

# ■ Making Sense of Statistics for Family Practitioners ■

## “Understanding the score”

Durrheim DN, Ogunbanjo GA

A plethora of scoring systems have evolved in medicine. Ranks or scores are a common feature of quantitative research conducted by family practitioners. In addition, ranking is prevalent in the laboratory, e.g. malaria parasitaemia ranked as 0, +, ++, or +++, and at the patient's bedside, that is, grading the severity of stroke or leprosy lesions as mild, moderate or severe (+, ++, +++) or of a cancer as Stage I, II, III or IV. A convenient way of combining the ranks of a group of related criteria is the creation of an “index”. Basic indices result from simply adding the individual ranks of all components, while more sophisticated indices provide individual weightings to each component before they are summed. It is useful to create indices but they should be used with circumspection. We will use the Apgar scoring system, which is used for the evaluation of newborn babies, to highlight features commonly shared by indices that should always be explicitly considered before use.

You will recall that in an earlier article in this series, we clearly distinguished between categorical and discrete numerical data. Discrete numerical data legitimately lend themselves to mathematical manipulation, for example four naevi are twice as many as two naevi, and the difference between one and two naevi is the same as the difference between two and three naevi. It is obvious that we cannot however say that a Stage IV cancer is twice as bad as a Stage II cancer, or that the difference between a Stage II and Stage III cancer is the same as the difference between a Stage III and a Stage IV cancer. Although this is clear for “categorical” data, when ranks are combined into a score, the score is often incorrectly treated as discrete “numerical” data.

Dr. Virginia Apgar developed the Apgar scoring system in 1952 to systematically evaluate the condition of neonates at Sloane Hospital for Women in New York City, at 1 and 5 minutes after birth.<sup>2</sup> The Apgar scoring system has enjoyed global application for predicting infant survival and

for assessing the effectiveness of resuscitation efforts. The score consists of adding the rating: 0, 1 or 2 for each of five clinical variables (Table I). A score of less than 3 at one minute is considered very bad, while a score of greater than or equal to 7 is considered good. A recent investigation concluded that the “Apgar scoring system remains as relevant for the prediction of neonatal survival today as it was almost 50 years ago”.<sup>3</sup> However, why should we be a little wary of the Apgar score and other similar scoring systems?

Firstly, most of the clinical signs are somewhat subjective, e.g. “weak” vs. “strong”, “some movement” vs. “cry”, etc and thus open to individual interpretation. Secondly, the score treats each factor as of equal importance with no differential weighting. But, this is invalid, as limpness in a baby has a different prognostic value to absence of a heartbeat at five minutes. Finally, differences between 0 and 1, or 1 and 2 may be treated as though they are equally important. Prognosis however differs markedly between absent respiratory effort (0) and a weak cry (1), and between the latter and a strong cry (2).

From these brief comments on a well known scoring system, it serves to caution family practitioners against impulsive combination of ordered categorical data into indices that may not offer greater insight than careful consideration of the individual factors. In particular we should be wary of falling into the trap of treating categorical data as numerical data.

### References

1. Durrheim DN, Ogunbanjo GA. Making sense of statistics for family practitioners: “Well proportioned”. *S A Fam Pract* 1999; 21(2): 55-56.
2. Apgar V. A proposal for a new method of evaluation of the newborn infant. *Curr Res Anesth Analg* 1953; 32: 260-267.
3. Casey BM, McIntire DD, Leveno KJ. The continuing value of the Apgar score for assessment of newborn infants. *N Eng J Med* 2001; 344: 467-471.

**Table I:** Apgar scoring system of newborn babies

Sign	Score		
	0	1	2
Heart rate	Absent	< 100	> 100
Respiratory effort	Absent	Weak cry	Strong cry
Reflex irritability	Absent	Some movement	Cry
Muscle tone	Limp	Some flexion	Well flexed
Colour	Blue/pale	Pink body, blue extremities	Completely pink

Professor David N. Durrheim MBChB, DTM&H, DCH, FACTM, MPH & TM, MACTM School of Public Health and Tropical Medicine, James Cook University Australia & Consultant: Communicable Disease Control, Department of Health, Mpumalanga.

Professor Gboyega A. Ogunbanjo MBBS, MFGP (SA), M Fam Med (MEDUNSA) Department of Family Medicine and Primary Health Care, Medunsa & School of Public Health and Tropical Medicine, James Cook University Australia.