Clinician's Commentary on Dang et al.[1]

Dang et al.'s article in this issue of *Physiotherapy Canada* investigates whether the predictive equations from the well-known and widely used Chedoke-McMaster Stroke Assessment (CMSA) are accurate enough for prognostication.[1] The reader is likely familiar with this assessment approach, which is based on classifying the recovery of hand, arm, leg, foot, and postural control into seven categories that correspond to seven stages of motor recovery.[2] The aim of this classification system is to predict what stage of recovery the person with stroke is likely to attain, so that discharge planning from rehabilitation can be optimized and treatments can target progression to higher stages.

Re-evaluating these regression equations is a good idea for several reasons. First, at the time the original studies were published in the 1980s, the statistical assumptions underlying the use of statistical models were poorly understood by clinical researchers. Interestingly, I was the original reviewer of the Gowland's 1984 article,[3] and at the time raised statistical issues about the model's assumptions (I was then a PhD student in the Department of Epidemiology and Biostatistics). Using a regression equation to predict a future stage was considered quite novel in the physiotherapy field at the time, and the work on the development and testing of the CMSA was otherwise exemplary; for this reason, these first equations were published. Revised equations were subsequently generated on a new sample of 182 patients, and these were reported in the CMSA manual, published in 1995.[4]

In the decades that followed Gowland's original publication,[3] new theories of stroke recovery emerged, including motor control and task practice, which raised questions about the need to think in terms of stages of recovery and synergies.[5,6] This era also brought modern psychometric methods to develop and refine measures of motor skills, and the rehabilitation field became richer for the application of Rasch modelling and Item Response Theory to retrofit existing measures and develop new ones.[7–16]

Now is the time to re-examine our existing measures in the light of changes to theory and to statistical and psychometric methods, as Dang et al.[1] have done by undertaking a revalidation of the predictive equations from the CMSA manual.

The authors' methodology was essentially to apply the original equations to new data. For the impairment outcomes—hand, arm, foot, leg, shoulder pain, and postural control—the old equations showed acceptable "fit" to the new data (i.e., there was little "shrinkage"), but with very wide confidence bands, indicating that each prediction could be off by ± 2 stages. Since the Impairment Inventory is scored on a seven-point scale (1 = flaccid paralysis; 2 = spasticity is present and felt as a resistance to passive movement; 3 = marked spasticity but voluntary movement present within synergistic patterns; 4 = spasticity decreases; 5 = spasticity wanes but is evident with rapid movement at the extremes of range; 6 = coordination and patterns of movement are near normal; 7 = normal movement), this degree of error is too large to provide a confident prediction of a patient's outcome. The fit of the equations for gross motor function, walking, and activity greatly exceeded the threshold for reproducibility (i.e., showed high shrinkage), indicating that prediction of these domains is poor.

Several assumptions must be satisfied for correct interpretation of the results of a linear regression model.[17] Linear regression is used when the outcome of interest is measured on a continuous scale (i.e., when it can theoretically have any value within a range). For the CMSA impairment inventories, however, the outcome is not continuous but ordinal (i.e., 0, 1, 2, 3, etc.). Gowland originally identified this as a statistical issue, but in 1984, doing ordinal regression was not a realistic option, given the existing software.[18,19] Ordinal regression is now part of standard statistical packages, however. The key assumption of linear regression is *linearity*—that is, that the relationship between the outcome (discharge status) and the predictor (baseline status) is linear. In the study by Dang et al., this assumption did not hold for the hand and arm domains, and the authors nicely show the curvilinear relationship.[1]

These are not the only assumptions of linear regression. The distribution of the outcome variable (discharge stage) needs to be normal within levels of the predictor variable (baseline stage); information on this assumption is not presented. In addition, predictors are fixed variables, and are optimally measured without error; if there is measurement error on the predictors, these errors must not be correlated with the measurement error of the outcome. In this case, correlated errors exist, as measurements at baseline and discharge stages share the same sources of imprecision.

