SUMMARY: In medical research analyses, continuous variables are often converted into categoric variables by grouping values into ≥2 categories. The simplicity achieved by creating ≥2 artificial groups has a cost: Grouping may create rather than avoid problems. In particular, dichotomization leads to a considerable loss of power and incomplete correction for confounding factors. The use of data-derived “optimal” cut-points can lead to serious bias and should at least be tested on independent observations to assess their validity. Both problems are illustrated by the way the results of a registry on unruptured intracranial aneurysms are commonly used. Extreme caution should restrict the application of such results to clinical decision-making. Categorization of continuous data, especially dichotomization, is unnecessary for statistical analysis. Continuous explanatory variables should be left alone in statistical models.

ABBREVIATIONS: ACA = anterior cerebral artery; CHUM = Centre hospitalier de l’Université de Montréal; ICA = internal carotid artery; ISUIA = International Study of Unruptured Intracranial Aneurysms; MCA = middle cerebral artery; PcomA = posterior communicating artery; SAH = subarachnoid hemorrhage; UIA = unruptured intracranial aneurysms
ment of continuous variables arise very widely across the spectrum of medical and surgical specialties. We will illustrate these issues by using the case of UIAs as assessed in the second ISUIA study.7

How Size Was Analyzed in the Prospective ISUIA Study
The prospective ISUIA study7 is a significant example in our field. This study represents the most important contribution to the knowledge of the elusive natural history of UIAs to date. Published in 2003, it included 4060 patients enrolled during a 7-year period from 60 centers in the United States, Canada, and Europe. Multiple categories of patients and aneurysms were created according to clinical history, size, location of aneurysms, and management of patients. The 4060 patients studied included 983 who had a history of aneurysmal SAH (group 2) and 3077 without such a history (group 1). A total of 2368 patients were treated, 1917 by clipping and 451 by endovascular therapy, while 1692, who did not have aneurysm treatment, were considered the “natural history group” and will be the target of the current discussion. There were 2686 aneurysms in these 1692 patients; 40% of patients had multiple aneurysms. It is unclear how patients were categorized and how eventual ruptures were allocated when patients had multiple aneurysms because categories concern patients, but size is a characteristic of aneurysms. We can only presume that the largest of multiple aneurysms, along with its location, served to categorize the patient, at least in Table 4 (which seems to add up for patients, not for aneurysms).7 There is room here for different interpretations.

The hypothesis in the ISUIA study was that size was an important risk factor for future ruptures. One must at first consider the uncertainty of measurements, the imprecision of the methods of correction for magnification and for various projections, from imaging studies performed between 1991 and 1998 from multiple institutions with diverse equipment of widely ranging quality. Patients who were observed were selected according to clinical judgment, and it is fair to assume that this decision was partly based on the variable to be studied: size. Hence not only are the observations biased but they cannot be representative of the natural history of patients whom physicians want to treat.

In the ISUIA study, the resulting group of patients left untreated was divided into 4 size categories as seen in the Table. Categories of sizes were determined in a data-dependent fashion: “the running average for successive 3-mm-size categories showed optimum cut-points at diameters <7 mm, 7–12 mm, 13–24 mm, and 25 mm or larger”7 for the 5-year cumulative rupture rates, illustrated in Table 4 but grouped in 3 categories (<7, 7–12, and >12) for the multivariate analyses of predictors of hemorrhage. The 3 categories further changed for clinical outcomes 1 year after treatment: Here the categories were ≤12, 13–24, and >25. It is unclear how and why groupings were constructed or dismantled because sometimes the size categories were split according to the previous history of the patients (group 1 or 2 for aneurysms <7 mm) and sometimes they were lumped (>7 mm). The same problem resurfaces when location is considered a characteristic to be used for categorizing in combination with size, with the further problem that some anatomic locations are unorthodox (with some carotid aneurysms that have presented an event, the PcomA, now being lumped with the posterior circulation aneurysms). Remember that these problems are compounded by the fact that 40% of patients had multiple aneurysms of various sizes and locations.

It remains impossible to work out precise numbers because they are not provided in Table 4, which included only percentages.7 It is even unclear if we are here dealing with actual events or estimated events because Table 4 presented 5-year cumulative percentages, while the mean follow-up period was 4 years.7 At least 15 different categories are presented in Table 4 (and probably 3 times as many have been played with), while there were only 49 (or 51) events, far too few to provide any confidence in the attribution of risks to any individual category.7 Confidence intervals cannot be calculated because precise numbers were not provided, but they must be very wide. If rates deprived of confidence intervals are then projected over the lifetime of an individual to provide a natural history or prognosis, with the presumption that risk remains constant, incalculable error is multiplied beyond control. Even if ISUIA 2003 did not explicitly use 2 groups split at 7 mm—the authors never gave the overall rupture rate for <7 mm aneurysms— the 5-year cumulative rupture rates rise above 7 mm for all locations. Consequently many clinicians now use the 7-mm limit as a threshold to guide clinical decisions. Multiple methodologic flaws are involved in this conclusion, which is, in our opinion, unjustified and risky, but our discussion will be restricted to categorization of continuous data.

