**Lab 8: Bias in Algorithms**

*Methods/concepts: algorithmic bias; choice of “labels” vs. “predictors”*

**LAB DESCRIPTION**

In this lab, we will dive deeper into bias in algorithms, following [Obermeyer, Powers, Vogeli, and Mullainathan (2019](https://science-sciencemag-org.ezp-prod1.hul.harvard.edu/content/366/6464/447)). We will train several prediction algorithms, some including the patient’s race and others explicitly leaving out the patient’s race. We will see how the choice of “label” – either patient costs, patient health, or patient avoidable costs – affects the performance of the models. Finally, we will examine the racial composition of patients predicted to have high risk according to the algorithms.

**QUESTIONS**

1. Start by randomly splitting the 48,748 patients included in the **health.dta** data set into a **10%** training data set and a **90%** test data set. [Table 1](#table1) describes the data. There are two reasons we are using such a small fraction of the data to train the models. First, estimating random forests on a larger fraction of the data would be prohibitively time consuming. Second, we require a large number of observations in the test data set so that we can study differences in risk score by race.
2. Estimate the following statistical models using the training data set:
3. Random forest to predict the “label” of patient costs (cost\_t) using the full set of predictors consisting of all variables starting with tm1\_, but *excluding* the patient’s race
4. Random forest to predict the “label” of patient costs (cost\_t) using the full predictor set, now *including* the patient’s race
5. Random forest to predict the “label” of patient health (gagne\_sum\_t) using the full predictor set, *excluding* the patient’s race
6. Random forest to predict the “label” of patient health (gagne\_sum\_t) using the full predictor set, *including* the patient’s race
7. *Optional - Random forest to predict the “label” of avoidable patient costs (cost\_avoidable\_t) using the full predictor set, excluding the patient’s race*
8. *Optional - Random forest to predict the “label” of avoidable patient costs (cost\_avoidable\_t) using the full predictor set, including the patient’s race*

Note that random forests with lots of observations and predictors (150) will take a long time to run. You should therefore only use around 100 trees in your forests.

1. Calculate and compare the root mean squared prediction error for your models that include patient race vs. those that exclude patient race in the **training sample**.
2. Calculate and compare the root mean squared prediction error for your models that include patient race vs. those that exclude patient race in the **test sample**.
3. Export a data set with **the test data** and your predictions as a .dta file. If you are in a Stata lab, you can exit Python and load this file into Stata for further analysis.
4. As in Lab 1 and Lab 2, convert the predictions in the test sample from each of your prediction algorithms into percentile ranks, normalized so that the top rank is equal to 100. The percentile rank is the “risk score” from the algorithm.
5. Now consider a program that makes patients eligible for extra resources if their “risk score” is above the 55th percentile (i.e., the top 45 percent).
   1. Overall, what fraction of all Black patients would be eligible for the program using each of the four algorithms?
   2. Among patients eligible for the program, what fraction are Black (using each of the four algorithms)?
6. Now we will replicate the key figures from [Obermeyer, Powers, Vogeli, and Mullainathan (2019](https://science-sciencemag-org.ezp-prod1.hul.harvard.edu/content/366/6464/447)). Produce binned scatter plots of patient costs and patient health vs. the percentile rank “risk score” from each algorithm, with White and Black patients plotted separately. This is a total of 8 graphs: 4 models x 2 outcomes.

In Stata, use a connected line type in binscatter, which is controlled by the option linetype(connect). To plot Black and White patients separately, use the by(race) option:

binscatter outcome\_variable percentile\_rank, by(race) linetype(connect)

In R, you could do the same in ggplot by using the geom="line" option and geom="point" option in stat\_binmean from the statar package, and set the color option to the race variable to plot Black and White patients separately:

ggplot(dat, aes(x = percentile\_rank , y = outcome\_variable, color = race)) +

stat\_binmean(n = 20, geom = "line") +

stat\_binmean(n = 20, geom = "point")

