# UCSF

**UC San Francisco Previously Published Works**

**Title**
A Novel Metric for Developing Easy-to-Use and Accurate Clinical Prediction Models

**Permalink**
https://escholarship.org/uc/item/7x46m63k

**Journal**
Medical Care, 59(5)

**ISSN**
0025-7079

**Authors**
Lee, Sei J
Smith, Alexander K
Diaz-Ramirez, L Grisell
et al.

**Publication Date**
2021-05-01

**DOI**
10.1097/mlr.0000000000001510

Peer reviewed

# A Novel Metric for Developing Easy-To-Use and Accurate Clinical Prediction Models: The Time-Cost Information Criterion

**Sei J. Lee, MD MAS**[1,2], **Alexander K. Smith, MD MS MPH**[1,2], **L. Grisell Ramirez-Diaz, MS**[1,2], **Kenneth E. Covinsky, MD MPH**[1,2], **Siqi Gan, MPH**[1,2], **Catherine L. Chen, MD MPH**[3,4], **W. John Boscardin, PhD**[1,2]

[1]University of California, San Francisco, Department of Medicine, Division of Geriatrics, San Francisco, CA

[2]San Francisco Veterans Affairs Medical Center, San Francisco, CA

[3]University of California, San Francisco, Department of Anesthesia & Perioperative Care, San Francisco, CA

[4]Philip R. Lee Institute for Health Policy Studies at UCSF, San Francisco, CA

## Abstract

**Background:** Guidelines recommend that clinicians use clinical prediction models to estimate future risk to guide decisions. For example, predicted fracture risk is a major factor in the decision to initiate bisphosphonate medications. However, current methods for developing prediction models often lead to models that are accurate but difficult-to-use in clinical settings.

**Objective:** To develop and test whether a new metric that explicitly balances model accuracy with clinical usability leads to accurate, easier-to-use prediction models.

**Methods:** We propose a new metric called the Time-cost Information Criterion (TCIC) that will penalize potential predictor variables that take a long time to obtain in clinical settings. To demonstrate how the TCIC can be used to develop models that are easier-to-use in clinical settings, we use data from the 2000 wave of the Health and Retirement Study (n=6311) to develop and compare time to mortality prediction models using a traditional metric (Bayesian Information Criterion or BIC) and the TCIC.

**Results:** We found that the TCIC models utilized predictors that could be obtained more quickly than BIC models while achieving similar discrimination. For example, the TCIC identified a 7-predictor model with a total time-cost of 44 seconds while the BIC identified a 7-predictor model with a time-cost of 119 seconds. The Harrell's c-statistic of the TCIC and BIC 7-predictor models did not differ (0.7065 vs 0.7088, p=0.11)

**Conclusion:** Accounting for the time-costs of potential predictor variables through the use of the TCIC led to the development of an easier-to-use mortality prediction model with similar discrimination.

## INTRODUCTION

Much of modern medical decision-making hinges on predictions of future events.[1] For example, the decision to initiate treatment for osteoporosis depends in part on a woman's future fracture risk;[2,3] the decision to initiate treatment for hyperlipidemia depends in part on the predicted risk of cardiovascular events;[4] and the decision to screen for colorectal cancer is dependent on predicted life expectancy.[5,6] This focus on prediction to target interventions is conceptually appealing since most interventions have risks, burdens and/or costs. Exposing low-risk patients to these interventions may lead to overall harm rather than benefit, while high-risk patients are much more likely to benefit. Thus, prediction models hold tremendous potential as a way to individualize clinical decision making to target interventions to those patients most likely to benefit.[1]

However, many accurate prediction models are not widely adopted into clinical practice.[7,8] One reason for this may be that clinical researchers develop prediction models using large datasets (often from cohort studies) where data elements have already been collected. For researchers, since the data has already been collected, all data elements are equally easy to incorporate into prediction models. In contrast, providers trying to use prediction models in clinical practice must obtain the value of each data element for each patient, before inputting the information into prediction models, often using online calculators such as http://www.cvriskcalculator.com/ or https://eprognosis.ucsf.edu/. Thus, for the end-users of prediction models, some data elements are easy to obtain (e.g. age), while other data elements are difficult or take a long time to obtain (e.g. delayed word recall). To develop prediction models that are easier-to-use in clinical practice, model development methods should account for the data collection burdens of clinical users and favor data elements that are easier to obtain in clinical settings.

In this clinical research paper, we will provide a brief overview of current prediction model development methods, highlighting the role that the Bayesian Information Criterion (BIC) commonly plays in model development. Then we will describe how the BIC can be modified to account for the time it would take a clinician to obtain a potential predictor variable for a patient, resulting in a modified metric we have named the Time-Cost Information Criterion (TCIC). We will use data from the 2000 wave of Health and Retirement Study (HRS)[9] to illustrate how the TCIC leads to the development of models that are as accurate but take substantially less time for clinicians to use than models developed using BIC.