Problems with Dichotomization
The first problem that categorizing a continuous variable causes is loss of information; that loss is small with several groups and is most severe with just 2 groups. Several studies16,17 have demonstrated that 100 continuous observations are statistically equivalent to at least 157 dichotomized observations. Selvin18 derives a formula to calculate the efficiency loss due to categorizing a continuous variable. The problem increases in importance when one attempts to take into account confounding factors to reveal the relation between the event rate and the variable. Becher et al19 found that models with a categorized exposure variable removed only 67% of the confounding controlled when the continuous version of the variable was used. A further important disadvantage of dichot-
omization: it does not make use of within-category information: Everyone above or below the cut-point (often the sample median) is treated as equal, yet their prognosis may vary considerably. Dichotomization by necessity leads to pooling groups with different risks and to an unrealistic “step function” for risk. The risk of misclassification because of measurement error is high. In addition, comparing studies that used different cut-points becomes impossible. Hence, for most statisticians, dichotomization is not a natural way of analyzing continuous data.

In the ISUIA, the cut-points were chosen after inspecting the data, rendering all results exploratory and in need of confirmation from data of a different source before being seriously considered for clinical decisions. If size is to be dichotomized, the choice of a cut-point should be made before analysis and, if possible, with some theoretic or clinical justification. The size of the value or odds ratio should never influence the choice of cut-points. This procedure leads to $P$ values that are too small and overestimates the prognostic value of the variable.

Some researchers have argued that loss of power and efficiency is not important if statistically significant effects are still found with dichotomous variables. Why should we use a regression coefficient, a concept the meaning of which may be less likely to be understood by typical research consumers? Indeed, if type I and II errors are all that matter and the results achieve statistical significance, researchers might feel justified in dichotomizing such measures so as to simplify conclusions. The objective of such research should be to get reliable estimates of risk, not just to ensure a statistically significant result. What is not well appreciated in this reasoning is that categorizing continuous variables may not only miss the message, it can also get it wrong: Under some circumstances, categorizing continuous variables can give biased results. In a simulation study, Taylor and Yu found that categorizing 1 continuous variable can artificially make another variable appear associated with the outcome. More generally, the cut-point chosen by looking into data during the categorization of continuous variables significantly changed the calculated odds ratio. Information loss and bias from categorizing continuous variables explain why statisticians frequently warn us to leave continuous variables alone.

There is thus much support for using regression analysis in which variables are kept continuous. Although this approach is simplest when there is a reasonably linear relation between the variable and the outcome, it is not difficult to extend the approach to nonlinear relations. It appears that this advice was lost to investigators—ourselves included—who have developed risk stratification schemes for patients with UIAs. Most rupture-risk stratification schemes have categorized size by using arbitrary or data-driven cut-points. Much the same practices have been documented in other areas, so it is clear that the many warnings by statisticians against this type of analysis have not been heeded.

**Conclusions and Recommendations**

Forcing individuals into 2 groups, such as <7 mm and ≥7 mm, is widely perceived to simplify analyses and facilitate presentation and interpretation of findings. In fact, size is so frequently dichotomized that some may believe this to be a recommended practice. Use of categories yields results that are readily understood by colleagues, policy makers, and the interested public. Nevertheless such results may distort the reported relationship and poison the clinical meaning of findings. Modern regression models do not require categorization. In general, continuous variables should remain continuous in regression models designed to study the effects of the variable on the outcome of interest.

We recommend the following:

1) If a continuous variable such as size is to be dichotomized, the choice of cut-point should be made before analysis and with some theoretic or clinical justification. Data-driven cut-points should be avoided. Never choose an optimal cut-point based on minimizing the $P$ value or maximizing statistics such as odds ratios.

2) If a continuous variable is categorized, having ≥3 groups is preferable to just 2. Again, prespecification of cut-points is strongly recommended.

3) If the assumption of a linear relation (eg, between rupture risk and size) is supported, size can best be used as a continuous variable in a regression model. Careful interpretation of the resulting odds ratio as representing the effect for each additional millimeter can be readily understood by readers. Nonlinear relations can also be quite easily modeled.

