 34726 |
| 2006 | 3729 | 4278 | 35806 | 2239 | 2788 | 34316 |
| 2007 | 3451 | 4623 | 34979 | 1982 | 3154 | 33510 |
| 2008 | 3416 | 4977 | 33772 | 1933 | 3494 | 32289 |
| 2009 | 3456 | 5436 | 32251 | 1813 | 3793 | 30608 |
| 2010 | 3354 | 6146 | 30169 | 1598 | 4390 | 28413 |
| 2011 | 3605 | 7212 | 27628 | 1578 | 5185 | 25601 |
| 2012 | 3097 | 7334 | 23513 | 1318 | 5555 | 21734 |
| 2013 | 2633 | 6763 | 18812 | 1059 | 5189 | 17238 |
| 2014 | 2569 | 7773 | 14618 | 890 | 6094 | 12939 |
| 2015 | 2208 | 9053 | 9053 | 505 | 7350 | 7350 |

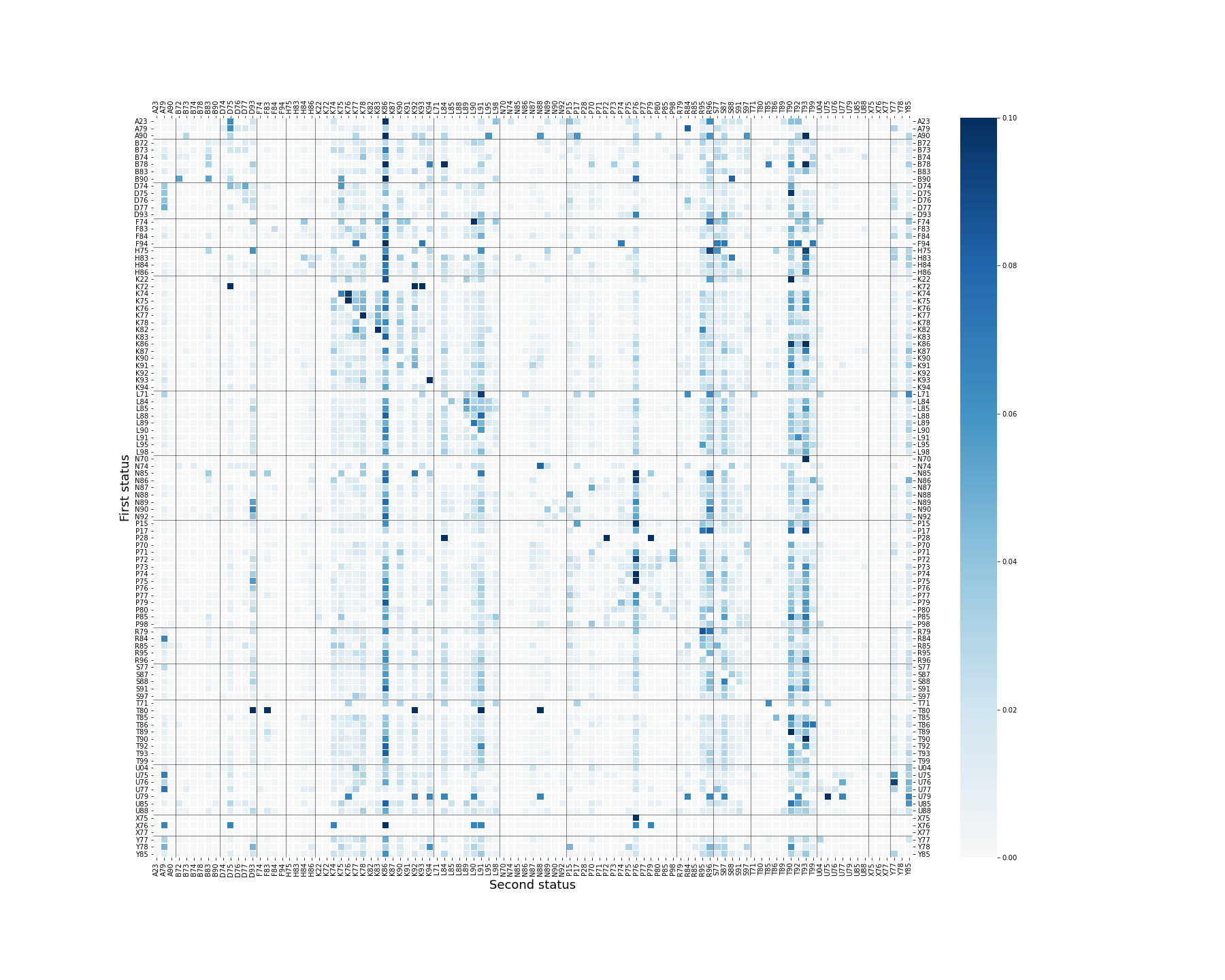
# **eTable2. Selection of 103 chronic diseases with their prevalence in Intego**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ICPC | Disease Name | Prevalence | ICPC | Disease Name | Prevalence |
| A23 | Risk factor NOS | 0.115 | N88 | Epilepsy | 1.344 |
| A79 | Malignancy NOS | 2.277 | N89 | Migraine | 5.247 |
| A90 | Congenital anomaly OS/multiple | 0.104 | N90 | Cluster headache | 0.470 |
| B72 | Hodgkin's disease/lymphoma | 0.509 | N92 | Trigeminal neuralgia | 0.901 |
| B73 | Leukaemia | 0.437 | P15 | Chronic alcohol abuse | 2.915 |
| B74 | Malignant neoplasm blood other | 0.220 | P17 | Tobacco abuse | 1.794 |
| B78 | Hereditary haemolytic anaemia | 0.121 | P28 | Limited function/disability (p) | 0.014 |
| B83 | Purpura/coagulation defect | 1.033 | P70 | Dementia | 2.734 |
| B90 | HIV-infection/aids | 0.103 | P71 | Organic psychosis other | 0.737 |
| D74 | Malignant neoplasm stomach | 0.305 | P72 | Schizophrenia | 0.577 |
| D75 | Malignant neoplasm colon/rectum | 2.748 | P73 | Affective psychosis | 0.414 |
| D76 | Malignant neoplasm pancreas | 0.363 | P74 | Anxiety disorder/anxiety state | 2.629 |
| D77 | Malignant neoplasm digest other/NOS | 0.791 | P75 | Somatization disorder | 2.214 |
| D93 | Irritable bowel syndrome | 10.145 | P76 | Depressive disorder | 16.239 |
| F74 | Neoplasm of eye/adnexa | 0.057 | P77 | Suicide/suicide attempt | 0.720 |
| F83 | Retinopathy | 0.943 | P79 | Phobia/compulsive disorder | 0.825 |
| F84 | Macular degeneration | 0.627 | P80 | Personality disorder | 0.464 |
| F94 | Blindness | 0.035 | P85 | Mental retardation | 0.136 |
| H75 | Neoplasm of ear | 0.064 | P98 | Psychosis NOS/other | 0.632 |
| H83 | Otosclerosis | 0.192 | R79 | Chronic bronchitis | 2.048 |
| H84 | Presbyacusis | 0.881 | R84 | Malignant neoplasm bronchus/lung | 1.254 |
| H86 | Deafness | 1.308 | R85 | Malinant neoplasm respiratory, other | 0.225 |
| K22 | Risk factor cardiovascular disease | 0.348 | R95 | Chronic obstructive pulmonary dis | 6.288 |
