

Problèmes méthodologiques liés à l'utilisation de différentes limites d'inclusion pour catégoriser les variables de résultats

Mayer M. El-Masri et Susan M. Fox-Wasylyshyn

La connaissance des facteurs qui contribuent aux temps d'attente dans le recours au traitement médical en cas d'infarctus aigu du myocarde (IAM) permettra de déterminer des interventions visant à faciliter le recours rapide aux soins. Toutefois, les définitions opérationnelles des temps d'attente varient selon les recherches. L'utilisation de limites d'inclusion différentes pour faire la distinction entre les retards et les non-retards risque de compromettre la comparabilité et le caractère généralisable des résultats obtenus dans les différentes études. L'objectif de cet article est d'examiner l'impact d'une opérationnalisation incohérente des temps d'attente, en termes de limites d'inclusion, sur la validité des résultats des recherches visant à déterminer les prédicteurs desdits temps d'attente. Une analyse des données secondaires a été effectuée à partir d'un échantillon de 73 patients qui avaient récemment subi un IAM hors de l'hôpital et conclu que leurs symptômes étaient d'ordre cardiaque. Plusieurs modèles de régression ont été élaborés afin d'examiner l'influence de l'usage de différentes limites d'inclusion (temps d'attente médian de 1, 2, 3, 6 et 12 heures) sur le nombre et la nature des prédicteurs des temps d'attente dans la recherche de soins en cas d'IAM. L'impact des différentes limites d'inclusion sur la variance expliquée, la sensibilité, la spécificité et le coefficient de prévision associés à chaque modèle de régression a été examiné. L'utilisation de différentes limites d'inclusion a donné lieu à différentes séries de prédicteurs indépendants, qui variaient en nombre et en nature. La variance expliquée par les différents modèles de régression ainsi que leurs indices de classification n'étaient pas les mêmes d'une fois à l'autre. L'utilisation de différentes limites d'inclusion pour la définition du temps d'attente a donné lieu à des résultats incohérents. Par conséquent, il est recommandé que des critères soient établis parmi les cliniciens et les chercheurs en ce qui a trait à la définition opérationnelle du temps d'attente dans le recours aux soins en cas d'IAM.

Mots clés : Définitions opérationnelles, limites d'inclusion, infarctus aigu du myocarde, temps d'attente dans le recours aux soins, sensibilité et spécificité du modèle, résultat

Best Practices in Research Methods

Methodological Issues Associated with Using Different Cut-off Points to Categorize Outcome Variables

Mayer M. El-Masri and Susan M. Fox-Wasylyshyn

Knowledge of the factors that contribute to delay in seeking medical treatment for acute myocardial infarction (AMI) provides the basis for interventions that are intended to facilitate prompt care-seeking behaviour. However, operational definitions of delay time vary across research studies. The use of inconsistent cut-off times to distinguish between delayers and non-delayers is likely to compromise comparability and generalizability of the findings across studies. The purpose of this paper is to examine the impact of inconsistent operationalization of delay, in terms of cut-off times, on the validity of research findings pertaining to identifying its predictors. Secondary data analysis was performed using a sample of 73 patients who had recently experienced out-of-hospital AMI and concluded that their symptoms were related to the heart. Several regression models were built to examine the influence of using different cut-off times (1, 2, 3, 6, and 12 hours, median delay) on the number and nature of predictors of AMI care-seeking delay. The impact of varying cut-off times on the explained variance, sensitivity, specificity, and predictive values associated with each regression model was examined. The use of different cut-off times produced different sets of independent predictors, which varied in number and nature. The variance explained by the different regression models as well as their classification indices varied. Use of different cut-off times for the definition of delay time led to inconsistent results. Thus, it is recommended that criteria be established among clinicians and researchers with regard to operationally defining care-seeking delay for AMI.