1. In the pre-recorded video for this lab, Professor Ziad Obermeyer said that it is the left-hand side variable (i.e., the “label” or target parameter) that is the source of bias in algorithms, not the right-hand side variables (i.e., the predictors). Explain what he meant, and evaluate whether you agree with him using your binned scatters above.
2. For this Lab, please submit the following:
3. Your final do-file and Python Jupyter notebook; or .R script file to Gradescope
4. A single PDF document with the answers and graphs submitted to Gradescope. You can print the Python Jupyter notebook as a PDF using the file drop down menu.
5. There will also be a Google form that is projected to the screen in Lab

**DATA DESCRIPTION, FILE: health.dta**

The data consist of 48,784 patient records. Variables that start with tm1\_ were measured in the prior year (time ). Variable that end with \_t are measured in the current year. For more details on the construction of the variables included in this data set, please see [Obermeyer, Powers, Vogeli, and Mullainathan (2019](https://science-sciencemag-org.ezp-prod1.hul.harvard.edu/content/366/6464/447)).

TABLE 1

Variable Definitions

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |
| *Variable* | Description | mean | sd | min | max |
| *(1)* | (2) | (3) | (4) | (5) | (6) |
|  |  |  |  |  |  |
| ***patient\_id*** | Patient identification number | n/a | n/a | n/a | n/a |
| ***gagne\_sum\_t*** | Total number of active chronic illnesses | 1.354 | 1.942 | 0 | 17 |
| ***cost\_t*** | Total medical expenditures, rounded to the nearest 100 | 7,660 | 17,990 | 0 | 550,500 |
| ***cost\_avoidable\_t*** | Total avoidable (emergency + inpatient) medical expenditures, rounded to nearest | 2,435 | 12,058 | 0 | 642,700 |
| ***race*** | String variable containing the words “black” and “white” | n/a | n/a | n/a | n/a |
| ***tm1\_dem\_black*** | 1 = Black  0 = White | 0.114 | 0.318 | 0 | 1 |
| ***tm1\_dem\_female*** | 1 = Female  0 = Male | 0.631 | 0.483 | 0 | 1 |
| ***tm1\_dem\_age\_band\_1824*** | Indicator for patient age between 18-24 | 0.0369 | 0.188 | 0 | 1 |
| ***tm1\_dem\_age\_band\_2534*** | Indicator for patient age between 25-34 | 0.110 | 0.313 | 0 | 1 |
| ***tm1\_dem\_age\_band\_3544*** | Indicator for patient age between 35-44 | 0.194 | 0.396 | 0 | 1 |
| ***tm1\_dem\_age\_band\_4554*** | Indicator for patient age between 45-54 | 0.239 | 0.427 | 0 | 1 |
| ***tm1\_dem\_age\_band\_5564*** | Indicator for patient age between 55-64 | 0.197 | 0.397 | 0 | 1 |
| ***tm1\_dem\_age\_band\_6574*** | Indicator for patient age between 65-74 | 0.142 | 0.349 | 0 | 1 |
| ***tm1\_dem\_age\_band\_75*** | Indicator for patient age 75+ | 0.0703 | 0.256 | 0 | 1 |
| ***tm1\_alcohol\_elixhauser*** | Indicator for alcohol abuse | 0.00892 | 0.0940 | 0 | 1 |
| ***tm1\_anemia\_elixhauser*** | Indicator for deficiency anemia | 0.0636 | 0.244 | 0 | 1 |