## BACKGROUND:

### BIC AND CURRENT MODEL DEVELOPMENT METHODS

Information criterion (IC) are often used in prediction model development to balance maximizing model fit with minimizing model complexity[10]. IC is a metric that seeks to gauge whether an additional factor improves prediction enough to justify its inclusion in the prediction model. The most commonly used IC is the Bayesian IC (BIC)[11] which is defined as:

$$BIC = -2\log L + k\log n \qquad (\textit{lower is better})$$

where the log likelihood (log $L$) is the measure of model fit, decreasing in magnitude with better fitting models (and correlates strongly with improved predictive accuracy), k is the degrees of freedom of a model (and correlates with the number of predictor variables and thus model complexity) and n is the sample size.

Model selection with BIC generally leads to the development of a prediction model that balances maximizing model fit while minimizing model complexity by identifying a much more parsimonious set of predictor variables than would be identified by AIC or likelihood ratio testing.[12,13,14] For example, the Lee Index combined BIC with backward stepwise selection to identify a 12 predictor model from a starting set of 41 potential predictor variables.[15] In moving from a 41 variable model to 40 variable model, the decrease in model fit (as measured by the log likelihood) is trivial since the smaller model still has many predictors. However, the complexity penalty is moderate, resulting in a decrease in the magnitude of the BIC, favoring the smaller (40 predictor) model. In contrast, when moving from a 5 variable model to a 4 variable model, the decrease in model fit is substantial while the complexity penalty remains moderate. Since the decrease in model fit is larger than the complexity penalty, the BIC would favor the larger (5 predictor) model. Between the 40 and 5 predictor model, the BIC identified the 12 predictor model that balanced model fit and model complexity.

The central problem is that traditional information criteria, including the BIC, view all variables as equivalent. However, some predictor variables with one degree of freedom are easy for clinicians to obtain, while other predictor variables with one degree of freedom are difficult for clinicians to obtain. Standard model development methods often select difficult-to-obtain predictor variables to maximize predictive accuracy when easy-to-obtain variables may have led to a model with nearly the same level of predictive accuracy.

To encourage development of prediction models that are accurate and easier to use, a clinical prediction model development methodology should distinguish between easy-to-obtain predictors and hard-to-obtain predictors, so that easy-to-obtain predictors are favored. The ideal model development methodology would include hard-to-obtain predictor variables only if the predictor increased the model fit substantially. Conversely, an easy-to-obtain predictor variable could be included even if it only improved the model fit slightly. If there are 2 predictive factors with similar predictive properties, the new ideal model development methodology should preferentially choose the predictive factor that will be easier for clinical

staff to obtain. Applied serially over the many decision points in the prediction model development process, this new prediction model development methodology should lead to a model that balances predictive accuracy and ease-of-use. In our example below, we show how the Time-Cost Information Criterion (TCIC) leads to models that require less time to use with minimal decreases in discrimination.

## METHODS:

### QUANTIFYING THE DEGREE TO WHICH A VARIABLE IS EASY- OR HARD-TO-OBTAIN

To develop a model development strategy that penalizes the inclusion of hard-to-obtain variables, we must define and operationalize what is meant by "hard-to-obtain." Previous research showing that the time required to use prediction models was a common reason for not using these tools[16] suggests that <u>time</u> may be the most important factor that makes a predictor "hard-to-obtain" in a clinical setting. For a factor (such as age) that is universally available, the "time-cost" would be minimal. For a factor (such as a scale) that requires administering a series of questions and recording answers, the time-cost would be much higher. By estimating a time-cost of each potential predictor variable, we would have the raw data that can be used to quantify how easy- or hard-to-obtain a potential predictor variable is.

## METHODS:

### DEFINITION AND CHARACTERISTICS OF THE TCIC

To incorporate time-cost into model development, we have developed the time-cost information criterion (TCIC), a modified BIC that incorporates time-costs into the complexity penalty term. We start by rewriting the standard BIC formula given above as:

$$BIC = -2\log L + M\frac{k}{M}\log n \qquad (lower\ is\ better)$$

M: degrees of freedom in full model

k: degrees of freedom in reduced model

M refers to the degrees of freedom in the full model, before any predictor selection has occurred. In contrast, k refers to the model that is being considered, which would usually be a model with fewer predictor variables. To incorporate time-costs, the TCIC modifies the BIC:

$$TCIC = -2\log L + M\frac{time\ for\ k}{time\ for\ M}\log n \qquad (lower\ is\ better)$$

Time for k: time it takes to gather data for all variables in reduced model

Time for M: time it takes to gather data for all variables in full model

Consider the scenario where a high time-cost predictor is being evaluated for whether it should be eliminated or retained in the model. For the reduced model without the high time-cost predictor, the time-cost ratio will be relatively low and the complexity term will be smaller, resulting in a lower (better) TCIC. Since lower TCIC scores are favored, the reduced model without the predictor will be favored, and the high time-cost predictor will be eliminated from the model. Conversely, consider the scenario where a low time-cost predictor is being evaluated for whether it should be eliminated or retained in the model. For the reduced model without the low time-cost predictor, the time-cost ratio will be relatively high and the complexity term will be larger, resulting in a higher (worse) TCIC. Since lower TCIC scores are favored, the reduced model without the predictor will be rejected and the low time-cost predictor will be retained in the model. Incorporating the TCIC into prediction model development should lead to the selection of predictors that improve model fit while having a relatively low time-cost.

