Scaling Back on Scales with a Scale of Scales

H.J. Cloft and D.F. Kallmes

AJNR Am J Neuroradiol published online 23 December 2010
http://www.ajnr.org/content/early/2010/12/23/ajnr.A2432.citation

Published December 23, 2010 as 10.3174/ajnr.A2432

EDITORIAL

Scaling Back on Scales with a Scale of Scales

An ever-increasing number of articles are published introducing clinical scales to describe neurovascular diseases. Unfortunately, unless you are some kind of idiot savant, there are now too many scales to remember. It would be helpful to have some way of knowing which scales are worth remembering and which are not. Most would agree that those worth remembering are those that are useful. A scale may be useful in a number of ways: First, it might allow us to predict outcomes for patients in our practice. Predicting outcomes helps us to counsel patients and choose the best therapy. Second, a scale might be useful if it can be used in clinical trials to objectively select patients to enroll or to provide objective descriptions of patient outcomes. For any scale to be useful for these purposes, however, it needs to be clinically relevant, valid, and easy to use. Unfortunately, many scales being published currently lack 1 or more of these qualities and are, therefore, not useful in the world of real patients. Thus, we propose a Scale of Scales to help you decide which are worth remembering and which are not (Table).

The Scale of Scales is founded on the fundamental principles of clinical relevance, validity, and reliability. "Clinical relevance" means that the scale informs us about something that is important to patient outcome. "Validity" means that it actually measures or describes what it is supposed to measure or describe. "Reliability" means that the scale is reproducible, with little variation among users of the scale. The Scale of Scales incorporates assessments of these key characteristics and assigns grades of I through V, with grade I being the most benign beneficial form of scale and grade V being the most malignant useless form of scale. In addition, we added subgrades a and b, to classify the scale as easy or difficult to remember and/or use.

Physicians are predisposed to memorize and follow rules, so it is perhaps easy to get us to blindly follow along in categorizing diseases according to some inane scale. We may think we are practicing medicine when we classify our patient's disease according to scales in the literature, but maybe we are, in fact, just engaging in a pointless pretense of understanding. Just because we can classify something with alphanumeric symbols does not mean that we understand it. The scale gives us an answer, but it is not always clear if it answers an important question. Scales and classifications can both make the simple seem complex and the complex seem simple, which could easily lead to distraction from relevant clinical issues. With the rapid growth of the medical literature, it is difficult for practicing physicians to keep up with important developments, so it is increasingly important that we not clutter our minds with scales of dubious value.

Classification scales of dural fistulas, carotid cavernous fistulas, and spinal vascular malformations and fistulas are abundant and redundant, as well as abundantly and redundantly confusing. Many of these classifications neither predict natural history nor guide therapy. They serve only to confuse conversation with coded language. It is preferable to simply state that the patient has a direct carotid cavernous fistula than to cryptically state that the patient has a Barrow type A fistula (a grade IIIa scale).[1] When we speak or write, we should strive to use terminology that people understand.

The Scale of Scales

Grades	Description
I	Clinically relevant, reliable, and valid
II	Clinically relevant but not yet validated or shown to be reliable
III	Valid and reliable, but not clinically relevant
IV	Clinically relevant but shown to be invalid or unreliable
V	Not clinically relevant, not simple to remember and use, and not validated
Subgrades	
a	Easy to remember and/or use
b	Not easy to remember and/or use

Our field has still not matured to the point that we have many well-developed scales, but papers describing new scales will be essential to progress. Many scales related to neurointerventions are based on angiographic appearances and attempt to divide a continuum of variability into discrete categories (ie, perform analog to digital conversion). Dr Tomsick[2] wrote eloquently about this problem in the conduct of stroke trials. We wrote about the reliability of angiography scales used in research of endovascular aneurysm treatments.[3] The purpose of these previous articles was to point out that new scales must be developed with respect for proper scientific methodology and the basic issue of clinical relevance, but these articles seem to have been largely ignored. For example, a scale of endovascular aneurysm coiling results has been put forth as a multisociety-approved reporting standard for future research,[4] but it has never been tested for reliability. Not only has it been neither tested nor used in any study of any kind, but an earlier study of aneurysm occlusion scales[3] indicates that such a new complex scale will undoubtedly have so much interobserver and intraobserver variability that it will be utterly useless. Thus, we would grade the newly proposed aneurysm occlusion grading scale as a grade IVb scale.