| ***tm1\_arrhythmia\_elixhauser*** | Indicator for arrhythmia | 0.0922 | 0.289 | 0 | 1 |
| ***tm1\_arthritis\_elixhauser*** | Indicator for arthritis | 0.0466 | 0.211 | 0 | 1 |
| ***tm1\_bloodlossanemia\_elixhauser*** | Indicator for blood loss anemia | 0.00246 | 0.0495 | 0 | 1 |
| ***tm1\_coagulopathy\_elixhauser*** | Indicator for coagulopathy | 0.0115 | 0.107 | 0 | 1 |
| ***tm1\_compdiabetes\_elixhauser*** | Indicator for diabetes, complicated | 0.0217 | 0.146 | 0 | 1 |
| ***tm1\_depression\_elixhauser*** | Indicator for depression | 0.0621 | 0.241 | 0 | 1 |
| ***tm1\_drugabuse\_elixhauser*** | Indicator for drug abuse | 0.00623 | 0.0787 | 0 | 1 |
| ***tm1\_electrolytes\_elixhauser*** | Indicator for electrolyte disorder | 0.0329 | 0.178 | 0 | 1 |
| ***tm1\_hypertension\_elixhauser*** | Indicator for hypertension | 0.332 | 0.471 | 0 | 1 |
| ***tm1\_hypothyroid\_elixhauser*** | Indicator for hypothyroid | 0.0938 | 0.292 | 0 | 1 |
| ***tm1\_liver\_elixhauser*** | Indicator for liver disease | 0.0159 | 0.125 | 0 | 1 |
| ***tm1\_neurodegen\_elixhauser*** | Indicator for neurodegenerative disease | 0.0280 | 0.165 | 0 | 1 |
| ***tm1\_obesity\_elixhauser*** | Indicator for obesity | 0.0929 | 0.290 | 0 | 1 |
| ***tm1\_paralysis\_elixhauser*** | Indicator for paralysis | 0.000574 | 0.0240 | 0 | 1 |
| ***tm1\_psychosis\_elixhauser*** | Indicator for psychoses | 0.0325 | 0.177 | 0 | 1 |
| ***tm1\_pulmcirc\_elixhauser*** | Indicator for pulmonary circulation disorders | 0.00558 | 0.0745 | 0 | 1 |
| ***tm1\_pvd\_elixhauser*** | Indicator for peripheral vascular disorders | 0.0263 | 0.160 | 0 | 1 |
| ***tm1\_renal\_elixhauser*** | Indicator for renal failure | 0.0367 | 0.188 | 0 | 1 |
| ***tm1\_uncompdiabetes\_elixhauser*** | Indicator for diabetes, uncomplicated | 0.0987 | 0.298 | 0 | 1 |
| ***tm1\_valvulardz\_elixhauser*** | Indicator for valvular disease | 0.0315 | 0.175 | 0 | 1 |
| ***tm1\_wtloss\_elixhauser*** | Indicator for weight loss | 0.00139 | 0.0373 | 0 | 1 |
| ***tm1\_cerebrovasculardz\_romano*** | Indicator for cerebrovascular disease | 0.0283 | 0.166 | 0 | 1 |
| ***tm1\_chf\_romano*** | Indicator for congestive heart failure | 0.0319 | 0.176 | 0 | 1 |
| ***tm1\_dementia\_romano*** | Indicator for dementia | 0.00949 | 0.0970 | 0 | 1 |
| ***tm1\_hemiplegia\_romano*** | Indicator for hemiplegia | 0.00266 | 0.0516 | 0 | 1 |
| ***tm1\_hivaids\_romano*** | Indicator for HIV/AIDS | 0.00305 | 0.0552 | 0 | 1 |
| ***tm1\_metastatic\_romano*** | Indicator for metastasis | 0.00613 | 0.0780 | 0 | 1 |
| ***tm1\_myocardialinfarct\_romano*** | Indicator for myocardial infarction | 0.0169 | 0.129 | 0 | 1 |
| ***tm1\_pulmonarydz\_romano*** | Indicator for pulmonary disease | 0.102 | 0.302 | 0 | 1 |
| ***tm1\_tumor\_romano*** | Indicator for tumor | 0.0944 | 0.292 | 0 | 1 |
| ***tm1\_ulcer\_romano*** | Indicator for ulcer | 0.00480 | 0.0691 | 0 | 1 |
| ***tm1\_cost\_dialysis*** | Total costs for dialysis, rounded to nearest 10 | 26.72 | 976.6 | 0 | 63,410 |