For the most part, having the wrong model is expected. According to the well-known statistician George Box, "Essentially, all models are wrong, but some are useful."[20(p.424)] The usefulness of the CMSA model appeared on my radar again when, in designing guidelines for the SCORE project,[21] the team considered using the CMSA prediction equations as a way to decide what type of

therapy a person in rehabilitation for stroke would receive. If treatment decisions are made based on these prediction equations, people at CMSA stages ≤2—those predicted not to progress—would not receive "active" task-oriented treatment but, rather, would be taught compensatory strategies. At an investigators' meeting, I presented data from a study conducted by our research group[14,15] that included data on the CMSA and other measures. The study sample consisted of data from 235 people at 3–7 days and 3 months post stroke; we found that 44% of people measured at stage 2 or below for the arm 3–7 days post stroke did in fact progress to stage 4 or beyond by 3 months (versus only 11% for people at stage 1). This proportion is equivalent to a "number needed to treat" (NNT) of 2.3—meaning, essentially, that 2.3 people at stage 2 would need to be treated for 1 person to progress. This seems to me a fairly high yield of good outcomes for the effort invested. In addition, people at this stage also made clinically relevant gains in terms of other important outcomes.

My conclusion from these data was that there is insufficient evidence to recommend changing current practice and withholding active treatment from people at CMSA stage 2, unless, of course, we assume that all the progress is made between 3 days post stroke and the start of rehabilitation (at about 18 days post stroke), after which no further progress will occur.

We subsequently also applied the original regression equations to this data set, and found that the R^2 for the arm stage was 0.61—showing a shrinkage of 0.18, past the threshold for reproducibility. Shrinkage for the other impairment inventories ranged from 0.18 to 0.31, except for postural control, which showed no shrinkage.

Dang et al.[1] are to be congratulated for "myth busting" with respect to prediction using the CMSA. Many of our measures are undergoing revision in light of modern psychometric and statistical methods, and ultimately this effort will produce better measures. Better measures will lead to better treatments, because—to quote Lord Kelvin—"If you cannot measure it, you cannot improve it."*

* The Rt Hon Lord Kelvin was first president of the International Electrotechnical Commission (IEC), founded in 1906.

Nancy E. Mayo, BSc(PT), MSc, PhD
James McGill Professor
Fellow of the Canadian Academy of Health Sciences
Department of Medicine
School of Physical and Occupational Therapy
McGill University
Division of Clinical Epidemiology,
Division of Geriatrics
McGill University Health Center Royal Victoria Hospital Site
Ross Pavilion R4.29
687 Pine Ave W
Montreal, QC, H3A 1A1
Tel. (514) 934 1934 36922
FAX (514) 843 1493
nancy.mayo@mcgill.ca

