

Stick your Neck Out, but Be Careful Where you Stick it

**Jerome R Hoffman,
MA MD**

**William R Mower,
MD PhD**

From the UCLA Emergency
Medicine Center, and the
Department of Medicine,
UCLA School of Medicine

Corresponding Author:
Jerome R Hoffman
UCLA Emergency Medicine
Center
10833 Le Conte Avenue
LA CA 90024, USA
jrh@ucla.edu

Emergency physicians (EPs), fearful of missing a rare patient with occult cervical spine injury, have traditionally used liberal imaging in the evaluation of blunt trauma patients. Until recently most EPs felt compelled to order films even for patients about whom they felt certain there was essentially no risk of injury. Although ostensibly "safe," such an approach has substantial costs, in terms of time, money and radiation exposure. Fortunately, two decision instruments have recently been developed which appear to identify successfully patients who have very low risk of cervical spine injury and, consequently, no need for radiography.

In the February issue of IJEM (1), Dickinson and Stiell comment on the two different instruments, one of which, the Canadian C-Spine Rule (CCR) they derived in Canada (2), and the other of which (NEXUS) we derived (3) and validated (4) in the United States. For reasons that are unclear to us, they take this opportunity to suggest that EPs - ostensibly uncomfortable with a sudden embarrassment of riches - need to *choose* one or the other. Not surprisingly, they then go on to criticize NEXUS, and thus imply that the CCR should be the one chosen.

We believe that EPs should be aware of the evidence that is available in support of each of the decision aids, and - *if* the Canadian instrument is validated in a separate group of patients - thus be able to take advantage of their separate strengths as each EP feels is most appropriate for a given patient or environment. There is no obvious reason why one has to be "chosen" and the other discarded, by all EPs, or in all cases.

That being said, we feel it is important to comment on a number of specific claims and, indeed, misstatements made by Dickinson and Stiell. First, they assert that "following the NEXUS criteria in Canada would require "a 50% increase in the rate of C-spine radiography." This is patently ridiculous, as it is based on a fundamentally false assumption about the differences in specificity *measured* in the two separate studies. As Dickinson and Stiell note, NEXUS included only patients in whom C-spine imaging was already being performed, while the CCR included many patients who did not undergo imaging. If we had included such patients in our calculation of specificity, in precisely the manner done in the Canadian study, our *measured* specificity would have increased accordingly. Furthermore (contrary to the assertion by Dickinson and Stiell), we published strong circumstantial evidence that there were no missed fractures among non-imaged patients seen at the NEXUS sites, so the vast majority (and probably all) such patients were almost certainly true negative according to the NEXUS decision instrument. (Even in the unlikely event that there was a missed injury in one or several such patients, it would have no impact on specificity and thus would not effect this particular discussion.) While it is true that we cannot state with certainty how many patients were seen and evaluated without x-rays at the 21 NEXUS hospitals, it is likely that the number was close to that in the Canadian study, given that among patients who were imaged in the two studies the percentage who proved to have an injury was very

similar. If so, the *true* impact of the two instruments in decreasing imaging is also very likely to be the same. In fact, the Canadians claim to be able to decrease imaging by 10.7% from their baseline rate of 68.9%, which is again similar to the 12.6% decrease, *from baseline*, found in NEXUS.

Dickinson and Stiell also complain that the NEXUS criteria are imprecise and cannot be applied widely because it is impossible to know what NEXUS EPs were thinking when they applied our individual criteria. This is, contrary to their assertions, a distinct strength of the NEXUS methodology. First of all - and most important - the criteria worked - regardless of how the EPs came to their individual conclusions! There were hundreds of EPs involved in NEXUS. They came from enormously varied backgrounds and brought very diverse training and experience to this task. Nevertheless, they applied the instrument successfully in an enormous cohort of patients. Had NEXUS failed, we could speculate about whether it was due to *misapplication* of our criteria rather than any innate defect of the instrument itself. Indeed, one of the two “missed” clinically significant injuries in our series appeared to be due to just such misapplication, as there was ample evidence on that patient’s chart that two of the five NEXUS criteria were indeed present. (The other “miss” appeared to involve an old fracture in a patient who did well despite refusing immobilization.)