| ***tm1\_cost\_emergency*** | Total costs for emergency, rounded to nearest 10 | 423.7 | 1,572 | 0 | 67,090 |
| ***tm1\_cost\_home\_health*** | Total costs for home health, rounded to nearest 10 | 220.5 | 1,396 | 0 | 56,830 |
| ***tm1\_cost\_ip\_medical*** | Total costs for inpatient medical, rounded to nearest 10 | 638.8 | 4,570 | 0 | 282,300 |
| ***tm1\_cost\_ip\_surgical*** | Total costs for inpatient surgical, rounded to nearest 10 | 978.5 | 6,575 | 0 | 279,930 |
| ***tm1\_cost\_laboratory*** | Total costs for laboratory, rounded to nearest 10 | 330.9 | 949.4 | -490 | 62,720 |
| ***tm1\_cost\_op\_primary\_care*** | Total costs for outpatient primary care, rounded to nearest 10 | 473.9 | 1,872 | 0 | 240,290 |
| ***tm1\_cost\_op\_specialists*** | Total costs for outpatient specialists, rounded to nearest 10 | 866.2 | 1,546 | 0 | 41,720 |
| ***tm1\_cost\_op\_surgery*** | Total costs for outpatient surgery, rounded to nearest 10 | 846.6 | 2,659 | 0 | 75,790 |
| ***tm1\_cost\_other*** | Total other costs, rounded to nearest 100 | 1,569 | 4,639 | 0 | 193,200 |
| ***tm1\_cost\_pharmacy*** | Total costs for pharmacy, rounded to nearest 10 | 342.5 | 3,995 | -10 | 153,250 |
| ***tm1\_cost\_physical\_therapy*** | Total costs for physical therapy, rounded to nearest 10 | 167.2 | 534.0 | 0 | 10,240 |
| ***tm1\_cost\_radiology*** | Total costs for radiology, rounded to nearest 10 | 241.1 | 580.8 | 0 | 20,710 |
| ***tm1\_lasix\_dose\_count*** | Number of Lasix doses | 0.0182 | 0.228 | 0 | 9 |
| ***tm1\_lasix\_min\_daily\_dose*** | Minimum daily dose of Lasix | 0.353 | 4.370 | 0 | 200 |
| ***tm1\_lasix\_mean\_daily\_dose*** | Mean daily dose of Lasix | 0.378 | 4.535 | 0 | 160 |
| ***tm1\_lasix\_max\_daily\_dose*** | Maximum daily dose of Lasix | 0.418 | 5.247 | 0 | 200 |
| ***tm1\_cre\_tests*** | Number of c-reatinine tests | 1.237 | 3.396 | 0 | 166 |
| ***tm1\_crp\_tests*** | Number of c-reactive protein tests | 0.000471 | 0.0226 | 0 | 2 |
| ***tm1\_esr\_tests*** | Number of erythrocyte sedimentation rate tests | 0.113 | 0.538 | 0 | 13 |
| ***tm1\_ghba1c\_tests*** | Number of GHbA1c tests | 0.385 | 0.748 | 0 | 9 |
| ***tm1\_hct\_tests*** | Number of hematocrit tests | 1.089 | 3.140 | 0 | 164 |
| ***tm1\_ldl\_tests*** | Number of LDL tests | 0.520 | 0.701 | 0 | 10 |
| ***tm1\_nt\_bnp\_tests*** | Number of BNP tests | 0.0305 | 0.257 | 0 | 10 |
| ***tm1\_sodium\_tests*** | Number of sodium tests | 1.156 | 3.237 | 0 | 122 |
| ***tm1\_trig\_tests*** | Number of triglycerides tests | 0.483 | 0.681 | 0 | 12 |
| ***tm1\_cre\_minlow*** | Indicator for low (< 0.84) minimum creatinine test result | 0.222 | 0.416 | 0 | 1 |
| ***tm1\_cre\_minhigh*** | Indicator for high (> 1.21) minimum creatinine test result | 0.0391 | 0.194 | 0 | 1 |
| ***tm1\_cre\_minnormal*** | Indicator for normal minimum creatinine test result | 0.236 | 0.424 | 0 | 1 |