| K72 | Neoplasm cardiovascular | 0.022 | R96 | Asthma | 12.264 |
| K74 | Ischaemic heart disease w. angina | 4.955 | S77 | Malignant neoplasm of skin | 2.512 |
| K75 | Acute myocardial infarction | 3.946 | S87 | Dermatitis/atopic eczema | 10.716 |
| K76 | Ischaemic heart disease w/o angina | 3.064 | S88 | Dermatitis contact/allergic | 6.318 |
| K77 | Heart failure | 3.674 | S91 | Psoriasis | 3.036 |
| K78 | Atrial fibrillation/flutter | 6.030 | S97 | Chronic ulcer skin | 2.402 |
| K82 | Pulmonary heart disease | 0.536 | T71 | Malignant neoplasm thyroid | 0.105 |
| K83 | Heart valve disease NOS | 3.514 | T80 | Congenital anom endocrine/metab. | 0.043 |
| K86 | Hypertension uncomplicated | 2.731 | T85 | Hyperthyroidism/thyrotoxicosis | 2.374 |
| K87 | Hypertension complicated | 0.389 | T86 | Hypothyroidism/myxoedema | 2.549 |
| K90 | Stroke/cerebrovascular accident | 4.632 | T89 | Diabetes insulin dependent | 0.920 |
| K91 | Cerebrovascular disease | 0.387 | T90 | Diabetes non-insulin dependent | 14.513 |
| K92 | Atherosclerosis/PVD | 4.527 | T92 | Gout | 5.216 |
| K93 | Pulmonary embolism | 1.176 | T93 | Lipid disorder | 18.515 |
| K94 | Phlebitis/thrombophlebitis | 6.187 | T99 | Endocrine/metab/nutrit. dis. other | 3.880 |
| L71 | Malignant neoplasm musculoskeletal | 0.063 | U04 | Incontinence urine | 4.291 |
| L84 | Back syndrome w/o radiating pain | 10.059 | U75 | Malignant neoplasm of kidney | 0.307 |
| L85 | Acquired deformity of spine | 1.946 | U76 | Malignant neoplasm of bladder | 0.485 |
| ICPC | **Disease Name** | **Prevalence** | **ICPC** | **Disease Name** | **Prevalence** |
| L88 | Rheumatoid/seropositive arthritis | 1.926 | U77 | Malignant neoplasm urinary other | 0.134 |
| L89 | Osteoarthrosis of hip | 4.439 | U79 | Neoplasm urinary tract NOS | 0.027 |
| L90 | Osteoarthrosis of knee | 8.842 | U85 | Congenital anomaly urinary tract | 0.304 |
| L91 | Osteoarthrosis other | 13.724 | U88 | Glomerulonephritis/nephrosis | 0.415 |
| L95 | Osteoporosis | 5.424 | X75 | Malignant neoplasm cervix | 0.264 |
| L98 | Acquired deformity of limb | 4.780 | X76 | Malignant neoplasm breast female | 3.129 |
| N70 | Poliomyelitis | 0.003 | X77 | Malignant neoplasm genital other (f) | 0.584 |
| N74 | Malignant neoplasm nervous | 0.216 | Y77 | Malignant neoplasm prostate | 1.776 |
| N85 | Congenital anomaly neurological | 0.081 | Y78 | Malign neoplasm male genital other | 0.076 |
| N86 | Multiple sclerosis | 0.277 | Y85 | Benign prostatic hypertrophy | 3.174 |
| N87 | Parkinsonism | 1.452 |  |  |  |