## METHODS:

### QUANTIFYING TIME-COSTS OF POTENTIAL PREDICTOR VARIABLES

To determine the "time-cost" of obtaining each potential predictive factor, we used the automated reading function of Microsoft Excel (Microsoft Corp, Redmond, WA) and measured the time it took the automated reader to read the HRS survey question for each potential predictive factor. For example, for the variable "Cancer", we measured the time it takes the Excel reader to read aloud verbatim the HRS question: "Has a doctor ever told you that you have cancer or a malignant tumor, excluding minor skin cancer?" For a data element such as delayed recall, which requires 1) initial presentation of things to remember, 2) an intervening time period and 3) follow-up questions to ascertain how many items are recalled, we summed the time required for 1) the initial presentation and 3) follow-up questions, since the intervening time could be spent asking other questions or providing other needed care.

True clinical time-costs, incorporating patient response times and the heterogeneity across patients, may be difficult to ascertain during the clinical prediction model development. Thus, we conducted a series of sensitivity analyses varying time-costs to determine how model development and variable selection would differ with differing time-costs. First, we divided our predictors into 3 groups with low time-costs (<10 seconds via automated reader), moderate time-costs (10–30 seconds) and high time-costs (>30 seconds). Then, we assigned the same time-costs to all predictors in each category, using the mean time-cost of predictors in that category (low = 5.1 seconds, moderate = 18.6 seconds, high = 68.5 seconds). Second, using the same low/moderate/high time-cost groupings, we assigned a time-cost of 1 to low-cost predictors, a time-cost of 2 to moderate cost predictors and a time-cost of 3 to high-cost predictors. Finally, we varied the time-costs for a single predictor (delayed word recall) and varied its time-cost 1000- fold from 0.78 seconds to 780 seconds to determine how varying a single predictor's time-costs affects how it is retained or excluded during variable selection and model development.

## METHODS:

### STUDY POPULATION, PREDICTOR AND OUTCOME VARIABLES, STATISTICAL ANALYSES

To demonstrate how prediction models developed using the TCIC differs from models developed using BIC, we used data from the 2000 wave of the Health and Retirement Study (HRS) to develop and compare BIC and TCIC time to mortality prediction models. The HRS is a nationally-representative cohort of US adults over age 50.[9] We focused on community-dwelling HRS respondents over age 70 (n=6311). We identified 19 potential predictor variables that have been shown to be important risk factors for mortality including demographic factors such as age and gender, chronic conditions such as heart failure and diabetes, functional factors such as dependence in activities of daily living (ADL) and cognitive factors such as 10-word delayed recall. For each of these factors, we used the Microsoft Excel automated reading function to estimate the time it would take for clinical staff to ask this question to a patient. Our outcome of interest was time to death, ascertained through usual HRS procedures including linkage with the National Death Index.[17]

To compare the BIC and TCIC models, we used backward stepwise selection to develop BIC and TCIC Cox survival models. Specifically, we used BIC and backward selection to identify the optimal BIC 18 variable model, BIC 17 variable model, etc. We calculated the Harrell's c-statistic[18] and total time-cost of all predictor variables in the model at each step. Harrell's c-statistic is similar to the c-statistic but can accommodate censoring, making it useful in the evaluation of survival models. We identified the BIC minimum model as the overall BIC-optimal model. We followed a parallel strategy with TCIC, resulting in the optimal TCIC 18 variable model, 17 variable model, etc, calculating the Harrell's c-statistic and total time-costs for each model. As a sensitivity analysis, we re-conducted our analysis using the Akaike Information Criterion (AIC) instead of the BIC.

All analyses were performed using statistical software SAS/STAT 15.1 (SAS Institute Inc, Cary, NC, 2016). The Committee on Human Research at the University of California, San Francisco declined to review this study since it relied on de-identified data and thus did not meet criteria for human research.

## RESULTS:

The baseline characteristics of our cohort are shown in Table 1. For each of our 19 potential predictor variables, we present the time-cost of the variable and how each variable was specified. The potential predictor variable with the highest time-cost was delayed 10-word recall at 78.5 seconds. The potential predictor variable with the lowest time-cost was age at 3.0 seconds. The mortality rate at 10 years was 52%.

The discrimination (Harrell's c-statistic) and time-costs of the BIC and TCIC models are shown in Figure 1. We found minimal decreases in discrimination with decreasing model size initially, with Harrell's c-statistic decreasing from 0.7213 to 0.7199 when decreasing model size from 19 variables to 13 variables. In contrast, substantial decreases in discrimination was observed when decreasing model size from 5 to 1 variable (Harrell's c-statistic 0.6992 to 0.6489). As expected, time-costs decreased more quickly for TCIC

models (red lines) compared to BIC models (blue lines). Figure 1 shows that BIC and TCIC identified the same optimal models for some model sizes (where blue and red solid lines converge/overlap). Figure 1 also shows that at other model sizes, BIC and TCIC identified different optimal models (where blue and red solid lines diverge).