Keywords: Operational definitions, cut-off times, acute myocardial infarction, care-seeking delay, model sensitivity and specificity, dichotomization, outcome

Operational definitions of outcome variables have a significant impact on the validity and generalizability of research findings. Dichotomization of continuous variables represents one situation in which generalizability, comparability, and synthesis of findings across studies can be compromised. This is because different authors may select varying criteria to determine the cut-off point at which subjects are classified as having or

not having the outcome of interest. The purpose of this paper is to demonstrate how the use of different cut-off points to dichotomize a continuous variable, delay in deciding to seek care for acute myocardial infarction (AMI), compromises the validity of research findings.

Given that effectiveness of reperfusion therapies used in the treatment of AMI is time-dependent, many researchers have investigated the phenomenon of delay in seeking treatment for AMI. However, review of the literature suggests that investigators (Burnett, Blumenthal, Mark, Leimberger, & Califf, 1995; Clark, Bellam, Shah, & Feldman, 1992; Dracup & Moser, 1997; Johnson & King, 1995; King & McGuire, 2000; McKinley, Moser, & Dracup, 2000; Wu, Zhang, Li, Hong, & Huang, 2004) have used a wide variety of approaches to operationalizing delay time. This is because, when used as a continuous variable, delay among AMI patients tends to have severe positive skew due to a common tendency for a small proportion of patients to delay seeking medical attention for a relatively long period (i.e., days vs. hours). One approach to the management of skewed data is mathematical transformation. However, mathematical transformation produces scores that can be difficult to interpret because they no longer carry the unit of analysis of the original data. For example, in our data set, the base log of delay time of 0.5 hours was $-.30$, which is clearly difficult to explain and/or compare in terms of actual/exact time. In addition, mathematical transformation procedures may sometimes fail to produce a normal distribution when the departure from normality is severe (Tabachnick & Fidell, 2001). In such cases, investigators often choose to dichotomize the variable into two categories (i.e., delayers and non-delayers). Unfortunately, the selection of the cut-off time used to distinguish subjects as delayers or non-delayers differs widely among studies. A review of the literature suggests that 1 hour (Al-Hassan & Omran, 2005; Carney, Fitzsimons, & Dempster, 2002), 2 hours (Dempsey, Dracup, & Moser, 1995; Turi et al., 1986), 3 hours (King & McGuire; Reilly, Dracup, & Dattolo, 1994), 6 hours (Schmidt & Borsch, 1990; Sheifer et al., 2000), and 12 hours (Ruston, Clayton, & Calnan, 1998) have all been used as cut-offs for defining care-seeking delay. These inconsistent operationalizations of delay across studies present a challenge to the understanding and generalizability of the research findings that pertain to delay among AMI patients. Therefore, the purpose of this study was to examine the impact of using varied cut-off values for operationalizing delay on the validity of research findings pertaining to identifying its predictors. Specifically, this study examined how the use of various cut-off times impacts: (a) the number and nature of predictors, (b) the magnitude and direction of relationships, (c) the amount of explained variance, and (d) classification indices (i.e., sensitivity, specificity, positive predictive value, and negative predictive value) of a regression model.

Methods

Sample

Secondary data analyses were performed on data that were collected after ethical approval was granted to examine the factors that impact care-seeking delay among patients who had experienced a recent AMI (Fox-Wasylyshyn, 2005). The database represented a sample of 73 AMI patients who concluded, prior to seeking medical care, that their symptoms were heart-related. Patients were recruited from two hospitals in southeastern Michigan and two in southwestern Ontario. The participants were over 18 years, deemed to be in stable physiological condition, and able to speak English. Data were collected using a structured interview technique during the first 24 to 96 hours post-admission.