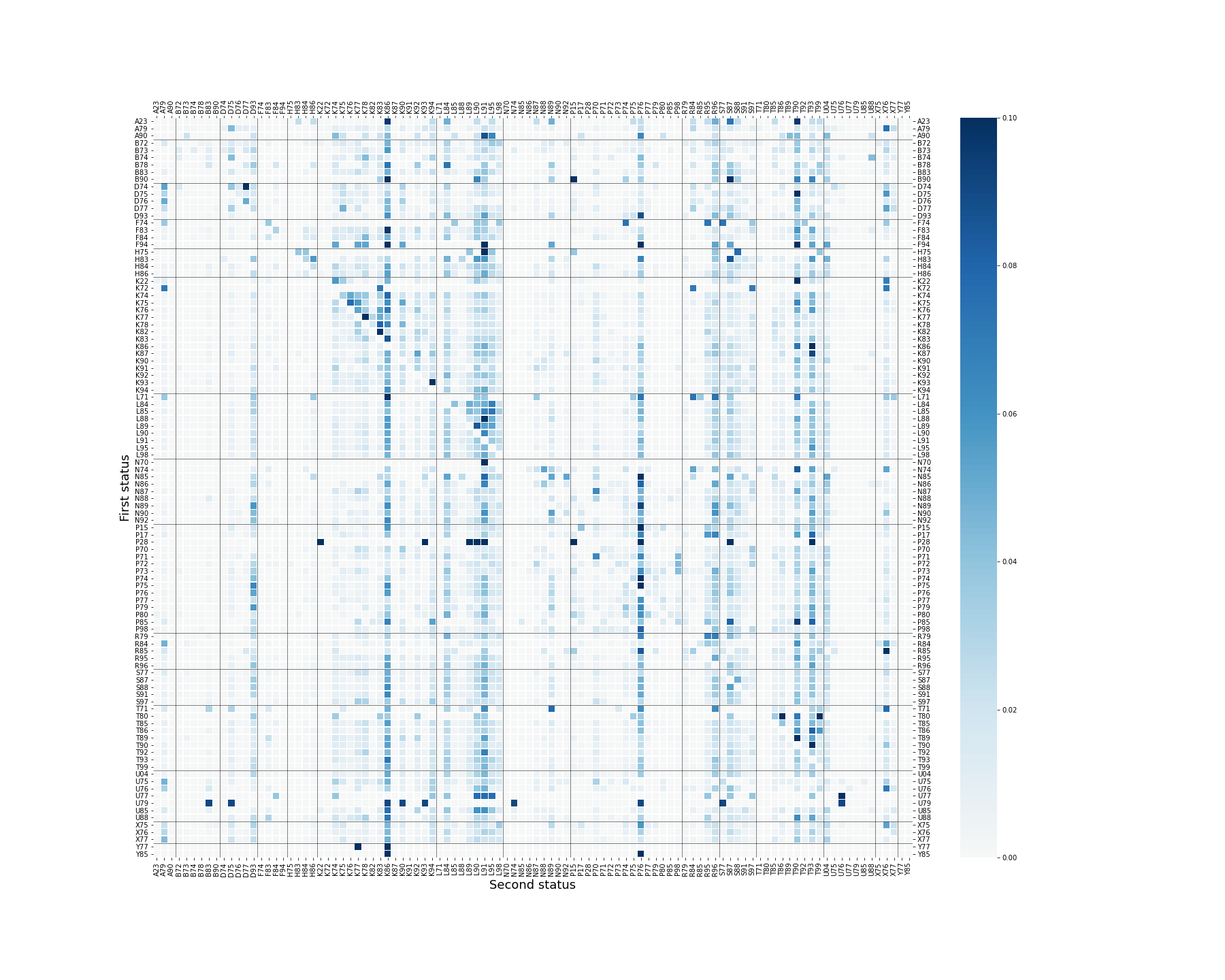
# **eTable3. The important associations from the Markov Chain results**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Antecedents | Disease | Consequents | Antecedents | Disease | Consequents |
| K22 P72 P80 | A23 | K86 R96 S87 T90 | N74 T80 U79 | N88 | K86 |
| R84 U75 U77 | A79 | D75 R84 | A23 N90 T71 | N89 | D93 K86 L91 P76 R96 T93 |
| H84 P98 R85 | A90 | K86 L91 L95 P76 T93 | N74 N89 N92 | N90 | D93 K86 L91 N89 P76 R96 T93 |
| B73 B74 B90 | B72 |  | A23 N85 N90 | N92 | D93 K86 L91 P76 R96 |
| A90 B74 U77 | B73 | K86 | B90 P28 Y78 | P15 | K86 P17 P76 T90 T93 |
| A79 B73 N74 | B74 | K78 | A90 N90 P15 | P17 | P76 R95 R96 T90 T93 |
| B72 N92 P98 | B78 | K86 L84 T90 | L90 N88 N89 | P28 | K22 K93 L84 L89 L90 L91 P15 P72 P76 P79 S87 T93 |
| B78 B90 U79 | B83 | K86 | K90 N87 P71 | P70 |  |
| B72 U85 X75 | B90 | K86 P15 P76 S87 S88 T90 | K91 N87 P70 | P71 | P70 P76 P98 |
| A79 B72 H83 | D74 | A79 D75 D77 K75 | N86 P28 P98 | P72 | P76 P98 |
| A79 D74 K72 | D75 | T90 | B78 P72 P80 | P73 | P76 T93 |
| A79 B73 D74 | D76 | A79 | F74 F94 P79 | P74 | D93 P76 T93 |
| D74 D76 R85 | D77 | K86 T90 | P74 T80 Y78 | P75 | D93 K86 P76 T93 |
| N89 P75 T80 | D93 | K86 L91 P76 R96 S87 T93 | N85 P74 P75 | P76 | D93 K86 T93 |
| H86 K93 U77 | F74 | L90 S77 | P73 P76 P80 | P77 | P76 |
| T80 T89 U88 | F83 | K86 T90 T93 | N85 P28 P74 | P79 | D93 K86 P76 T93 |
| B78 F83 H84 | F84 | K78 K86 | A90 P73 P77 | P80 | K86 P76 T93 |
| F83 K87 P98 | F94 | K77 K86 L91 P76 S87 T90 | N88 P72 P80 | P85 | K86 S87 T90 T93 |
| N87 P72 P77 | H75 | L91 R96 T93 | P71 P72 P73 | P98 | P76 |
| A23 H75 H86 | H83 | H86 K86 L84 L90 L91 P76 S87 T93 | B74 N86 R95 | R79 | K86 P76 R95 R96 T93 |
| F74 H75 H83 | H84 | K86 | A79 K72 L71 | R84 | A79 R95 T90 |
| H83 H84 L71 | H86 | K86 L91 T93 | D76 L71 Y78 | R85 |  |
| H83 P28 P85 | K22 | K74 K86 T90 | K82 P17 R79 | R95 | K86 R96 T90 T93 |
| A79 D76 X76 | K72 | A79 D75 K83 K92 K93 R84 S97 X76 | H75 L71 P17 R79 | R96 | K86 L91 T90 T93 |