The optimal models identified through the BIC and TCIC methodologies are shown in detail in Table 2. Using the BIC methodology, the STROKE variable is removed first, with the resultant 18 predictor model having a time-cost of 360 seconds and a Harrell's c-statistic of 0.7209. Similarly, using the TCIC methodology, the 10 WORD IMMEDIATE RECALL variable is removed first, with the resultant 18 predictor model having a time-cost of 289 seconds and a Harrell's c-statistic of 0.7205. The 12 variable model was identified as optimal by both the BIC and TCIC metrics; however, while the BIC 12-predictor model included IADL DEPENDENCY, the TCIC 12-predictor model substituted EDUCATION for IADL DEPENDENCY. While the Harrell's c-statistic for the BIC and TCIC 12-variable models differed (0.7197 vs 0.7184, p=0.01), the Harrell's c-statistic for the 18 other models did not differ (p>0.11).

Varying time-costs by grouping predictors into low, moderate and high time-costs (compared to individual predictor time-costs from automated reader) did not dramatically change our results (Supplemental Table 1). For example, while DEPRESSION was second variable eliminated when using automated reader time-costs, it was the first variable eliminated when using categorized time-costs. The same 5 final predictors were retained regardless of time-costs. As the range of time-costs decreased from automated reader time-costs (range: 3 – 78 seconds), to categorized time-costs (range: 1 – 3), we found that the differences between BIC selection and TCIC selection narrowed (Supplemental Figure 1).

Varying the time-cost of a single predictor showed that increasing a predictor's time-cost led the predictor being removed earlier (Supplemental Table 2). When the 10-word delayed recall predictor time-cost was increased from 78 seconds to 780 seconds, the predictor went from being the 11th eliminated predictor to the first eliminated predictor. Conversely, when the time-cost was decreased from 78 seconds to 7.8 seconds, the predictor went from being the 11th eliminated predictor to being the 14th eliminated predictor.

Our sensitivity analysis utilizing the Akaike Information Criterion (AIC) instead of the BIC resulted in minimal differences in the resultant model. (Supplemental Table 3) While the BIC identified the 12 variable model as optimal, the AIC identified the 15 variable model as optimal. However, the identified variables (and the order of elimination of variables) were nearly identical between the AIC and BIC models.

## DISCUSSION:

We developed a new metric (Time-Cost Information Criterion or TCIC) from the BIC and demonstrated how it can be used to develop prediction models that are nearly as accurate but take less time to use. For example, the BIC-optimal 15 variable model had a Harrell's c-statistic of 0.7208 with a time-cost of 325 seconds. In contrast, the TCIC-optimal 15 variable model has a Harrell's c-statistic of 0.7202 with a time-cost of 209 seconds. Thus, the TCIC

was able to identify a model that achieved similar discrimination (0.7202 vs 0.7208, p=0.20) but took much less time (209 vs 325 seconds). In our example, use of the TCIC nearly always led to models with no statistically significant loss of discrimination, with substantial decreases in time-costs.

Our sensitivity analyses showed that our results were fairly robust to varying time-cost ascertainment methods or using AIC instead of BIC. Specifically, our results showed the grouping time costs did not dramatically change our resultant models. In addition, when the range of time-costs across potential predictors decreases, TCIC models become increasingly similar to BIC models (Supplemental Figure 1). Finally, as noted by previous investigators, [12,13,14] we found that BIC favored more parsimonious models while AIC favored larger models. However, the order of elimination of variables were nearly identical when using AIC rather than BIC.

## BEST PREDICTION IN UNDER A MINUTE: USING THE TCIC TO IDENTIFY THE OPTIMAL MODEL THAT TAKES LESS THAN A PRE-SPECIFIED AMOUNT OF TIME

For clinical researchers focused on developing prediction models that are easy-to-use in clinical settings, the TCIC allows model developers to identify the optimal model under a certain time-cost. For example, while our mortality index identified the 12 variable model as optimal, clinical researchers may want to identify the optimal model that takes less than 1 minute. In our mortality prediction example, the time-cost methodology would identify a 7 variable model that took 44 seconds and had a Harrell's c-statistic of 0.7065. In contrast, the traditional methodology would identify a 5 variable model that takes 32 seconds and had a Harrell's c-statistic of 0.6992. Thus, rather than comparing models with the same number of variables, a more clinically relevant comparison may be models with the same time-costs.

## TCIC FUTURE DIRECTIONS

The TCIC has several advantageous characteristics suggesting it could be the basis of future information criteria. First, by having a ratio of 2 time-costs, our complexity term remains unitless. Second, by focusing on comparing ratios, the TCIC should be relatively robust to small errors in time-cost ascertainment, since the errors likely affect both the numerator and denominator similarly. Third, the TCIC reduces to the BIC when all time-costs are equal.