Variable Definitions

History of AMI was defined in terms of whether or not a participant experienced AMI prior to the one for which he or she was currently admitted. *Symptom congruence* was defined as the degree to which one's AMI symptom experience was consistent with pre-conceived notions about the nature of AMI symptoms. It was measured using a 10 cm horizontal visual analogue scale on which participants were asked to indicate with a vertical line how similar or different their heart attack symptoms were in relation to how they had previously thought a heart attack would feel (0 = not at all similar; 10 = exactly as expected). *Emotion-focused coping* was measured using the Coping with Heart Attack Symptoms Scale (CHASS), a five-item Likert scale (Fox-Wasylyshyn, 2005). Factor analysis suggested that the CHASS measured a single concept in which item loadings ranged from .54 to .87, a reflection of good construct validity. The CHASS items were internally consistent (Cronbach's alpha: 0.76) (Kline, 1998). *Time to cardiac symptom attribution (CSA)* refers to the time interval between AMI symptom onset and when the participants concluded that their symptoms were heart-related. This variable was operationally defined as the time (in hours) from symptom onset until the participant determined that the probable cause of his or her symptoms was the heart. *AMI care-seeking delay* was defined as the time interval (in hours) between the onset of AMI symptoms and the decision to seek medical treatment. AMI care-seeking delay comprised the intervals between: (a) symptom onset and CSA, and (b) CSA and the decision to seek medical care.

Data Analysis

Data analysis procedures were performed using the Statistical Package for Social Sciences (SPSS) computer program (version 14.0; SPSS Inc.). Prior

to the analysis, data were explored for missing data, multicollinearity, and deviation from normality. Time to CSA and AMI care-seeking delay were positively skewed. Time to CSA was transformed to the Base log (Lg10) of its value. For the purpose of this paper, AMI care-seeking delay was dichotomized using five different cut-off times (1, 2, 3, 6, and 12 hours). Basic descriptive statistics such as means, medians (*mdn*), standard errors (*se*), frequencies, and percentages were performed to describe the sample. Five logistic regression models, one for each cut-off time, were built to assess the impact of different cut-off times on the results. The same independent variables were entered into each regression model using a forward stepwise approach. These variables were time to CSA, emotion-focused coping, history of AMI, symptom congruence, age, gender, country of residence, and level of education. The variables ethnic background and medical insurance were excluded from analysis because their categories had more than a 90% split, which could have truncated correlation coefficients with other variables (Tabachnick & Fidell, 2001). The regression models were compared in terms of their independent predictors, variance explained, and predictive accuracy. The predictive accuracy of each model was determined by the specificity, sensitivity, positive predictive value, and negative predictive value of that model. A 95% confidence interval (95% CI) was set as the criterion to establish significance.

Results

The sample mean age was 59 years ($SE \pm 1.42$). The majority of participants were non-Hispanic white ($n = 66$; 90.4%), with men making up 71.9% ($n = 97$) of the sample. Most participants ($n = 51$; 69.9%) were in the hospital for their first AMI. Fifty-six participants (76.7%) lived in Canada, while the remaining 22 (30.1%) lived in the United States. With respect to education level, 44 participants (60.3%) had a high school education or less. The mean AMI care-seeking delay time was 8.48 hours ($mdn = 1.0$ hour; $SE \pm 2.30$), while the mean time to CSA was 5.97 hours ($mdn = .33$ hour; $SE \pm 2.2$).

Table 1 presents the results of the five logistic regression models that were built to identify the predictors of care-seeking delay using different cut-off times to classify participants as delayers or non-delayers. The results show that Model 4 explained the smallest proportion of variance (13.9%), whereas Models 1 and 5 explained the largest proportion of variance (40.8% and 47.8%, respectively) in care-seeking delay. The results also show that each model identified a different set of independent predictors of care-seeking delay, which varied from as few as one predictor (Models 2 and 4) to as many as five predictors (Model 5).

Table 1 *Summary of Results of Logistic Regression Models Using Varied Cut-off Times to Dichotomize AMI Care-Seeking Delay*

Model	Cut-off Time (hours)	Variance (%)	Independent Predictors of Delay	OR	95% CI
1	1*	40.8	Country of residence	0.14	0.02–0.81
			History of AMI	10.57	2.12–52.85
			Education	5.93	1.53–22.98
			Time to CSA	3.89	2.02–7.51
2	2	21.3	Time to CSA	2.67	1.40–4.81
3	3	25.0	History of AMI	5.16	1.40–19.04
			Education	3.50	1.04–11.76
			Time to CSA	2.38	1.40–4.05
4	6	13.9	Time to CSA	2.33	1.33–4.09
5	12	47.8	Country of residence	13.00	1.44–117.35
			History of AMI	0.07	0.01–0.40
			Education	0.09	0.02–0.45
			Emotion-focused coping	0.87	0.75–0.99
			Time to CSA	0.20	0.09–0.45

*Median time delay.

