

Collapsing high-end categories of comorbidity may yield misleading results

Timothy L Lash

Department of Epidemiology, Boston University School of Public Health, Boston, MA, USA; Department of Clinical Epidemiology, Aarhus University Hospital, Aarhus, Denmark

Abstract: Adequate control of comorbidity has long been recognized as a critical challenge in clinical epidemiology. Comorbidity scales reduce information about coexistent disease to a single index that is easy to comprehend and statistically efficient. These are the main advantages of an index over incorporating each disease into an analysis as an individual variable. Many study populations have a low prevalence of subjects with high comorbidity scores, so it is common to combine subjects with some score above a threshold into a single open-ended category. This paper examines the impact of collapsing comorbidity scores into these categories. It shows analytically and by synthetic example that collapsing the high-end categories of a comorbidity scale changes the pattern of effect of comorbidity. Furthermore, collapsing the high-end categories biases analyses that control for comorbidity as a confounder or analyze modification of an exposure's effect by comorbidity. Each of these results specific to comorbidity scoring derives from more general epidemiologic principles. The appeal of collapsing categories to facilitate interpretation and statistical analysis may be offset by misleading results. Analysts should assure the uniformity of outcome risk in collapsed categories, informed by judgment and possibly statistical testing, or use analytic methods, such as restriction or spline regression, which can achieve similar goals without sacrificing the validity of results.

Keywords: epidemiologic factors, comorbidity, epidemiologic factors, bias (epidemiology)

Introduction

A recent US National Institute on Aging Task Force defined comorbidity as “the co-occurrence of preexisting age-related health conditions (eg, disability, anemia, impairments, urinary incontinence) or diseases (eg, diabetes, heart disease, hypertension) in reference to an index disease (eg, cancer, Parkinson's disease, diabetes).”¹ Adequate measurement and analytic control of comorbidity has long been recognized as a critical challenge in clinical epidemiology.² The aforementioned task force has reviewed the methodology of measurement of comorbidity,³ including the nosology of disease classification⁴ and strategies to include disease severity in comorbidity scales.⁵

Collapsing comorbid diseases into a single scale provides an index that is easy to comprehend and statistically efficient, which are the main advantages of an index over incorporating each disease into an analysis as an individual variable.⁶ A simple sum of the number of comorbid diseases treats each disease equivalently, thereby ignoring differences in the severity of the component diseases and differences in the severity of the disease state in different patients. Weighting schemes have been proposed and implemented to address each of these shortcomings.^{5,7} Whether summing diseases included in the index or weighting them by severity, all comorbidity schemes inevitably misclassify study subjects with respect to the idealized true scale of comorbidity.³ The impact of this misclassification on the analytic results depends on whether comorbidity is the exposure of interest, study outcome, a confounder, or modifier.³

Correspondence: Timothy L Lash
Boston University School of Public Health, 715 Albany St., TE3, Boston, MA 02118, USA
Tel +1 617 638 8384
Fax +1 317 638 4458
Email tlash@bu.edu

In most study populations, the prevalence of subjects with high comorbidity scores is low. It is common, therefore, to combine subjects with some score above a threshold into a single open-ended category. For example, the Charlson Index can theoretically range from 0 to 33 but was collapsed into categories of 0, 1–2, 3–4, and ≥ 5 in its initial presentation.⁷ Similar examples, particularly examples of collapsing the scores in the highest categories, are easy to find, even in this author's own work.⁸ The rationale for collapsing these categories is the same as the rationale for using an index of comorbid diseases: ease of comprehension and statistical efficiency. The effect of collapsing these categories is also the same as the effect of collapsing disparate comorbid diseases: introduction of classification errors. In this paper, we show analytically and by synthetic example that collapsing the high-end categories of a comorbidity scale changes the estimate(s) of effect(s) of comorbidity and biases analyses that control for comorbidity as a confounder or analyze modification of an exposure's effect by comorbidity.