The first future direction may be the development of other information criteria that focuses on other "costs" beyond time, such as financial cost. For example, imaging studies or laboratory tests may not impose high time-costs, but may have high financial costs to patients or society. If the goal is to develop accurate prediction models without imposing high financial costs to obtain predictor data, a financial-cost information criteria could be constructed, using a parallel method to the TCIC.

A second future direction may be to incorporate TCIC into other prediction model development methods beyond stepwise selection such as Least Absolute Shrinkage and Selection Operator (LASSO).[19] One advantage of deriving TCIC from BIC is that BIC is widely used and has been incorporated into LASSO[20] in several statistical software programs,[21–22] providing a roadmap for incorporating TCIC into LASSO.

Additional prior work shows the possibility of individual variable weights into LASSO regression to differentially penalize specific variables. Bergersen et al showed how to incorporate individual variable weights into LASSO regression to more heavily penalize the selection of poor quality variables.[23] Our approach could build on the same methodology, penalizing higher time-cost variables so that they are more heavily penalized and less likely to be selected. Additional authors have also discussed more complex weighted LASSO approaches in the bioinformatics literature,[24,25] suggesting several methodologic approaches for incorporating variable-specific time-cost weights into LASSO. Thus, while incorporating TCIC into LASSO is beyond the scope of this manuscript, prior work on BIC-LASSO and weighted LASSO suggests that TCIC could be incorporated into LASSO.

## CONCLUSION

Current prediction model development strategies focus on improving predictive accuracy. Although this is clearly a critical goal, the lack of attention to the clinical usability of prediction models have led to the development of many accurate models which are difficult to use in clinical settings. We have introduced the concept of using time-costs as a way of identifying predictors that are easier to obtain in clinical practice. By using the time-cost information criteria (TCIC) in prediction model development, we have shown that prediction models with similar discrimination but decreased time-costs can be developed. Use of the TCIC in clinical prediction model development may lead to models that are nearly as accurate and easier to use in routine clinical practice.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## ACKNOWLEDGEMENT:

## REFERENCES:

1. Karlawish J Desktop medicine. JAMA. 2010;304(18):2061–2062. [PubMed: 21063017]

2. Camacho PM, Petak SM, Binkley N, et al. American association of clinical endocrinologists and american college of endocrinology clinical practice guidelines for the diagnosis and treatment of postmenopausal osteoporosis. Endocrine Practice. 2016;22(9):1111–1118. [PubMed: 27643923]

3. Kanis JA, McCloskey EV, Johansson H, et al. European guidance for the diagnosis and management of osteoporosis in postmenopausal women. Osteoporos Int. 2013;24(1):23–57. [PubMed: 23079689]

4. Bibbins-Domingo K, Grossman D, Curry S, et al. Statin use for the primary prevention of cardiovascular disease in adults. JAMA. 2016;316(19):1997–2007. [PubMed: 27838723]

5. Lin JS, Piper MA, Perdue LA, et al. Screening for colorectal cancer: updated evidence report and systematic review for the US preventive services task force. JAMA. 2016;315(23):2576–2594. [PubMed: 27305422]

6. Lee SJ, Leipzig RM, Walter LC. Incorporating lag time to benefit into prevention decisions for older adults. JAMA. 2013;310(24):2609–2610. [PubMed: 24322396]

7. Redelmeier DA, Lustig AJ. Prognostic indices in clinical practice. JAMA. 2001;285(23):3024–3025. [PubMed: 11410103]

8. Gill TM. The central role of prognosis in clinical decision making. JAMA. 2012;307(2):199–200. [PubMed: 22235093]

9. Health and Retirement Study. Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant number NIA U01AG009740). Accessed May 20, 2018, Ann Arbor, MI (2000).

10. Steyerberg EW. Clinical prediction models: a practical approach to development, validation, and updating (statistics for biology and health). 9th ed. New York, NY: Springer; 2009.

11. Schwarz G Estimating the dimension of a model. Ann Statist. 1978;6(2):461–464.

12. Dziak JJ, Coffman DL, Lanza ST, Runze L, Jermiin LS. Sensitivity and Specificity of Information Criteria. Brief Bioinform. 2020 3 23;21(2):553–65. doi: 10.1093/bib/bbz016 [PubMed: 30895308]

13. Hastie T, Tibshirani R, Friedman J. (2009) The Elements of Statistical Learning: Data Mining, Inference and Prediction. (2nd ed.). Spring Series in Statistics.

14. Vrieze SI. Model selection and psychological theory: A discussion of the differences between Akaike information criterion (AIC) and the Bayesian information criterion (BIC). Psychol Methods. 2012 6;17(2):228–43. doi: 10.1037/a0027127. [PubMed: 22309957]

15. Lee SJ, Lindquist K, Segal MR, Covinsky KE. Development and validation of a prognostic index for 4-year mortality in older adults. JAMA. 2006;295(7):808–808.