| K22 K75 K76 | K74 | K75 K76 K78 K86 T93 | F74 R85 U79 | S77 | K86 |
| D74 D77 K74 | K75 | K76 K77 K86 K90 T90 T93 | H83 P28 P85 | S87 | K86 L91 P76 S88 |
| K74 K75 U79 | K76 | K77 K78 K83 K86 K92 T90 T93 | B90 H83 S87 | S88 | K86 L91 S87 T93 |
| F94 K75 K76 | K77 | K78 K83 | S87 S88 U85 | S91 | K86 L91 T90 T93 |
| F84 K76 K77 | K78 | K83 K86 K90 | K72 P70 Y78 | S97 | T90 |
| B73 K77 K78 | K82 | K77 K83 R95 | H83 L71 N74 | T71 | K86 L91 N89 P76 R96 X76 |
| K77 K78 K82 | K83 | K86 | L88 P17 T85 | T80 | D93 K92 L91 T86 T90 T99 |
| A23 B90 F94 | K86 | L91 T90 T93 | B78 T71 T80 | T85 | K86 T86 T90 T93 |
| P80 U85 U88 | K87 | K86 K92 T93 | H75 T80 T85 | T86 | T90 T93 T99 |
| K75 K78 U79 | K90 | K86 T90 | A23 A90 F83 | T89 | K86 T90 T93 |
| F74 K76 K90 | K91 | T90 | D75 K22 T89 | T90 | K86 T93 |
| K72 K87 T80 | K92 | T93 | F74 L91 U79 | T92 | K86 L91 T90 T93 |
| Antecedants | **Disease** | **Consequents** | **Antecedants** | **Disease** | **Consequents** |
| K72 P28 U79 | K93 | K94 | K86 N70 T90 | T93 | K86 T90 |
| K93 U79 Y78 | K94 | K86 L91 | B78 T80 T86 | T99 | K86 L91 |
| A79 D75 D76 | L71 | K86 L91 R84 R96 | H83 N87 X77 | U04 | K86 L91 |
| B78 H83 P28 | L84 | K86 L85 L89 L91 L95 T93 | D74 T71 U79 | U75 | A79 |
| F74 L84 N90 | L85 | K86 L89 L91 L95 P76 T93 | U75 U77 U79 | U76 | K86 Y77 |
| H83 L91 N85 | L88 | K86 L91 T90 | U75 U76 U79 | U77 | A79 U76 |
| L84 L85 P28 | L89 | K86 L90 L91 | B73 K91 U75 | U79 | U75 |
| F74 L89 P28 | L90 | K86 L91 | K75 L85 U75 | U85 | K86 L91 T90 |
| H75 L88 N70 | L91 | K86 T93 | A90 B74 K87 | U88 | K86 T90 T93 |
| A90 L84 L85 | L95 | K86 L91 T93 | N86 R84 T71 | X75 | K86 P76 X76 |
| F74 L85 X75 | L98 | K86 L91 P76 T93 | K72 T71 X75 | X76 | K86 |
|  | N70 | L91 T93 | L71 X75 X76 | X77 | A79 K86 |
| A23 D77 U79 | N74 | N88 T90 | U75 U76 Y85 | Y77 | K86 |
| H83 L85 P85 | N85 | K86 L91 P76 S87 | B72 P77 U88 | Y78 | A79 D93 K94 P15 R96 T90 |
| K87 K91 L71 N74 | N86 | K86 L91 P76 R96 | L71 U76 U79 | Y85 | K86 |
| K91 L71 P72 | N87 | P70 T90 |  |  |  |