In fact, NEXUS was overwhelmingly successful, with near-perfect sensitivity and even higher negative predictive value, *despite* the fact that it was fundamentally a real-world study and that it relied on the judgment of many different EPs. Readers may not know exactly what each of those EPs was thinking in each and every case, but they do know that a great number of EPs, in many different environments, were able to apply the criteria with virtually perfect sensitivity. Thus they can feel confident that other EPs like themselves, outside the NEXUS sites, will almost certainly be able to do the same. The CCR, on the other hand, was derived in a group of large teaching hospitals. Even if it is validated in these centers, it will still need to be tested in many other environments before clinicians can be confident of its real-world application. This is in distinct contrast to the proven wide clinical applicability of the NEXUS instrument.

Several other points need to be made with regard to the precision of the criteria used in the two studies. First, the claim by Dickinson and Stiell that “the inter-observer agreement on ... two (NEXUS) items was found to be too poorly reproducible for them to be usefully incorporated into a clinical decision rule” is inaccurate, as all five of our criteria, and the rule as a whole, were proven to have very high inter-rater reliability (5). More importantly, the proof is in the pudding! The fact that hundreds of different EPs applied the NEXUS criteria successfully ultimately provides far greater evidence of the decision instrument’s reliability than any testing of a surrogate marker like inter-observer agreement.

In addition, we do not believe there is any meaningful way to define our criteria with the type of “precision” Dickinson and Stiell seem to prefer. Most EPs know that some long-bone fractures can be relatively painless, while some patients appear to focus all their attention on a small abrasion (or road rash). Thus attempting to define a “distracting painful injury” based on a list of precise findings would be impossible. Furthermore, it would be self-defeating, for even if it were possible, it would make use of the entire instrument unfeasible, because the list would have to be far too complex for busy practitioners to apply (which may indeed be a problem for CCR, with its complicated algorithm). Once again, this would be a major problem for NEXUS if

different practitioners *failed* when they tried to implement our criteria, or if it only succeeded in a special group of EPs (like those who work at academic centers); the fact that such a large and various group of EPs *succeeded* in their application of the NEXUS instrument, in over 34,000 patients, including over 800 with cervical spine injury, can only increase the confidence we can have that other real-world physicians will be able to apply these clinical criteria to other real-world patients with similar success.

Finally, we are struck by the implication that (unlike NEXUS) the Canadian criteria are completely "objective." For example, their definition of "dangerous mechanism" may at first glance appear precise, but in what percentage of patients in clinical practice does anyone actually *know* if they fell from a height of two and a half feet, as opposed to three and a half feet? How reliable are patients (or paramedics or witnesses) at estimating the speed of vehicles at the time of impact? Isn't there also a need for judgment in deciding whether some version of "yes, doc, my right shoulder and arm also feel funny" does or doesn't represent "paresthesia?" What is the exact interpretation of the CCR exclusion criteria of "minor injuries" or "known vertebral disease?"

Not that this is a problem - almost all of what we do in medicine is based on some interpretation of historical and physical findings and any attempts to improve our skills *must* also include some reliance on our ability to do this well. To the extent that we can simplify this process, we may be able to standardize and even improve skills, but we are only fooling ourselves if we claim that we can take "subjectivity" out of what we do.

Simplicity is indeed another key difference between the two instruments, as Dickenson and Stiell acknowledge. Readers can judge for themselves, but we take strong issue with their contention that "these (NEXUS) criteria are too simplistic to be of practical value in clinical practice outside of the litigious medical atmosphere of the United States." Quite the contrary - the simplicity of the NEXUS instrument makes it likely that physicians *everywhere* will be both able and willing to use it in their own practice (just like the large and diverse group of EPs who successfully applied it in our study). On the other hand, there is plenty of evidence that most clinicians are unwilling and unable to apply complex algorithms in busy patient care settings. *If* the CCR is validated, and then *if* it is successfully applied in a different *study setting*, it would still need to be shown to work in the real world, where for "novice users" (whom we prefer to think of as busy clinicians) it would be "best (if they) followed ... a (complicated) chart until one is familiar with the definitions of each component."