16. Kilsdonk E, Peute LW, Jaspers MWM. Factors influencing implementation success of guideline-based clinical decision support systems: a systematic review and gaps analysis. Int J Med Inform. 2017;98:56–64. [PubMed: 28034413]

17. Weir DR. Validating Mortality Ascertainment in the Health and Retirement Study. 2016 Nov; Accessed 2020 Oct.

18. Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med. 1996;15(4):361–387 [PubMed: 8668867]

19. Tibshirani R Regression shrinkage and selection via the Lasso. J R Statist Soc B. 1996;58(1):267–288.

20. Zou H, Hastie T, Tibsharani R. On the "Degrees of Freedom" on the LASSO. Ann Statis. 2007;25(5):2173–92

21. Ahrens A, Hansen CB, Shaffer ME. LASSOPACK: Model selection an prediction with regularized regression in STATA. The STATA Journal. 2020;20(1):176–235.

22. SAS Institute Inc. 2016. SAS/STAT 14.2 User's Guide. The HPGENSELECT Procedure. Cary, NC: SAS Institute Inc. Retrieved from: https://support.sas.com/documentation/onlinedoc/stat/142/hpgenselect.pdf

23. Bergersen LC, Glad IK, Lyng H. Weighted Lasso with data integration. Stat Appl Genet Mol Biol. 2011;10(1).

24. Tharmaratnam K, Sperrin M, Jaki T, et al. Tilting the lasso by knowledge-based post-processing. BMC bioinformatics. 2016 12;17(1):344. [PubMed: 27590269]

25. Van De Wiel MA, Lien TG, Verlaat W, et al. Better prediction by use of co-data: adaptive group-regularized ridge regression. Statistics in medicine. 2016 2 10;35(3):368–81. [PubMed: 26365903]
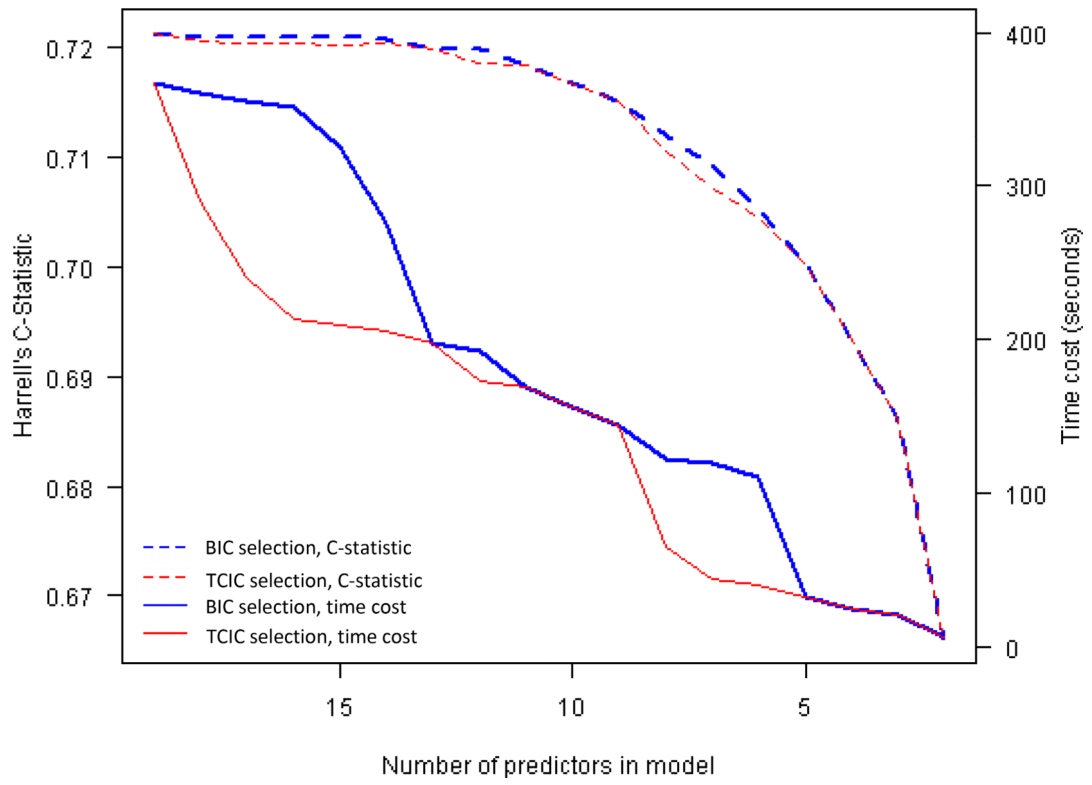
**Figure 1:**
Discrimination and Time-Costs for Models of Varying Number of Predictors.