Table 2 *Classification Indices^a Associated with Logistic Regression Models*

Model	Cut-off Time (hours)	Classification Indices			
		Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
1	1	75.0	81.1	79.4	76.9
2	2	76.7	83.7	76.7	83.7
3	3	57.7	80.9	62.5	77.6
4	6	33.3	100.0	100.0	82.1
5	12	79.4	79.5	77.1	81.6

^aClassification indices are based on a threshold of 0.5.

Results of Models 1 and 3 suggest that a history of AMI is a risk factor for prolonged delay ($OR = 10.57$ and 5.16 , respectively), while Model 5 suggests that it was inversely related to delay ($OR = .07$). Table 2 shows that the models had different classification indices, with sensitivity ranging from 33.3% (Model 4) to 79.4% (Model 5) and specificity ranging between 79.5% (Model 5) and 100% (Model 4).

Discussion

The findings of this study demonstrate that the selection of various cut-off times produces varying results with regard to (a) the nature and number of independent predictors of AMI care-seeking delay, (b) the direction and magnitude of relationships between predictors and AMI care-seeking delay, (c) the amount of explained variance, and (d) the classification indices of the overall regression models. Such variations present a challenge to the comparability, generalizability, and synthesis of research findings, and thus limit our understanding of the risk factors of care-seeking delay.

The findings demonstrate that using different cut-off times is likely to yield different results with regard to the nature and number of risk factors of delay. While Models 2 and 4 each had one independent predictor, Model 5 had five predictors (Table 1). In addition, the nature of predictors varied among models, as none were identical with regard to their independent predictors. For instance, emotion-focused coping was an independent predictor of delay only in Model 5, despite the fact that all five models were built from the same data. This finding is concerning when the primary purpose of a study is to identify the predictors of a given outcome so that they can be used to identify those who are at risk, to conduct intervention research, and to provide clinicians with specific targets upon which they can intervene.

The findings demonstrate that the magnitude and direction of the relationship between care-seeking delay and its risk factors are dependent on the cut-off time that is used to classify subjects as delayers or non-delayers. For example, Model 1 demonstrates that those who had a history of AMI were 10.6 times more likely to delay seeking care than those without a history of AMI, while Model 3 suggests that they were only five times more likely to delay. Although Models 1 and 5 had four independent predictors in common (Table 1), the relationships between the predictors and the outcome were in opposite directions in the two models. For instance, while country of residence was associated with shorter delay in Model 1, it was associated with longer delay in Model 5. Further, history of AMI, education, and time to CSA were positively associated with delay in Model 1 but were inversely related to delay in

Model 5. The aforementioned findings pertaining to the magnitude and direction of relationships are likely due to the fact that some subjects move from one category to another (delayers vs. non-delayers) each time a different cut-off time is established. For instance, a patient who waits 1.5 hours prior to deciding to seek medical care would be classified as a non-delayer when a cut-off time of 2 hours is selected, but would be classified as a delayer when a cut-off time of 1 hour is selected.

The results show that the amount of explained variance in care-seeking delay was different across all five models, ranging from 13.9% (Model 4) to 47.8% (Model 5). For example, although time to CSA was the only predictor of care-seeking delay in Models 2 and 4, the variance it explained in these two models was different (21.3% and 13.9%, respectively). Again, this difference in variance is likely a function of the change in patients' classification (delayers vs. non-delayers) among models, and the change in the nature and number of independent variables included in each model.

Logistic regression modelling provides researchers with classification indices from which the specificity, sensitivity, positive predictive values, and negative predictive values of the model can be assessed. These indices provide information to determine the clinical value of a model in terms of its ability to accurately classify patients with respect to an outcome variable. In our study, model specificity refers to the proportion of patients who did not delay and were correctly classified as such by that model. Sensitivity refers to the proportion of patients who delayed and were correctly identified as such by the model. The model's positive predictive value refers to the probability that it correctly predicts a patient as a delayer. The negative predictive value of the model refers to the probability that it correctly predicts that a patient will not delay.