| ***tm1\_cre\_meanlow*** | Indicator for low (< 0.84) mean creatinine test result | 0.200 | 0.400 | 0 | 1 |
| ***tm1\_cre\_meanhigh*** | Indicator for high (> 1.21) mean creatinine test result | 0.0512 | 0.220 | 0 | 1 |
| ***tm1\_cre\_meannormal*** | Indicator for normal mean creatinine test result | 0.245 | 0.430 | 0 | 1 |
| ***tm1\_cre\_maxlow*** | Indicator for low (< 0.84) maximum creatinine test result | 0.178 | 0.383 | 0 | 1 |
| ***tm1\_cre\_maxhigh*** | Indicator for high (> 1.21) maximum creatinine test result | 0.0674 | 0.251 | 0 | 1 |
| ***tm1\_cre\_maxnormal*** | Indicator for normal maximum creatinine test result | 0.252 | 0.434 | 0 | 1 |
| ***tm1\_crp\_minlow*** | Indicator for low (< 1) minimum c-reactive protein test result | 0.000164 | 0.0128 | 0 | 1 |
| ***tm1\_crp\_minhigh*** | Indicator for high (> 3) minimum c-reactive protein test result | 0.000164 | 0.0128 | 0 | 1 |
| ***tm1\_crp\_minnormal*** | Indicator for normal minimum c-reactive protein test result | 6.15e-05 | 0.00784 | 0 | 1 |
| ***tm1\_crp\_meanlow*** | Indicator for low (< 1) mean c-reactive protein test result | 0.000164 | 0.0128 | 0 | 1 |
| ***tm1\_crp\_meanhigh*** | Indicator for high (> 3) mean c-reactive protein test result | 0.000164 | 0.0128 | 0 | 1 |
| ***tm1\_crp\_meannormal*** | Indicator for normal mean c-reactive protein test result | 6.15e-05 | 0.00784 | 0 | 1 |
| ***tm1\_crp\_maxlow*** | Indicator for low (< 1) maximum c-reactive protein test result | 0.000164 | 0.0128 | 0 | 1 |
| ***tm1\_crp\_maxhigh*** | Indicator for high (> 3) maximum c-reactive protein test result | 0.000164 | 0.0128 | 0 | 1 |
| ***tm1\_crp\_maxnormal*** | Indicator for normal maximum c-reactive protein test result | 6.15e-05 | 0.00784 | 0 | 1 |
| ***tm1\_esr\_minlow*** | Indicator for low (< 1) minimum erythrocyte sedimentation rate test result | 0 | 0 | 0 | 0 |
| ***tm1\_esr\_minhigh*** | Indicator for high (> 20) minimum erythrocyte sedimentation rate test result | 0.0218 | 0.146 | 0 | 1 |
| ***tm1\_esr\_minnormal*** | Indicator for normal minimum erythrocyte sedimentation rate test result | 0.0514 | 0.221 | 0 | 1 |
| ***tm1\_esr\_meanlow*** | Indicator for low (< 1) mean erythrocyte sedimentation rate test result | 0 | 0 | 0 | 0 |
| ***tm1\_esr\_meanhigh*** | Indicator for high (> 20) mean erythrocyte sedimentation rate test result | 0.0245 | 0.155 | 0 | 1 |
| ***tm1\_esr\_meannormal*** | Indicator for normal mean erythrocyte sedimentation rate test result | 0.0487 | 0.215 | 0 | 1 |
| ***tm1\_esr\_maxlow*** | Indicator for low (< 1) maximum erythrocyte sedimentation rate test result | 0 | 0 | 0 | 0 |
| ***tm1\_esr\_maxhigh*** | Indicator for high (> 20) maximum erythrocyte sedimentation rate test result | 0.0265 | 0.161 | 0 | 1 |
| ***tm1\_esr\_maxnormal*** | Indicator for normal maximum erythrocyte sedimentation rate test result | 0.0470 | 0.212 | 0 | 1 |
| ***tm1\_ghba1c\_minlow*** | Indicator for low (< 4) minimum GHbA1c test result | 4.10e-05 | 0.00640 | 0 | 1 |