REFERENCES

1. Dang M, Ramsaran KD, Street ME, et al. Estimating the accuracy of the Chedoke–McMaster Stroke Assessment predictive equations for stroke rehabilitation. Physiother Can. 2011;63(3):334–41. doi:10.3138/ptc.2010-17
2. Brunnstrom S. Movement therapy in hemiplegia: a neurophysiological approach. New York: Harper & Row; 1970.
3. Gowland C. Predicting sensorimotor recovery following stroke rehabilitation. Physiother Can. 1984;36:313–20.
4. Gowland C, Van Hullenaar S, Torresin W, et al. Chedoke-McMaster Stroke Assessment: development, validation, and administration manual. Hamilton (ON): McMaster University; 1995.
5. Carr JH, Shepherd RB. A motor learning model for rehabilitation. In: Carr JH, Shepherd RB, Gordon J, et al., Editors. Movement science: foundations for physical therapy in rehabilitation. London: Heinemann; 1987. p. 31–91.
6. Woodbury M, Velozo CA, Thompson PA, et al. Measurement structure of the Wolf Motor Function Test: implications for motor control theory. Neurorehabil Neural Repair. 2010;24(9):791–801. Epub 2010 Jul 8. doi:10.1177/1545968310370749
7. Merbitz C, Morris J, Grip JC. Ordinal scales and foundations of misinference. Arch Phys Med Rehabil. 1989;70(4):308–12. Medline:2535599
8. Stucki G, Daltroy L, Katz JN, et al. Interpretation of change scores in ordinal clinical scales and health status measures: the whole may not equal the sum of the parts. J Clin Epidemiol. 1996;49(7):711–7. doi:10.1016/0895-4356(96)00016-9. Medline:8691219
9. Duncan PW, Bode RK, Min Lai S, et al.; Glycine Antagonist in Neuroprotection Americans Investigators. Rasch analysis of a new stroke-specific outcome scale: the Stroke Impact Scale. Arch Phys Med Rehabil. 2003;84(7):950–63. doi:10.1016/S0003-9993(03)00035-2. Medline:12881816
10. Hsueh IP, Wang WC, Sheu CF, et al. Rasch analysis of combining two indices to assess comprehensive ADL function in stroke patients. Stroke. 2004;35(3):721–6. doi:10.1161/01.STR.0000117569.34232.76. Medline:14963275
11. Haley SM, McHorney CA, Ware JE Jr. Evaluation of the MOS SF-36 physical functioning scale (PF-10): I. Unidimensionality and reproducibility of the Rasch item scale. J Clin Epidemiol. 1994;47(6):671–84. doi:10.1016/0895-4356(94)90215-1. Medline:7722580
12. Jenkinson C, Fitzpatrick R, Garratt A, et al. Can item response theory reduce patient burden when measuring health status in neurological

disorders? Results from Rasch analysis of the SF-36 physical functioning scale (PF-10). J Neurol Neurosurg Psychiatry. 2001;71(2):220–4. doi:10.1136/jnnp.71.2.220. Medline:11459897

13. Finch L, Higgins J, Wood-Dauphinee S, et al. Development of a measure of functioning for stroke recovery: the functional recovery measure. Disabil Rehabil. 2008;30(8):577–92. doi:10.1080/09638280701355892. Medline:17852294
14. Finch LE, Higgins J, Wood-Dauphinee S, et al. A measure of early physical functioning (EPF) post-stroke. J Rehabil Med. 2008;40(7):508–17. doi:10.2340/16501977-0200. Medline:18758666
15. Finch LE, Higgins J, Wood-Dauphinee SL, et al. A measure of physical functioning to define stroke recovery at 3 months: preliminary results. Arch Phys Med Rehabil. 2009;90(9):1584–95. doi:10.1016/j.apmr.2009.03.016. Medline:19735788
16. Higgins J, Finch LE, Kopec J, et al. Development and initial psychometric evaluation of an item bank created to measure upper extremity function in persons with stroke. J Rehabil Med. 2010;42(2):170–8. doi:10.2340/16501977-0501. Medline:20140414
17. Kleinbaum DG, Kupper LL, Muller KE. Applied regression analysis and other multivariable methods. 2nd ed. Boston: PWS-Kent; 1988.
18. Begg CB, Gray R. Calculation of polychotomous logistic regression parameter using individualized regressions. Biometrika. 1984;71(1):11–8. doi:10.1093/biomet/71.1.11
19. Scott SC, Goldberg MS, Mayo NE. Statistical assessment of ordinal outcomes in comparative studies. J Clin Epidemiol. 1997;50(1):45–55. doi:10.1016/S0895-4356(96)00312-5. Medline:9048689
20. Box GEP, Draper NR. Empirical model building and response surfaces. New York: Wiley; 1987.
21. SCORE (Stroke Canada Optimization of Rehabilitation through Evidence). Canadian Stroke Network [cited 2011 May 6]. Available from: http://www.canadianstrokenetwork.ca.

DOI:10.3138/physio.63.3.342