Two scales have been proposed this year to assess aneurysm appearance after placement of a flow-diversion device.[5,6] Both of these recently proposed scales offer descriptions of classification systems for angiograms but no correlation with any kind of clinical or radiographic outcome. It is hard to assign a grade to these scales, because we know so little about them. They might be grade II scales if the angiographic appearance immediately after flow-diversion treatment has any clinically relevant predictive value. A grade II scale has the potential to become a class I scale if it can be validated and shown to be reliable. A grade II scale is really just a good hypothesis waiting to be tested. Although it may be a good hypothesis, a grade II scale should not be published and certainly not memorized by the general public until it has been tested for validity and reliability.

Some scales simply tell us what we already know, which is really just translating common knowledge into an alphanumeric code. The Secondary Intracerebral Hemorrhage Score is an example of such a scale telling us something that we

already know, which is that patients who have an obvious vascular lesion on unenhanced CT as well as young patients are likely to have a definable vascular etiology for hemorrhage if we do further vascular imaging, whereas those with hypertension and/or on anticoagulation therapy are not.[7] Our Scale of Scales is similarly made up of valid concepts and thus also only tells us what we already know to be true.

We should strive to make the world a better place through scientific progress. Our grandchildren may someday read what our generation wrote in the scientific literature, and they will probably notice if the literature we leave behind is not only useless but a malignant waste of time to read. Before creating or propagating a scale or disease-classification system, we should ask ourselves if it is really going to help anyone in the future. Unfortunately, the perpetuation of useless scales is part of our medical culture. For example, the scale of carotid cavernous fistulas described by Barrow et al in 1985[1] has no real clinical utility, and yet a quick unscientific sampling of the literature shows us that it was nonetheless cited in 5 of 10 articles published about carotid cavernous fistulas from 2005 to 2010. For the common good, we should start actively trying not to generate or propagate bad scales. Our examples of bad scales all pertain to neurovascular diseases because that is our area of expertise, but bad scales have the potential of polluting all medical research.

So speaking of bad scales, how do we apply this Scale of Scales in the future? Grade I scales are good scales and are essential to the advancement of medicine. Grade II scales are hypotheses that should be tested for validity and reliability, and those results could be published. Grade III, IV, and V scales are all of no use to us and should be abandoned. We would suggest that the Scale of Scales is a grade IIIb scale, but some of you may disagree. We can accept that you may disagree because we are sure that there will be both inter- and intraobserver variability with the Scale of Scales. Nonetheless, because it is a grade IIIb scale, we strongly recommend that you do not bother to learn the Scale of Scales. We introduced the Scale of Scales simply to make the point that you can quickly dismiss it, just as you can quickly dismiss many of the scales published in the literature. Instead of learning the Scale of Scales, we encourage you to evaluate all scales carefully with regard to clinical relevance, validity, reliability, and if a scale does not have all 3 of these characteristics, you should waste no additional time thinking about it.

References

1. Barrow DL, Spector RH, Braun IF, et al. **Classification and treatment of spontaneous carotid-cavernous sinus fistulas.** *J Neurosurg* 1985;62:248–56
2. Tomsick T. **TIMI, TIBI, TICI: I came, I saw, I got confused.** *AJNR Am J Neuroradiol* 2007;28:382–84
3. Cloft HJ, Kaufmann T, Kallmes DF. **Observer agreement in the assessment of endovascular aneurysm therapy and aneurysm recurrence.** *AJNR Am J Neuroradiol* 2007;28:497–500
4. Meyers PM, Schumacher HC, Higashida RT, et al. **Reporting standards for endovascular repair of saccular intracranial cerebral aneurysms.** *AJNR Am J Neuroradiol* 2010;31:E12–24
5. Kamran M, Yarnold J, Grunwald IQ, et al. **Assessment of angiographic outcomes after flow diversion treatment of intracranial aneurysms: a new grading schema.** *Neuroradiology* 2010; Sep 14 [Epub ahead of print]
6. O'Kelly CJ, Krings T, Fiorella D, et al. **A novel grading scale for the angiographic assessment of intracranial aneurysms treated using flow diverting stents.** *Interv Neuroradiol* 2010;16:133–37
7. Delgado Almandoz JE, Schaefer PW, Goldstein JN, et al. **Practical scoring system for the identification of patients with intracerebral hemorrhage at highest risk of harboring an underlying vascular etiology: the Secondary Intracerebral Hemorrhage Score.** *AJNR Am J Neuroradiol* 2010;31:1653–60. Epub 2010 Jun 25

H.J. Cloft
Senior Editor
D.F. Kallmes
Editorial Board

DOI 10.3174/ajnr.A2432