**Table 1.**

Prevalence and time cost of predictors.

| Predictor | Time cost (seconds) | Levels of Predictor | N | Frequency (%) |
|---|---|---|---|---|
| Age | 3.0 | <=71 | 862 | 13.6 |
| | | >71–72 | 479 | 7.6 |
| | | >72–74 | 861 | 13.6 |
| | | >74–75 | 398 | 6.3 |
| | | >75–76 | 377 | 6.0 |
| | | >76–78 | 786 | 12.5 |
| | | >78–80 | 715 | 11.3 |
| | | >80–82 | 568 | 9.0 |
| | | >82–85 | 551 | 8.7 |
| | | >85 | 714 | 11.3 |
| Gender | 3.7 | Female | 3788 | 60.0 |
| | | Male | 2523 | 40.0 |
| ADL Dependency status | 21.5 | No difficulty | 5061 | 80.1 |
| | | Difficulty, no dependence | 794 | 12.6 |
| | | Dependence | 456 | 7.2 |
| Cancer | 13.1 | No cancer | 5261 | 83.46 |
| | | Cancer, but not diagnosed in last 2 years | 837 | 13.3 |
| | | Cancer, diagnosed in last 2 years | 213 | 3.4 |
| Delayed 10-word recall [*] | 78.5 | Fair (3–5 words) | 3352 | 53.1 |
| | | Bad (0–2 words) | 1976 | 31.3 |
| | | Good (6+ words) | 983 | 15.6 |
| Immediate 10-word recall [†] | 78.2 | Fair (3–5 words) | 3721 | 59.0 |
| | | Bad (0–2 words) | 608 | 9.6 |
| | | Good (6+ words) | 1982 | 31.4 |
| Diabetes | 8.7 | No diabetes / Diabetes, no meds | 5454 | 86.4 |
| | | Diabetes with oral meds/insulin | 857 | 13.6 |
| Eyesight | 27.1 | Good | 2218 | 35.1 |
| | | Bad | 4093 | 64.9 |
| Able to drive | 11.6 | No / never drove | 1586 | 25.1 |
| | | Yes | 4725 | 74.9 |
| Education | 5.1 | <12 years | 2086 | 33.1 |
| | | >=12 years | 4225 | 67.0 |
| Wear hearing aid | 3.4 | No | 5282 | 83.7 |
| | | Yes | 1029 | 16.3 |
| Heart failure | 3.2 | No heart disease/had angina/had heart attack | 6062 | 96.1 |
| | | Had heart failure | 249 | 4.0 |
| IADL Dependency status | 24.6 | No difficulty | 5304 | 84.0 |

| Predictor | Time cost (seconds) | Levels of Predictor | N | Frequency (%) |
|---|---|---|---|---|
| | | Difficulty, no dependence | 200 | 3.2 |
| | | Dependence | 807 | 12.8 |
| **Having Difficulty Climbing Stair** | 13.8 | No difficulty | 2678 | 42.4 |
| | | Difficulty climbing several flights, but no difficulty climbing one flight of stairs | 2014 | 31.9 |
| | | Difficulty walking one flight of stairs | 1619 | 25.7 |
| **Pain** | 4.8 | No pain /mild pain | 4993 | 79.1 |
| | | Moderate /severe pain | 1318 | 20.9 |
| **Depression** | 48.7 | No, CESD: 0–2 | 4669 | 74.0 |
| | | Yes, CESD: 3+ | 1642 | 26.0 |
| **Smoking** | 8.1 | Never | 2827 | 44.8 |
| | | Former | 2985 | 47.3 |
| | | Current | 499 | 7.9 |
| **Stroke** | 7.5 | No stroke | 5724 | 90.7 |
| | | Stroke, no remaining problems | 354 | 5.6 |
| | | Stroke, with remaining problems | 233 | 3.7 |
| **Volunteer** | 3.3 | No | 2812 | 44.6 |
| | | Yes | 3499 | 55.4 |

[*] Number of words from a 10-word list recalled correctly after 5 min

[†] Number of words from a 10-word list recalled correctly immediately

Table 2.