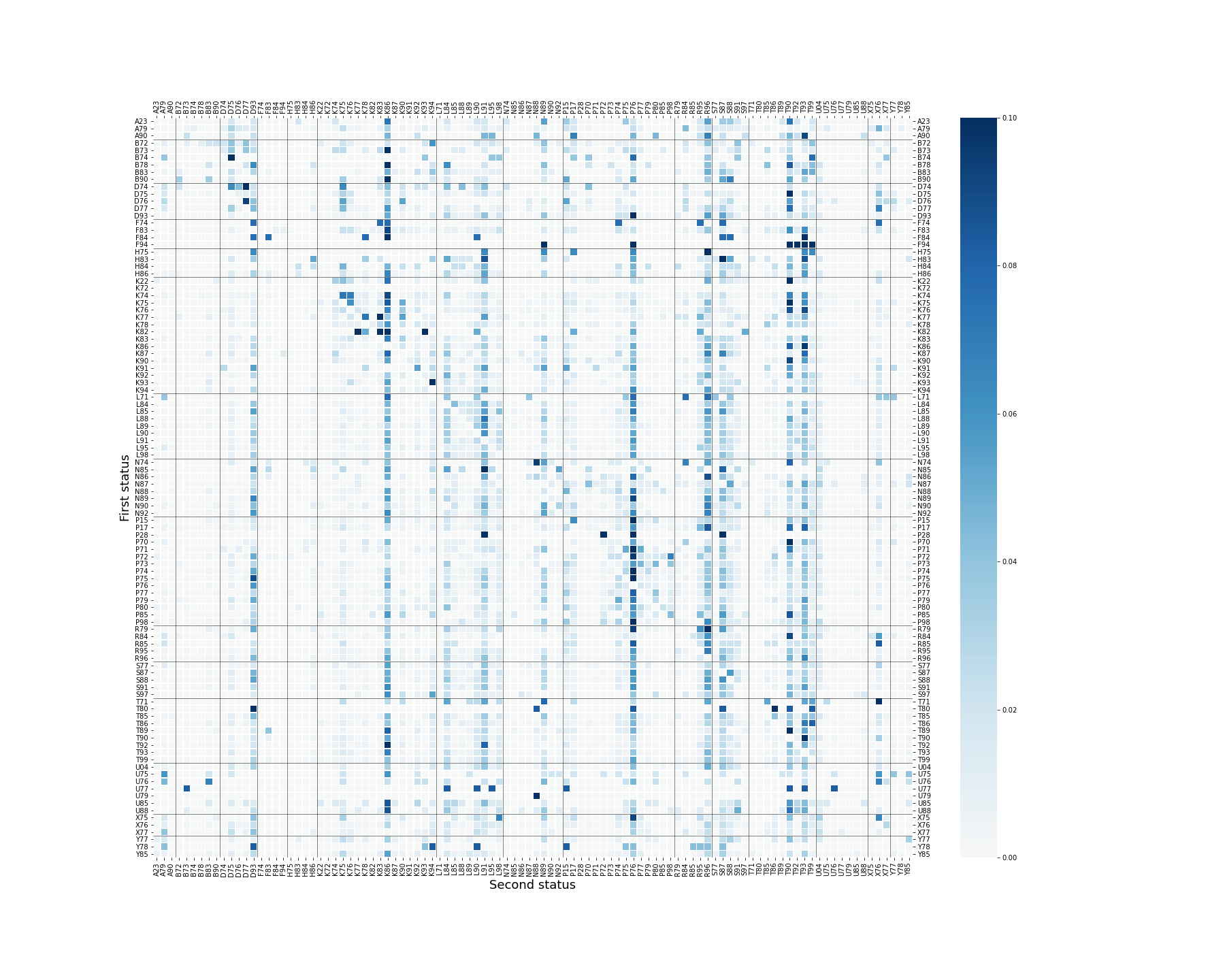
# **eFigure1. Subgroup analysis – Markov chain analysis**