| ***tm1\_ghba1c\_minhigh*** | Indicator for high (> 5.7) minimum GHbA1c test result | 0.123 | 0.329 | 0 | 1 |
| ***tm1\_ghba1c\_minnormal*** | Indicator for normal minimum GHbA1c test result | 0.146 | 0.353 | 0 | 1 |
| ***tm1\_ghba1c\_meanlow*** | Indicator for low (< 4) mean GHbA1c test result | 4.10e-05 | 0.00640 | 0 | 1 |
| ***tm1\_ghba1c\_meanhigh*** | Indicator for high (> 5.7) mean GHbA1c test result | 0.130 | 0.336 | 0 | 1 |
| ***tm1\_ghba1c\_meannormal*** | Indicator for normal mean GHbA1c test result | 0.140 | 0.347 | 0 | 1 |
| ***tm1\_ghba1c\_maxlow*** | Indicator for low (< 4) maximum GHbA1c test result | 4.10e-05 | 0.00640 | 0 | 1 |
| ***tm1\_ghba1c\_maxhigh*** | Indicator for high (> 5.7) maximum GHbA1c test result | 0.133 | 0.339 | 0 | 1 |
| ***tm1\_ghba1c\_maxnormal*** | Indicator for normal maximum GHbA1c test result | 0.137 | 0.344 | 0 | 1 |
| ***tm1\_hct\_minlow*** | Indicator for low (< 35.5) minimum hematocrit test result | 0.0639 | 0.245 | 0 | 1 |
| ***tm1\_hct\_minhigh*** | Indicator for high (> 48.6) minimum hematocrit test result | 0.00679 | 0.0821 | 0 | 1 |
| ***tm1\_hct\_minnormal*** | Indicator for normal minimum hematocrit test result | 0.375 | 0.484 | 0 | 1 |
| ***tm1\_hct\_meanlow*** | Indicator for low (< 35.5) mean hematocrit test result | 0.0424 | 0.202 | 0 | 1 |
| ***tm1\_hct\_meanhigh*** | Indicator for high (> 48.6) mean hematocrit test result | 0.00787 | 0.0884 | 0 | 1 |
| ***tm1\_hct\_meannormal*** | Indicator for normal mean hematocrit test result | 0.396 | 0.489 | 0 | 1 |
| ***tm1\_hct\_maxlow*** | Indicator for low (< 35.5) maximum hematocrit test result | 0.0242 | 0.154 | 0 | 1 |
| ***tm1\_hct\_maxhigh*** | Indicator for high (> 48.6) maximum hematocrit test result | 0.0119 | 0.109 | 0 | 1 |
| ***tm1\_hct\_maxnormal*** | Indicator for normal maximum hematocrit test result | 0.410 | 0.492 | 0 | 1 |
| ***tm1\_ldl\_minlow*** | Indicator for low (< 50) minimum LDL test result | 0.0155 | 0.124 | 0 | 1 |
| ***tm1\_ldl\_minhigh*** | Indicator for high (> 99) minimum LDL test result | 0.204 | 0.403 | 0 | 1 |
| ***tm1\_ldl\_minnormal*** | Indicator for normal minimum LDL test result | 0.198 | 0.398 | 0 | 1 |
| ***tm1\_ldlmeanlow*** | Indicator for low (< 50) mean LDL test result | 0.0127 | 0.112 | 0 | 1 |
| ***tm1\_ldlmeanhigh*** | Indicator for high (> 99) mean LDL test result | 0.211 | 0.408 | 0 | 1 |
| ***tm1\_ldlmeannormal*** | Indicator for normal mean LDL test result | 0.134 | 0.340 | 0 | 1 |
| ***tm1\_ldl\_maxlow*** | Indicator for low (< 50) maximum LDL test result | 0.0117 | 0.108 | 0 | 1 |
| ***tm1\_ldl\_maxhigh*** | Indicator for high (> 99) maximum LDL test result | 0.218 | 0.413 | 0 | 1 |
| ***tm1\_ldl\_maxnormal*** | Indicator for normal maximum LDL test result | 0.127 | 0.333 | 0 | 1 |
| ***tm1\_nt\_bnp\_minlow*** | Indicator for low (< 100) minimum BNP test result | 0.00488 | 0.0697 | 0 | 1 |