Methods and results

To depict the bias introduced by collapsing categories of a comorbidity scale, we created a scale with a strictly monotonically increasing risk of the outcome (r_i) with each increase in the ordinal comorbidity scale (indexed by i), and a strictly monotonically decreasing prevalence (p_i) of the comorbidity value with each increase in the ordinal comorbidity scale. Table 1 depicts this synthesized data. While the data are synthetic, we note that the risks of an outcome in the scale categories and the prevalence of the scale categories correspond well with values one might observe. For example, one might anticipate similar data if the population was an older population (say 70-years-old and older), if the index of comorbidity was the Charlson Index, and the outcome was a three-year risk of death.

Collapsing categories changes the estimate of comorbidity's effect

The prevalence of the comorbidity categories decreases as the ordinal value increases. The prevalence of comorbidity category 4 is only 5%. In many data sets, the number of

Table 1 Depiction of the prevalence of comorbidity index categories (p_i) and the risk of an outcome (r_i) within the categories

Comorbidity category (i)	0	1	2	3	4
Prevalence (p_i)	0.35	0.30	0.20	0.10	0.05
Risk (r_i)	0.010	0.027	0.076	0.213	0.603

persons with this value would be small, and the number of cases of some outcome within that category even smaller. Analysts might be tempted to collapse category $i = 4$ with category $i = 3$, for example, to avoid sparse data problems or to improve the precision of the estimate of association in the highest comorbidity category. The effect of this collapse is to set the risk for the combined category to a weighted average of the two individual categories. More generally, collapsing a set of the upper-end categories ranging from $i = v$ to the maximum ($i = 4$, in this example) generates a weighted average risk:

$$r_{v..4} = \frac{\sum_{i=v}^4 p_i r_i}{\sum_{i=v}^4 p_i}$$

Table 2 depicts the risk ratios ($RR_{C=x \text{ vs } C=0}$) estimated from the synthetic data when the high-end categories are collapsed together. The collapsed categories range from some value v , which can equal 1, 2, 3, or 4 to the maximum (4, in this example). Setting $v = 4$ therefore corresponds to the case in which there is no collapse. Note that collapsing categories does not introduce a bias; the estimate of risk and therefore risk ratio within each category is an accurate depiction of the effect in that category. With each additional combination, the risk in the highest category becomes more heavily weighted with the low-risk comorbidity categories because these low risk categories are more prevalent. When $v = 1$, which corresponds to a comparison of any comorbidity (collapsing categories 1 to 4) with no comorbidity (category $i = 0$), the risk ratio equals 12. This risk ratio is about five-fold lower than the risk ratio in the highest comorbidity category ($i = 4$, in which the risk ratio equals 60) and about five-fold higher than the risk ratio in the lowest category with any comorbidity ($i = 1$, in which the risk ratio equals 2.7). The risk ratio of 12 is not, in fact, a very good estimate of the effect

Table 2 Risk ratios associating the presence of comorbidity, compared with the absence of comorbidity, with the outcome within collapsed categories (ranging from v to 4)

v	Index of comorbidity (i)				
	0 (Reference)	1	2	3	4
4	1	2.7	7.6	21	60
3	1	2.7	7.6		34
2	1	2.7		19	
1	1			12	

of comorbidity in any of the most finely divided categories. Collapsing comorbidity categories can therefore diminish the ability to discern important patterns that are more apparent when categories are not collapsed.

Collapsing categories biases the relative risk due to confounding by comorbidity toward the null

Comorbidity data are frequently collected to control for confounding by underlying health indications. That is, comorbid diseases are likely to be more prevalent among patients with high risk conditions (eg, another health indicator such as frailty or disability) and likely also to be related to the outcome under study (eg, all-cause mortality). A scale of comorbid disease is therefore often a potential confounder and a candidate for analytic adjustment.