**References**

1. Rinkel GJ, Dijbbui M, Algra A, et al. Prevalence and risk of rupture of intracranial aneurysms: a systematic review. Stroke 1998;29:251–56
2. Cowan JA Jr, Zierler J, Dimick JB, et al. Use of endovascular coil embolization and surgical clip occlusion for cerebral artery aneurysms. J Neurosurg 2001;97:1530–35
3. Jutla S, Porruc M, Poussa K. Natural history of unruptured intracranial aneurysms: probability of and risk factors for aneurysm rupture. J Neurosurg 2008;108:1052–60
4. Ishibashi T, Murayama Y, Ikeyu S, et al. Unruptured intracranial aneurysms: natural history, risk of rupture, and clinical outcome of endovascular treatment as a first line of therapy. Int Stroke J 2007;38:595
5. Unruptured intracranial aneurysms: risk of rupture and risks of surgical intervention—International Study of Unruptured Intracranial Aneurysms Investigators. N Engl J Med 1998;339:1725–31
6. Bederson JB, Avad IA, Wisers DO, et al. Recommendations for the management of patients with unruptured intracranial aneurysms: a statement for healthcare professionals from the Stroke Council of the American Heart Association. Circulation 2000;102:2300–08
7. Wiebers DO, Whisnant JP, Huston J 3rd, et al. Unruptured intracranial aneurysms: natural history, clinical outcome, and risks of surgical and endovascular treatment. Lancet 2003;362:103–10
8. Yoshimoto Y. A mathematical model of the natural history of intracranial aneurysms: quantification of the benefit of prophylactic treatment. J Neurosurg 2006;104:195–200
9. Greving JP, Rinkel GJ, Buikens E, et al. Cost-effectiveness of preventive treatment of intracranial aneurysms: new data and uncertainties. Neurology 2009;73:258–65
10. Rosyn S, Altman DG, Saurerwli B. Dichotomizing continuous predictors in multiple regression: a bad idea. Stat Med 2006;25:127–41
11. Chen H, Cohen P, Chen S. Biased odds ratios from dichotomization of age. Stat Med 2007;26:3487–97
12. Steiner DL. Breaking up is hard to do: the heartbeat of dichotomizing continuous data. Can J Psychiatry 2002;47:262–66
13. van Walraven C, Hart RG. Leave ‘em alone why continuous variables should be analyzed as such. Neuroradiology 2008;38:39–49. Epub 2008 Apr 17
14. Vickers AJ, Lilja H. Cutpoints in clinical chemistry: time for fundamental reassessment. Clin Chem 2009;55:15–17
15. Heavner KK, Phillips CV, Burstyn I, et al. Dichotomization: 2 = 2 (x = 2 x 2 x 2?) categories: infinite possibilities. BMJ Med Res Methodol 2010:10:59
16. Fedorov V, Mannino F, Zhang R. Consequences of dichotomization. Pharm Stat 2009;8:50–61
17. Zhao LP, Kolonel LN. Efficiency loss from categorizing quantitative exposures.
into qualitative exposures in case-control studies. Am J Epidemiol 1992;136:464–74.
18. Selvin S. Statistical Power and Sample Size Calculations: Statistical Analysis of Epidemiological Data. New York: Oxford University Press; 2004:75–92.
19. Becher H, Grau A, Steindorf K, et al. Previous infection and other risk factors for acute cerebrovascular ischaemia: attributable risks and the characterisation of high risk groups. J Epidemiol Biostat 2000;5:277–83.
20. Ewald B. Post hoc choice of cut points introduced bias to diagnostic research. J Clin Epidemiol 2006;59:798–801.
21. Vargha A, Rudas T, Delaney H, et al. Dichotomization, partial correlation, and conditional independence. Journal of Educational and Behavioral Statistics 1996;21:284–82.
22. Taylor J, Yu M. Bias and efficiency loss due to categorizing an explanatory variable. Journal of Multivariate Analysis 2002;83:248–63.
23. Ragland DR. Dichotomizing continuous outcome variables: dependence of the magnitude of association and statistical power on the cutpoint. Epidemiology 1992;3:434–40.
24. Perel P, Edwards P, Wentz R, et al. Systematic review of prognostic models in traumatic brain injury. BMC Med Inform Decis Mak 2006;6:38.
25. Mushkudiani NA, Hukkelhoven CW, Hernandez AV, et al. A systematic review finds methodological improvements necessary for prognostic models in determining traumatic brain injury outcomes. J Clin Epidemiol 2008;61:331–43.
26. Leushuis E, van der Steeg JW, Steures P, et al. Prediction models in reproductive medicine: a critical appraisal. Hum Reprod Update 2009;15:537–52.
27. Mallett S, Royston P, Dutton S, et al. Reporting methods in studies developing prognostic models in cancer: a review. BMC Med 2010;8:20.
28. Royston P, Sauerbrei W. Building multivariable regression models with continuous covariates in clinical epidemiology: with an emphasis on fractional polynomials. Methods Inf Med 2005;44:561–73.
29. Royston P, Sauerbrei W. A new approach to modelling interactions between treatment and continuous covariates in clinical trials by using fractional polynomials. Stat Med 2004;23:2369–85.