The sensitivity and specificity of a model have a significant impact on its predictive values. For instance, models with greater sensitivity are likely to result in more false negative predictions (i.e., classifying true delayers as non-delayers). Likewise, an increase in model specificity tends to increase its positive predictive values, thus increasing the likelihood of false positive predictions (i.e., classifying true non-delayers as delayers). However, the predictive value of a model (positive and negative) is not completely dependent on the sensitivity and specificity of that model. The prevalence of the outcome (in this case, care-seeking delay) has a significant impact on a model's predictive value, especially when the outcome is a rare one. For instance, regardless of how specific the model is, positive predictive values tend to be largely false when the prevalence of the outcome is very low. Our data show that using different cut-off times yields vastly different positive predictive values that vary from 62.5% (Model 3) to 100% (Model 4), and sensitivity scores that vary from

33.3% (Model 4) to 79.4% (Model 5). This is likely due to the change across models in the prevalence of delay that results when the cut-off time used to define care-seeking delay is changed — that is, as the cut-off time increased from 1 through 12 hours, the proportion of subjects in the sample who were classified as delayers decreased from 55.6% to 22.2%.

The data show that Models 1, 2, and 5 have the best classification indices (all exceeded 75%). However, Model 5 has a cut-off time of 12 hours, and thus it may not be clinically meaningful due to the time-dependent nature of reperfusion therapies. In addition, contrary to expectations, it suggests an inverse relationship between delay and time to CSA such that individuals who take longer to attribute their symptoms to the heart are likely to have a shorter delay. This finding further highlights the problem with using 12 hours as a cut-off time for delay. Given that Model 1 explains almost twice the variance explained by Model 2, and that the cut-off times used in both models can be considered to be clinically relevant (1 and 2 hours), Model 1 is probably preferable, as it explains the greater proportion of variance, has reasonably high sensitivity and positive predictive values, and uses a clinically meaningful cut-off time.

The aforementioned discussion demonstrates that lack of consistency with regard to the cut-off time used to dichotomize care-seeking delay has a multifaceted impact on the validity of results. The findings demonstrate that variation in cut-off times impacts on the nature and number of predictors, the magnitude and direction of the relationship between a predictor and the outcome, and the amount of variance that a model explains. These variations are likely to compromise the validity of research findings pertaining to delay and limit meaningful comparisons among different studies. Further, they may limit our understanding of the nature of the true relationships between care-seeking delay and its risk factors, which may subsequently limit our ability to implement effective interventions intended to decrease AMI care-seeking delay. We therefore recommend that, whenever possible, researchers operationalize delay as a continuous variable (actual hours or minutes). If the variable is not normally distributed, then dichotomizing the variable can be used with the stipulation that selection of cut-off scores is based on sound theory or clinical judgement. When success of therapy is time-dependent, as in the case of reperfusion therapy, it is important that the cut-off point used to operationally define care-seeking delay be carefully considered — that is, researchers should use theoretical and/or clinical evidence to guide the definition of the time frame that constitutes delay. In fact, it might be helpful if the Heart and Stroke Foundation and the American Heart Association developed benchmarking criteria pertaining to the opera-

tional definition of delay. Such benchmarking is likely to allow clinicians and researchers to compare and synthesize findings of different studies and improve our understanding of the risk factors of delay.

Despite the study limitations related to retrospective data collection and small sample size, the results indicate that using different cut-off times has a significant impact on the validity of findings.