To examine the effect of collapsing comorbid categories when the comorbidity scale is used for analytic adjustment, we postulated a second dichotomous variable (E indexed by $k = 0$ or 1 within categories of the comorbidity scale) whose association with the outcome is of primary interest. We assumed that the prevalence of $E = 1$ depends on the category of the comorbidity scale, as depicted in Table 3. We assumed, however, that the risk of the outcome did not depend on the category of E within strata of the comorbidity scale. That is, after adjustment for the most finely divided comorbidity categories, the risk ratio associating $E = 1$ compared with $E = 0$ would be null ($RR_{E=1 \text{ vs } E=0} = 1$).

The crude risk in categories of E is the weighted average of the risks in Table 1, where now the weights correspond with the prevalence of comorbidity within category of E , as shown in Table 3. That is

$$r_k = \frac{\sum_{i=1}^4 p_{i,k} r_{i,k}}{\sum_{i=1}^4 p_{i,k}}$$

The risk equals 0.125 in $E = 1$ and 0.032 in $E = 0$, which yields a crude $RR_{E=1 \text{ vs } E=0}$ of 3.90. The substantial departure of this crude risk ratio from the true null association

Table 3 Depiction of the prevalence of comorbidity index categories within categories of E ($p_{i,k}$)

Comorbidity category (i)	0	1	2	3	4
Prevalence in $E = 1$ ($p_{i,1}$)	0.15	0.30	0.30	0.15	0.10
Prevalence in $E = 0$ ($p_{i,0}$)	0.55	0.30	0.10	0.05	0.00

is entirely due to confounding by comorbidity. The relative risk due to confounding (RR_c), which equals the ratio of the crude and adjusted estimates, provides a measure of the direction and magnitude of this confounding, and in this case $RR_c = 3.9/1 = 3.9$. To resolve the confounding, one can calculate the standardized risk ratio ($sRR_{E=1 \text{ vs } E=0}$), where the standard weights equal the prevalence of the comorbidity categories in $E = 1$ (these weights are $p_{i,1}$). That is:

$$sRR = r_1 \left/ \frac{\sum_{i=1}^4 p_{i,1} r_{i,0}}{\sum_{i=1}^4 p_{i,1}} \right.$$

when comorbidity categories are collapsed, the standardized risk in the denominator uses the weighted average risk in the collapsed category ($r_{v...4,0}$, where the weights come from the unexposed group) and the sum of corresponding weights in the exposed category (sum from v to 4 of $p_{i,1}$).

$$sRR' = r_1 \left/ \frac{\sum_{i=1}^{v-1} p_{i,1} r_{i,0} + \sum_{i=v}^4 p_{i,1} r_{v...4,0}}{\sum_{i=1}^4 p_{i,1}} \right.$$

The resulting sRR is incompletely adjusted for confounding by comorbidity. Table 4 depicts the $sRR_{E=1 \text{ vs } E=0}$ and RR_c for this scenario, and in a second scenario with the true $sRR_{E=1 \text{ vs } E=0} = 1.5$. In both cases, collapsing the upper-end categories of the comorbidity scales yields incomplete control for confounding by comorbidity. The result is a bias of RR_c toward the null, which can give rise to the appearance of an association between E and the outcome when the true association is null (scenario 1), can give rise to the appearance of a stronger association than is truly present (scenario 2), or can give rise to an underestimate of the true association

Table 4 Depiction of the true $sRR_{E=1 \text{ vs } E=0}$ and RR_c ($v = 4$) and biased $sRR_{E=1 \text{ vs } E=0}$ and RR_c ($v < 4$) when comorbidity categories are collapsed

v	Scenario 1		Scenario 2	
	$sRR_{E=1 \text{ vs } E=0}$	RR_c	$sRR_{E=1 \text{ vs } E=0}$	RR_c
4	1.00	3.90	1.50	4.36
3	1.45	2.68	2.32	2.81
2	1.63	2.40	2.60	2.51
1	2.42	1.61	3.96	1.65

if the true $RR_c < 1$ and the true association between E and the outcome is causal.