****

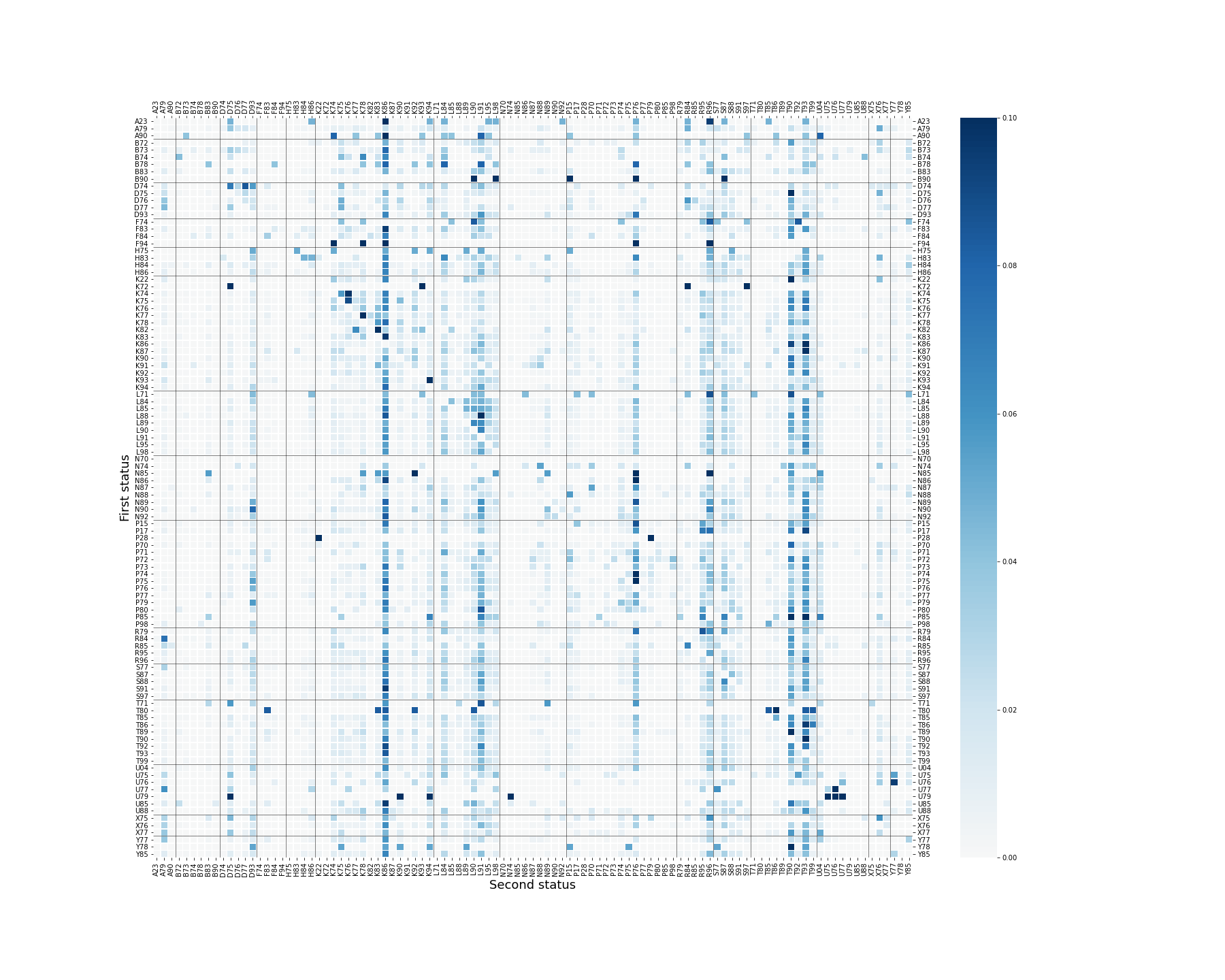
## Sex group Male

****

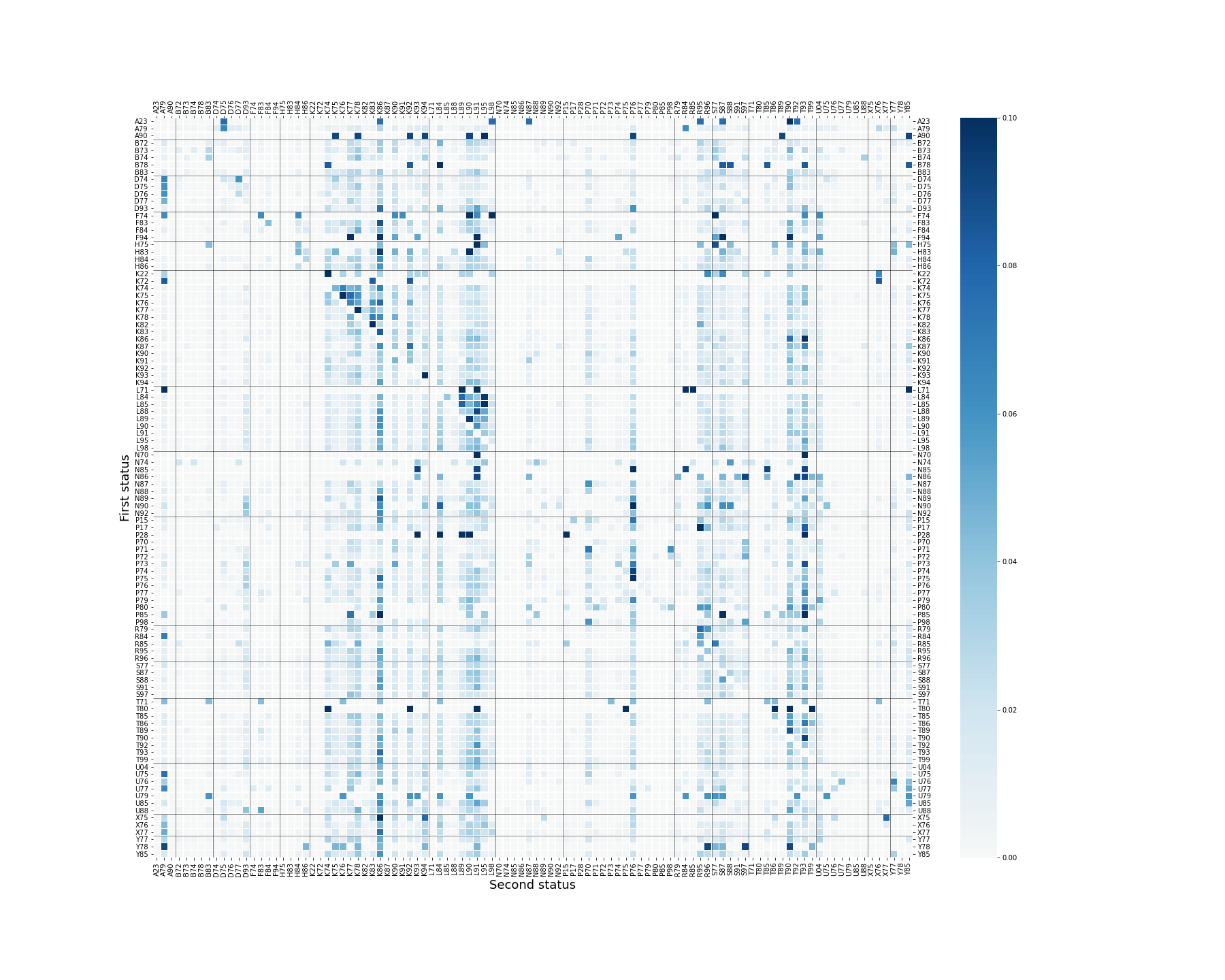
## Sex group Female



## Age group 40-59 years old



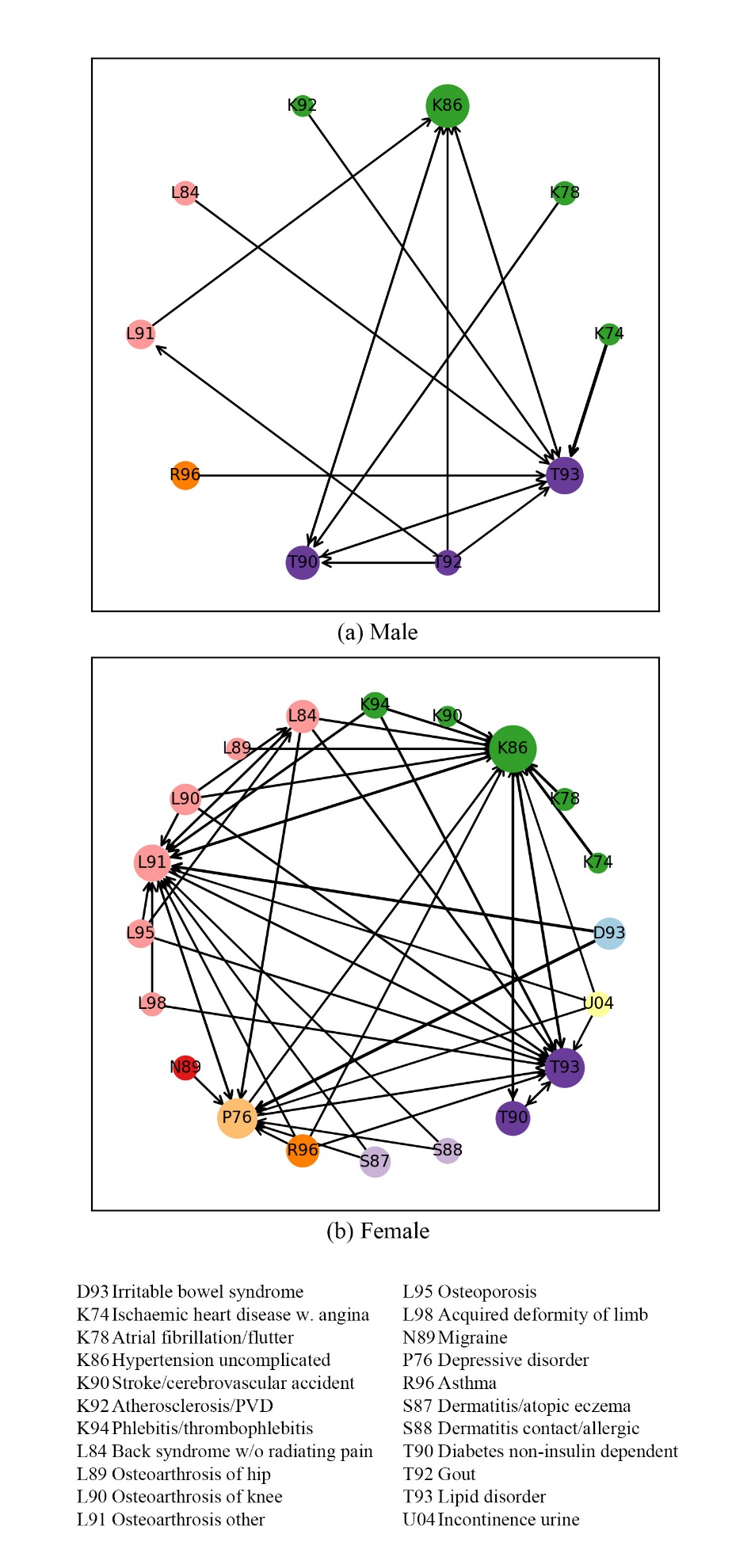
## Age group 60-74 years old



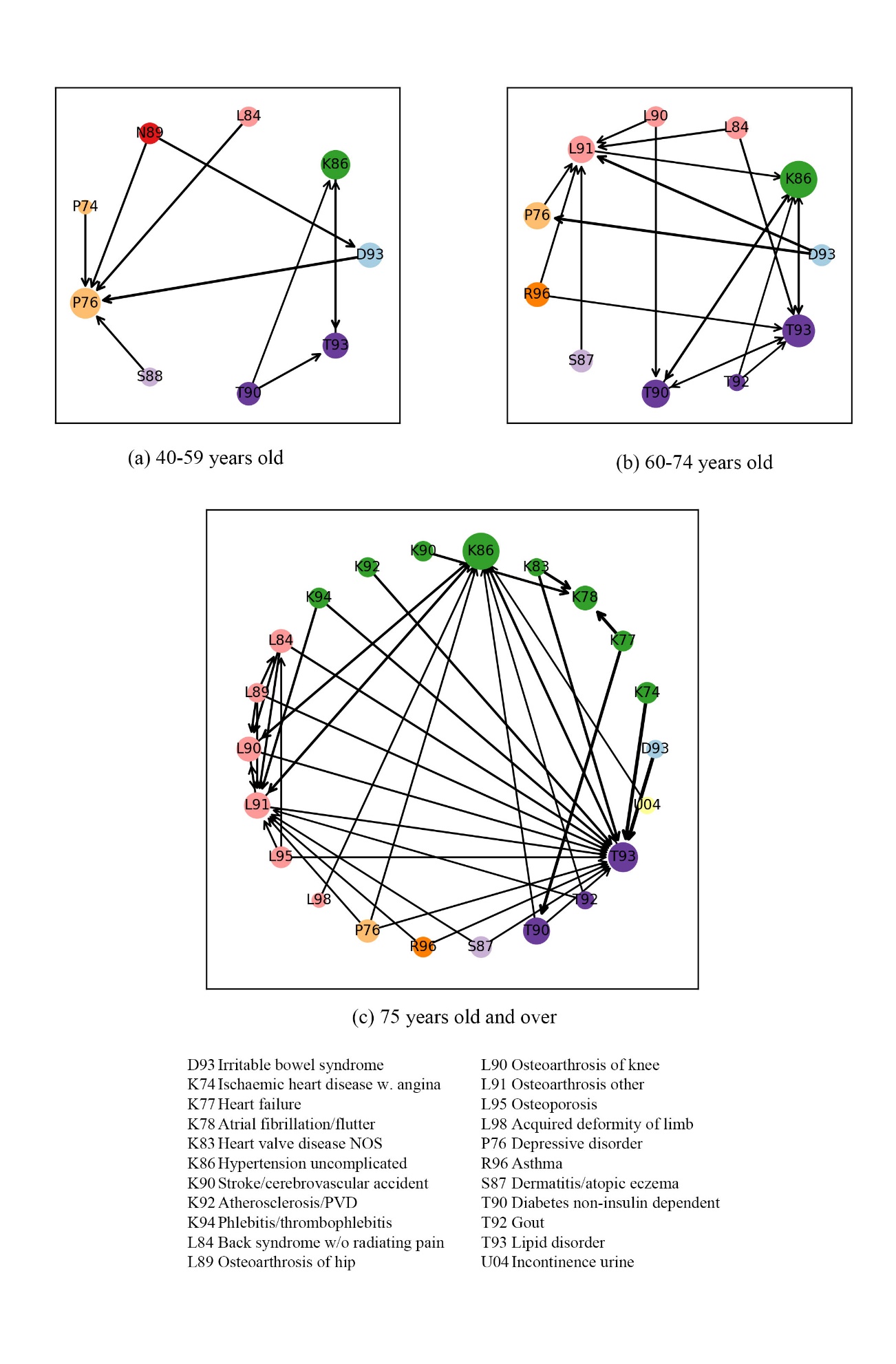
## Age group 75 years and over

**Figure 1.** Heatmap of the first-order transition probability matrix of a Markov chain by sex group and age group. The number in each square is the probability of developing the condition on the x-axis after the diagnosis of the condition on the y-axis. The scale of the color bar was defined between 0 and 0.1. Vertical and horizontal lines were added for disease groups. (Disease groups: A-General, B-Blood and Immune, D-Digestive, F-Eye, H-Ear, K-Cardiovascular, L-Musculoskeletal, N-Neurological, P-Psychological, R-Respiratory, S-Skin, T-Metabolic, U-Urological, X-Female Genital, Y-Male Genital)

# **eFigure2. Subgroup analysis – WARM**



eFigure2.1 Sex group