Dickinson and Stiell conclude that "*when validated* the CCR should be an effective tool" (italics added). We find this a strange statement, coming from someone who has been so instrumental in defining how decision instruments should be *developed* and tested (6). Readers need to know that the methodology involved in the derivation of such instruments *insures* extremely high sensitivity - because the data is dredged to find a set of criteria that together identify all the patients with injury. Thus it is absolutely mandatory to *validate* the instrument in a separate cohort of patients before claiming that it works (even aside from questions of wider applicability). What *may happen* if further study is successfully completed represents reasonable speculation; what *will happen when*, with regard to claims that cannot possibly be substantiated, should be strenuously avoided.

The NEXUS decision instrument is a fairly simple tool based on clinical criteria. In a validation set involving a huge cohort of real-world patients presenting to just about every type of facility and applied by many hundreds of diverse physicians, it had near-perfect sensitivity, and would have decreased imaging by a small but meaningful amount. The Canadian C-spine Rule, derived in a homogeneous set of teaching hospitals, is far more complicated. *If* it is successfully validated in that same set of facilities, clinicians may wish to consider whether they can or are willing to apply it in their own practice. If so and at that time, they would have the luxury of considering two different instruments to help them identify a small but important group of very low risk patients in whom imaging of the C-spine can be safely avoided.

References

1. Dickinson G, Stiell IG. Sticking Your Neck Out: Choosing a Decision Rule for Cervical Spine Radiography. *IJEM* 2003;1(3):17-21
2. Stiell IG, Wells GA, Vandemheen K, Clement C, Lesiuk H, De Maio VJ et al. The Canadian Cervical Spine Radiography Rule for alert and stable trauma patients. *JAMA* 2001;286(15):1841-1848
3. Hoffman JR, Schriger DL, Mower W, Luo JS, Zucker M. Low-risk criteria for cervical-spine radiography in blunt trauma: a prospective study. *Ann Emerg Med* 1992;21:1454-1460
4. Hoffman JR, Mower W, Wolfson AB, Todd K, Zucker M. Validity of a set of clinical criteria to rule out injury to the cervical spine in patients with blunt trauma. *N Engl J Med* 2000;343(2):94-99
5. Mahadevan S, Mower WR, Hoffman JR, Peeples N, Goldberg W, Sonner R: Inter-rater Variability of Cervical Spine Injury Criteria in Blunt Trauma Victims. *Ann Emerg Med* 1998;31:197-201
6. Stiell IG, Wells GA. Methodologic standards for the development of clinical decision rules in emergency medicine. *Ann Emerg Med* 1999;33:437-447

Response

C-Spine Decision Rules

Ian G. Stiell, MD, MSc, FRCPC & Garth Dickinson, MD, FRCPC

Department of Emergency Medicine and Ottawa Health Research Institute, University of Ottawa, Ottawa, Canada

We appreciate the response of Drs. Hoffman and Mower to our commentary (1,2). We agree that emergency physicians are blessed to have gone from rags to riches in having guidance for ordering C-spine radiography, with not one, but two clinical decision rules for C-spine radiography in blunt trauma patients. Unfortunately, these two clinical decision instruments are not identical and sometimes produce conflicting advice on whether to obtain C-spine radiography in individual patients. Emergency physicians do have to choose - do they follow one or the other, both or neither to assess if C-spine radiography is needed?

The NEXUS study is a landmark study in emergency medicine and should be the standard of care for ordering C-spine radiography in the US. However, because of different practice patterns and standards of care, we are not certain that the NEXUS criteria are sufficiently sensitive or specific to be adopted in Canada (where the rate of C-spine radiography is much lower). We would also suggest that prior to accepting and adopting NEXUS as a global standard, validation outside of the US is required. This applies equally to the CCR and any other clinical decision rule. It took many years and