| ***tm1\_nt\_bnp\_minhigh*** | Indicator for high (> 450) minimum BNP test result | 0.00980 | 0.0985 | 0 | 1 |
| ***tm1\_nt\_bnp\_minnormal*** | Indicator for normal minimum BNP test result | 0.00543 | 0.0735 | 0 | 1 |
| ***tm1\_nt\_bnp\_meanlow*** | Indicator for low (< 100) mean BNP test result | 0.00668 | 0.0815 | 0 | 1 |
| ***tm1\_nt\_bnp\_meanhigh*** | Indicator for high (> 450) mean BNP test result | 0.0103 | 0.101 | 0 | 1 |
| ***tm1\_nt\_bnp\_meannormal*** | Indicator for normal minimum BNP test result | 0.00344 | 0.0586 | 0 | 1 |
| ***tm1\_nt\_bnp\_maxlow*** | Indicator for low (< 100) maximum BNP test result | 0.00646 | 0.0801 | 0 | 1 |
| ***tm1\_nt\_bnp\_maxhigh*** | Indicator for high (> 450) maximum BNP test result | 0.0106 | 0.102 | 0 | 1 |
| ***tm1\_nt\_bnp\_maxnormal*** | Indicator for normal minimum BNP test result | 0.00344 | 0.0586 | 0 | 1 |
| ***tm1\_sodium\_minlow*** | Indicator for low (< 135) minimum sodium test result | 0.0403 | 0.197 | 0 | 1 |
| ***tm1\_sodium\_minhigh*** | Indicator for high (> 145) minimum sodium test result | 0.000615 | 0.0248 | 0 | 1 |
| ***tm1\_sodium\_minnormal*** | Indicator for normal minimum sodium test result | 0.438 | 0.496 | 0 | 1 |
| ***tm1\_sodium\_meanlow*** | Indicator for low (< 135) mean sodium test result | 0.0196 | 0.139 | 0 | 1 |
| ***tm1\_sodium\_meanhigh*** | Indicator for high (> 145) mean sodium test result | 0.000861 | 0.0293 | 0 | 1 |
| ***tm1\_sodium\_meannormal*** | Indicator for normal mean sodium test result | 0.459 | 0.498 | 0 | 1 |
| ***tm1\_sodium\_maxlow*** | Indicator for low (< 135) maximum sodium test result | 0.0109 | 0.104 | 0 | 1 |
| ***tm1\_sodium\_maxhigh*** | Indicator for high (> 145) maximum sodium test result | 0.00515 | 0.0715 | 0 | 1 |
| ***tm1\_sodium\_maxnormal*** | Indicator for normal maximum sodium test result | 0.464 | 0.499 | 0 | 1 |
| ***tm1\_trig\_minlow*** | Indicator for low (< 50) minimum triglycerides test result | 0.0318 | 0.176 | 0 | 1 |
| ***tm1\_trig\_minhigh*** | Indicator for high (> 150) minimum triglycerides test result | 0.0901 | 0.286 | 0 | 1 |
| ***tm1\_trig\_minnormal*** | Indicator for normal minimum triglycerides test result | 0.262 | 0.440 | 0 | 1 |
| ***tm1\_trig\_meanlow*** | Indicator for low (< 50) mean triglycerides test result | 0.0289 | 0.167 | 0 | 1 |
| ***tm1\_trig\_meanhigh*** | Indicator for high (> 150) mean triglycerides test result | 0.0972 | 0.296 | 0 | 1 |
| ***tm1\_trig\_meannormal*** | Indicator for normal mean triglycerides test result | 0.256 | 0.436 | 0 | 1 |
| ***tm1\_trig\_maxlow*** | Indicator for low (< 50) maximum triglycerides test result | 0.0279 | 0.165 | 0 | 1 |
| ***tm1\_trig\_maxhigh*** | Indicator for high (> 150) maximum triglycerides test result | 0.107 | 0.309 | 0 | 1 |
| ***tm1\_trig\_maxnormal*** | Indicator for normal maximum triglycerides test result | 0.251 | 0.434 | 0 | 1 |
| ***tm1\_gagne\_sum*** | Total number of active illnesses | 1.443 | 2.049 | 0 | 18 |
|  |  |  |  |  |  |