**TECHNICAL CORNER: WHY IS IT SO HARD TO STAGE SCORE RECORDINGS WITH FREQUENT APNEA?**
By Richard S. Rosenberg, PhD

Scoring an inter-scorer reliability record can cause considerable distress. A record was chosen from a 45-year-old man undergoing a diagnostic study with waveforms that were, for the most part, well defined and easy to categorize. Yet the overall agreement for stage scoring with the gold standard scorers (full disclosure – the author is one of the gold standard scorers for this record) was only 60.6%.

Daytime sleepiness is a hallmark of obstructive sleep apnea (OSA) and is included in the diagnostic criteria for patients with an apnea–hypopnea index between 5 and 15. Apneas are often associated with arousals, and arousals are associated with excessive daytime sleepiness. Some authors point to increases in stage N1 sleep with OSA, which is also associated with arousals, as an alternative cause of daytime sleepiness. Whatever the causes and consequences of sleepiness associated with arousal, it is clear that sleep is fragmented in patients with OSA and scoring is difficult.

When compared to scoring using Rechtschaffen and Kales, one of the major changes in sleep stage scoring with the AASM Manual for the Scoring of Sleep and Associated Events is the inclusion of this rule:

5.C.1. End stage N2 sleep when 1 of the following events occurs:

Subpart b: An arousal (change to stage N1 until a K complex unassociated with an arousal or a sleep spindle occurs)

Grigg-Damberger wrote that, “The 5.C.b. rule decreased NREM 2 time, increased NREM 1 time, and provided a subtle measure of sleep fragmentation by signaling a sleep stage shift.” The rule also increases scoring time and uncertainty. Scoring agreement for stage N1 sleep is typically the lowest of all stages.

One scorer indicated that her interpretation of this rule is that if there is an arousal followed at some point in an epoch by a K complex or sleep spindle then the epoch should be scored as stage N2 sleep. But the text of the rule states that after an arousal “change to stage N1 until a K complex … or a sleep spindle occurs.” This leads to the use of rule 2.B.3, which states that “if 2 or more stages coexist during a single epoch, assign the stage comprising the greatest portion of the epoch.” Having 3 stages in a single epoch occurs frequently in patients with OSA – the arousal is considered stage W, there is a period of time before the K complex or spindle that is scored as stage N1, and scoring of stage N2 begins with the K complex or spindle. Use of 2.B.3 requires a second-by-second scoring of the epoch. See Figure 1.

Consider the epoch in Figure 1. It begins with a K complex (A) and the first few seconds are scored as stage N2. An arousal occurs at the end of the respiratory event and several seconds of alpha rhythm (B) lead to the scoring of stage W. The alpha rhythm attenuates and is replaced by low amplitude mixed frequency EEG (C) resulting in a portion of the epoch scored as stage N1. But then a spindle is seen (D) resulting in the start of stage N2, as required by 5.C. 1.b. Proper scoring of this epoch requires counting the number of seconds of stages W, N1 and N2. This epoch was scored as stage W by 5.5%, stage N1 by 53.9%, stage N2 by 31.4%, stage N3 by 0.1% and stage R by 9.1% of scorers. I scored this epoch as stage N1. See Figure 2.

Now consider the epoch in Figure 2. The epoch begins with sleep spindles and the first few seconds are scored as stage N2. An arousal occurs but it is not associated with clear alpha rhythm activity. Two K complexes are embedded in the faster EEG activity. The subsequent recording includes a rudimentary sleep spindle and a waveform that appears to be a K complex but is most prominent in the central channel and poorly seen in the frontal channel. There are no firm rules in the AASM Manual for dealing with an epoch such as this. Should the portion of the epoch that comprises the arousal be scored as stage W? Should the K complexes during the arousal be used to start stage N2 scoring? Do the rudimentary waveforms meet criteria for a spindle and/or K complex and therefore restart stage N2? Most people scored this epoch as stage N2 sleep (30.0% N1, 69.5% N2). I went with stage N1 again. There really is no “correct” answer for this epoch. It depends on how you look at it.

The complexity of sleep scoring decisions clearly rises with the number of apneas. As a sleep center director I told my sleep technologists to “fly through the record” and score using a gestalt (mental picture) for sleep stages. One technologist took this to mean that he should listen to Ozzy Osborne at full volume when he scored records. He said that it helped him clear his mind. As a gold standard scorer I attempt to follow the AASM Manual rules as closely as possible. I measure every slow wave, contemplate each sleep spindle and parse the period where alpha activity is attenuated and replaced by low amplitude mixed frequency EEG. From a review of the results of an inter-scorer reliability program, it appears that whatever method scorers are using produces results that are similar to the level of agreement reported in the literature for two experienced scorers. Focused attention may not help with scoring of some epochs that have no correct answer. With some abnormal records, it may not matter whether you are scoring with Crazy Train on an iPod cranked up to 11 or in pristine silence with a copy of the AASM Manual by your side.

RICHARD S. ROSENBERG, PhD

Richard S. Rosenberg, PhD, has been in the sleep field since 1972 and currently lives in Long Beach, Calif., where he works remotely as AAST coordinator.
REFERENCES


Figure 1

Figure 2