Collapsing categories biases estimates of interaction unpredictably

Some analyses examine the interaction between comorbid disease and a second variable. These analyses investigate whether the effect of the exposure depends on the comorbidity category. Often the analysis compares the effect of the exposure in those with the highest comorbidity category to the effect of the exposure in those with the lowest comorbidity category. For example, one might calculate the interaction contrast (IC),⁹ which measures the departure of risk in those with the high risk category of the exposure ($E = 1$) and comorbidity ($I = 4$) from the risk expected given (a) the independent effect of the exposure in those without comorbidity ($r_{0,1} - r_{0,0}$), (b) the independent effect of higher comorbidity in those without the exposure ($r_{4,0} - r_{0,0}$), (c) the risk in those with the low risk category of the exposure ($E = 0$) and comorbidity ($I = 0$). This concept simplifies to the risk difference in those with high comorbidity less the risk difference in those without comorbidity. That is:

$$IC = r_{4,1} - (r_{0,1} - r_{0,0}) - (r_{4,0} - r_{0,0}) - r_{0,0} \\ = (r_{4,1} - r_{4,0}) - (r_{0,1} - r_{0,0})$$

A second measure of interaction is the ratio of the risk ratios, which we will call effect measure modification (EMM). That is:

$$EMM = \frac{\frac{r_{4,1}}{r_{0,1}}}{\frac{r_{4,0}}{r_{0,0}}}$$

When the highest categories of comorbidity are collapsed, however, $r_{4,1}$ will be replaced with $r_{v..4,1}$ and $r_{4,0}$ will be replaced with $r_{v..4,0}$. The result is an unpredictable bias in the estimates of the interaction between the exposure and comorbidity. In scenario 1, the exposure has no effect, so $r_{i,1} - r_{i,0} = 0$ and $r_{i,1}/r_{i,0} = 1$. Therefore, IC must equal 0 and EMM must equal 1. As depicted in Table 5, the collapsed categories ($v < 4$) all yield $IC > 0$ and $EMM > 1$, suggesting an interaction between E and comorbidity that does not exist. Furthermore, as v increases, the bias of IC decreases but the bias of EMM increases. In scenario 2, both the exposure and comorbidity affect the outcome. Collapsing the comorbidity categories can overestimate IC (when $v = 3$) or

Table 5 Depiction of the true IC and EMM ($v = 4$) and biased IC and EMM ($v < 4$) when comorbidity categories are collapsed

v	Scenario 1		Scenario 2	
	IC	EMM	IC	EMM
4	0	1	0.298	1
3	0.156	1.73	0.339	1.87
2	0.088	1.72	0.186	1.83
1	0.086	2.46	0.150	2.68

underestimate IC (when $v \leq 2$). On the other hand, EMM is most overestimated in scenario 2 when $v = 1$.

Discussion

The common practice of collapsing the highest categories of comorbidity into a single category has the advantages of increasing the prevalence of subjects in the highest category of comorbidity, thereby improving the ease of comprehension and the statistical efficiency of the analysis. These advantages, however, come at the price of misclassification of subjects. The impact of this misclassification depends on how the comorbidity variable is used in the analysis.

When comorbidity is an exposure or predictor of the outcome in the analysis, then the misclassification changes the pattern of the outcome response as a function of the “dose” of comorbidity. This result should be expected; miscategorization of dose – and in particular combining categories with dissimilar outcome risks – yields misleading dose-response patterns.¹⁰ Better analytic solutions are to collapse only adjacent comorbidity categories with similar risks⁶ or to use more sophisticated dose-response modeling, such as spline regression.¹⁰ The similarity of risks in adjacent categories is best left to judgment, perhaps informed by statistical testing, because of the poor power to detect important differences by statistical testing alone.¹¹

When comorbidity is a candidate confounder in the analysis, then the misclassification biases the relative risk due to confounding toward the null (assuming independent and nondifferential classification errors). The result is residual confounding of the association between the exposure of interest and the outcome. This result should also be expected; independent and nondifferential misclassification of a confounder is known to yield residual confounding.¹² Importantly, misclassification resulting from crude categorization of even a covariate that has been well-measured on a continuous scale can result in substantial bias.¹³ As above, better analytic solutions are to collapse only adjacent comorbidity categories with similar risks,⁶ to use spline regression,^{10,14} or to include

the comorbidity index as a single linear term in regression modeling.¹⁴ Restricting the study sample to subjects with comorbidity scores below the threshold where category collapsing will improve comprehensibility and statistical efficiency is also an alternative, although this restriction may reduce the generalizability of study results.¹⁵