## eFigure2.2 Age group

**eFigure 2.** Visualization of Weighted Association Rules Mining results. The color of the nodes stands for the disease group and the size of the nodes is determined by the frequency of the occurrence of the condition. The arrows on the edge show the direction of the relation and wider edges imply that the associations were closer and stronger. Support>0.04, confidence>0.3, lift>1

# **Reference**

1. Gagniuc PA. Markov Chains: From Theory to Implementation and Experimentation. USA, NJ: John Wiley & Sons; 2017; 1–235.
2. Meyn S, Tweedie RL. Markov Chains and Stochastic Stability. Cambridge University Press; 2009; 3. ISBN 978-0-521-73182-9.
3. Sonnenberg FA, Beck RJ. Markov Models in medical decision making: a practical guide. *Medical Decision Making*. 1993; 13:322–338.
4. Zemedikun DT, Gray LJ, Khunti K, Davies MJ, Dhalwani NN. Patterns of Multimorbidity in Middle-Aged and Older Adults: An Analysis of the UK Biobank Data. *Mayo Clinic Proceedings*. 2018; 93 (7): 857-866. https://doi.org/10.1016/j.mayocp.2018.02.012
5. Held FP, Blyth F, Gnjidic D, et al. Association Rules Analysis of Comorbidity and Multimorbidity: The Concord Health and Aging in Men Project. *The Journals of Gerontology: Series A*. 2016; 71(5): 625–631. https://doi.org/10.1093/gerona/glv181
6. Wang K, Su MY. Item Selection by Hub-Authority Profit Ranking. *Proc. ACM SIGKDD*. 2002. https://doi.org/10.1145/775047.775144
7. Sun K, Bai F. Mining Weighted Association Rules without Preassigned Weights. *IEEE Transactions on Knowledge and Data Engineering.* 2008; 4 (30): 489–495. https://doi.org/10.1109/TKDE.2007.190723
8. Zaki MJ. Scalable algorithms for association mining. *IEEE Transactions on Knowledge and Data Engineering*. 2000; 12 (3): 372–390. https://doi.org/10.1109/69.846291
9. Pei J, Han J, Mao R. CLOSET: An Efficient Algorithm for Mining Frequent Closed Itemsets. *ACM SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery (DMKD 2000*). 2000.