Comparison of BIC and TCIC backward selection models

| No. variables | Bayesian Information Criterion (BIC) models | | | | | Time-Cost Information Criterion (TCIC) models | | | | | p-value, BIC/TCIC c-statistics |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | Next variable to be removed | Reduction in time cost (sec) | Model time cost (sec) | BIC | Harrell's C-Statistic | Next variable to be removed | Reduction in time cost (sec) | Model time cost (sec) | TCIC | Harrell's C-Statistic | |
| 19 | -Stroke | 7 | 368 | 52167 | 0.7213 (0.7128–0.7298) | -10-Word immediate recall | 78 | 368 | 52167 | 0.7213 (0.7128–0.7298) | 1.00 |
| 18 | -Pain | 5 | 360 | 52153 | 0.7209 (0.7124–0.7294) | -Depression | 49 | 289 | 52117 | 0.7205 (0.7120–0.729) | 0.37 |
| 17 | -Hearing aid | 3 | 355 | 52144 | 0.7209 (0.7124–0.7294) | -Eyesight | 27 | 241 | 52078 | 0.7203 (0.7118–0.7288) | 0.27 |
| 16 | -Eyesight | 27 | 352 | 52136 | 0.7209 (0.7124–0.7294) | -Pain | 5 | 214 | 52057 | 0.7203 (0.7117–0.7288) | 0.23 |
| 15 | -Depression | 49 | 325 | 52128 | 0.7208 (0.7123–0.7294) | -Hearing aid | 3 | 209 | 52054 | 0.7202 (0.7117–0.7287) | 0.20 |
| 14 | -10-Word immediate recall | 78 | 276 | 52122 | 0.7206 (0.7121–0.7291) | -Stroke | 7 | 205 | 52051 | 0.7202 (0.7117–0.7287) | 0.40 |
| 13 | -Education | 5 | 198 | 52119 | 0.7198 (0.7113–0.7283) | -IADL dependency | 25 | 198 | 52048 | 0.7198 (0.7117–0.7287) | 1.00 |
| **12** | **-IADL dependency** | **25** | **193** | **52118** | **0.7197 (0.7112–0.7283)** | **-Education** | **5** | **173** | **52047** | **0.7184 (0.7099–0.727)** | **0.01** |
| 11 | -Cancer | 13 | 168 | 52120 | 0.7182 (0.7097–0.7268) | -Cancer | 13 | 168 | 52050 | 0.7182 (0.7097–0.7268) | 1.00 |
| 10 | -Ability to drive | 12 | 155 | 52129 | 0.7166 (0.708–0.7252) | -Ability to drive | 12 | 155 | 52065 | 0.7166 (0.708–0.7252) | 1.00 |
| 9 | -ADL dependency | 21 | 144 | 52145 | 0.7148 (0.7062–0.7233) | -10-Word delayed recall | 79 | 144 | 52081 | 0.7148 (0.7062–0.7233) | 1.00 |
| 8 | -Heart failure | 3 | 122 | 52182 | 0.7116 (0.7030–0.7202) | -ADL dependency | 21 | 65 | 52107 | 0.7101 (0.7014–0.7187) | 0.23 |
| 7 | -Diabetes | 9 | 119 | 52226 | 0.7088 (0.7002–0.7174) | -Heart failure | 3 | 44 | 52149 | 0.7065 (0.6979–0.7152) | 0.11 |
| 6 | -10-Word delayed recall | 79 | 110 | 52281 | 0.7047 (0.6960–0.7133) | -Diabetes | 9 | 41 | 52200 | 0.7037 (0.6951–0.7124) | 0.54 |
| 5 | -Smoking | 8 | 32 | 52367 | 0.6992 (0.6905–0.7079) | -Smoking | 8 | 32 | 52262 | 0.6992 (0.6905–0.7079) | 1.00 |
| 4 | -Volunteer | 3 | 24 | 52468 | 0.6914 (0.6826–0.7003) | -Volunteer | 3 | 24 | 52374 | 0.6914 (0.6826–0.7003) | 1.00 |

| No. variables | Bayesian Information Criterion (BIC) models | | | | | Time-Cost Information Criterion (TCIC) models | | | | | p-value, BIC/TCIC c-statistics |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Next variable to be removed | Reduction in time cost (sec) | Model time cost (sec) | BIC | Harrell's C-Statistic | Next variable to be removed | Reduction in time cost (sec) | Model time cost (sec) | TCIC | Harrell's C-Statistic | |
| 3 | -Difficulty climbing stairs | 14 | 20 | 52622 | 0.6827 (0.6738–0.6916) | -Difficulty climbing stairs | 14 | 20 | 52534 | 0.6827 (0.6738–0.6916) | 1.00 |
| 2 | -Gender | 4 | 7 | 52937 | 0.6575 (0.6481–0.6668) | -Gender | 4 | 7 | 52855 | 0.6575 (0.6481–0.6668) | 1.00 |
| 1 | -Age | 3 | 3 | 53005 | 0.6489 (0.6395–0.6583) | -Age | 3 | 3 | 52929 | 0.6489 (0.6395–0.6583) | 1.00 |

Abbreviations: sec - seconds; N/A - not applicable; IADL – instrumental activities of daily living; ADL – activities of daily living; BIC – Bayesian information criterion; TCIC – time-cost information criterion.

Rounding may lead to small inconsistencies in time cost totals within models.

Table shows that both traditional and time-cost selection started with a 19 variable model with a total time cost of 368 seconds. To get to the 18 variable model, traditional selection removed the Stroke variable which cost 7 seconds, resulting in an 18 variable model with a total time cost of 360 seconds. Alternatively, time-cost selection removed the 10-word immediate recall variable which cost 78 seconds, resulting in an 18 variable model with a total time cost of 289 seconds. The list below 18 shows the variables retained in the 18 variable model.

Both traditional and time-cost selection identified the 12 variable model as the optimum (bolded). However, the traditional method excludes education, while the TCIC method excludes IADL dependency.

If the goal is to identify the best model under 60 seconds, the traditional backward selection would identify a 5 variable model as optimal while the time-cost backward selection method would identify a 7 variable model as optimal (shaded).