When comorbidity is a candidate modifier in the analysis, then the misclassification can give rise to the appearance of interaction when no interaction exists, can mask true interaction, and can bias the estimate of interaction.³ Different combinations of these possibilities may appear depending on whether interaction is assessed as departure from additive or multiplicative effects, both of which have been proposed as important considerations in the examination of comorbidity.^{1,4} This result should also be expected; independent and nondifferential misclassification of a modifier is known to affect analyses of interaction unpredictably.¹² For most analyses of interaction, the best analytic solution is to restrict the analysis and inference to a category of comorbidity with uniform risk for the outcome.

The appeal of collapsing categories of comorbidity to facilitate interpretation and statistical analysis is often offset by misleading results. At a minimum, analysts should assure the uniformity of outcome risk in collapsed categories before collapsing them. Often times, more appropriate analytic methods can achieve similar goals without sacrificing the validity of the study's results.

Disclosure

The author reports no conflicts of interest in this work.

References

1. Yancik R, Ershler W, Satariano W, Hazzard W, Cohen HJ, Ferrucci L. Report of the National Institute on Aging Task Force on Comorbidity. *J Gerontol.* 2007;62A:275–280.
2. Feinstein AR. The pre-therapeutic classification of co-morbidity in chronic disease. *J Chron Dis.* 1970;23:455–468.
3. Lash TL, Mor V, Wieland D, Ferrucci L, Satariano W, Silliman RA. Methodology, design, and analytic techniques to address measurement of comorbid disease. *J Gerontol.* 2007;62A:281–285.
4. Karlamangla A, Tinetti M, Guralnik J, Studenski S, Wetle T, Reuben D. Comorbidity in older adults: Nosology of impairment, diseases, and conditions. *J Gerontol.* 2007;62A:296–300.
5. Boyd CM, Weiss CO, Halter J, Han KC, Ershler WB, Fried LP. Framework for evaluating disease severity measures in older adults with comorbidity. *J Gerontol.* 2007;62A:286–295.
6. Schneeweiss S, Maclure M. Use of comorbidity scores for control of confounding in studies using administrative databases. *Int J Epidemiol.* 2000;29:891–898.
7. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chron Dis.* 1987;40:373–383.
8. Lash TL, Thwin S, Horton NJ, Guadagnoli E, Silliman RA. Multiple informants: a new method to assess breast cancer patients' comorbidity. *Am J Epidemiol.* 2003;157:249–257.
9. Greenland S, Lash TL, Rothman KJ. Concepts of interaction. In: Rothman KJ, Greenland S, Lash TL, editors. *Modern Epidemiology*, 3rd edition. Philadelphia, PA: Lippincott, Williams and Wilkins; 2008.
10. Greenland S. Dose-response and trend analysis in epidemiology: alternatives to categorical analysis. *Epidemiology.* 1995;6:356–365.
11. Greenland S, Rothman KJ. Introduction to stratified analysis. In: Rothman KJ, Greenland S, Lash TL, eds. *Modern Epidemiology*, 3rd edition. Philadelphia, PA: Lippincott, Williams and Wilkins; 2008.
12. Greenland S. The effect of misclassification in the presence of covariates. *Am J Epidemiol.* 1980;112:564–569.
13. Brenner H. A potential pitfall in control of covariates in epidemiologic studies. *Epidemiology.* 1997;9:68–71.
14. Brenner H, Blettner M. Controlling for continuous confounders in epidemiologic research. *Epidemiology.* 1997;8:429–434.
15. Charlson ME, Horwitz RI. Applying results of randomised trials to clinical practice: Impact of losses before randomization. *Br Med J.* 1984;289:1281–1284.