References

- Al-Hassan, M. A., & Omran, S. M. (2005). The effects of health beliefs on health care-seeking decisions of Jordanian patients with myocardial infarction symptoms. *International Journal of Nursing Practice*, *11*(1), 13–20.
- Burnett, R. E., Blumenthal, J. A., Mark, D. B., Leimberger, J. D., & Califf, R. M. (1995). Distinguishing between early and late responders to symptoms of acute myocardial infarction. *American Journal of Cardiology*, *75*, 1019–1022.
- Carney, R., Fitzsimons, D., & Dempster, M. (2002). Why people experiencing acute myocardial infarction delay seeking medical assistance. *European Journal of Cardiovascular Nursing*, *1*(4), 237–242.
- Clark, L. T., Bellam, S. V., Shah, A. H., & Feldman, J. G. (1992). Analysis of prehospital delay among inner-city patients with symptoms of myocardial infarction: Implications for therapeutic intervention. *Journal of the National Medical Association*, *84*, 931–937.
- Dempsey, S. J., Dracup, K., & Moser, D. K. (1995). Women's decision to seek care for symptoms of acute myocardial infarction. *Heart and Lung*, *24*(6), 444–456.
- Dracup, K., & Moser, D. K. (1997). Beyond sociodemographics: Factors influencing the decision to seek treatment for symptoms of acute myocardial infarction. *Heart and Lung*, *26*(4), 253–262.
- Fox-Wasylyshyn, S. (2005). *Delay in seeking treatment for acute myocardial infarction*. Unpublished doctoral dissertation, Wayne State University.
- Johnson, J. A., & King, K. B. (1995). Influence of expectations about symptoms on delay in seeking treatment during a myocardial infarction. *American Journal of Critical Care*, *4*(1), 29–35.
- King, K. B., & McGuire, M. A. (2000). Symptom presentation and delay in seeking care for women and men with myocardial infarction. *Circulation*, *102*(18, Suppl 2), p. 671.
- Kline, R. B. (1998). *Principles and practice of structural equation modeling*. New York: Guilford.
- McKinley, S., Moser, D. K., & Dracup, K. (2000). Treatment-seeking behavior for acute myocardial infarction symptoms in North America and Australia. *Heart and Lung*, *29*(4), 237–247.
- Reilly, A., Dracup, K., & Dattolo, J. (1994). Factors influencing prehospital delay in patients experiencing chest pain. *American Journal of Critical Care*, *3*(4), 300–306.
- Ruston, A., Clayton, J., & Calnan, M. (1998). Patient's action during their cardiac event: Qualitative study exploring differences and modifiable factors. *British Medical Journal*, *316*, 1060–1064.

- Schmidt, S. B., & Borsch, M. A. (1990). The prehospital phase of acute myocardial infarction in the era of thrombolysis. *American Journal of Cardiology*, *65*(22), 1411–1415.
- Sheifer, S. E., Rathore, S. S., Gersh, B. J., Weinfurt, K. P., Oetgen, W. J., Breall, J. A., et al. (2000). Time to presentation with acute myocardial infarction in the elderly: Associations with race, sex, and socioeconomic characteristics. *Circulation*, *102*(14), 1651–1656.
- Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). Boston: Allyn & Bacon.
- Turi, Z. G., Stone, P. H., Muller, J. E., Parker, C., Rude, R. E., Rraabe, D. E., et al. (1986). Implications for acute intervention related to time of hospital arrival in acute myocardial infarction. *American Journal of Cardiology*, *58*, 203–209.
- Wu, Y., Zhang, Y., Li, Y. Q., Hong, B. L., & Huang, C. X. (2004). Factors associated with the extent of care-seeking delay for patients with acute myocardial infarction in Beijing. *Chinese Medical Journal*, *117*(12), 1772–1777.

Authors' Note

Comments or queries may be directed to Maher M. El-Masri, RN, PhD, Faculty of Nursing, University of Windsor, Health Education Centre, Room 322, 401 Sunset, Windsor, Ontario N9B 3P4 Canada. Telephone: 519-253-3000, ext. 2400. Fax: 519-973-7084. E-mail: melmasri@uwindsor.ca

Maher M. El-Masri, RN, PhD, is Associate Professor and Research Coordinator, Faculty of Nursing, University of Windsor, Ontario, Canada. Susan M. Fox-Wasylyshyn, RN, PhD, is Associate Professor, Faculty of Nursing, University of Windsor.