Accuracy of Ottawa Ankle Rules to exclude fractures of the ankle and mid-foot: systematic review.

Why Ottawa Ankle Rules?

- Never heard of them besides the internship
- Used to exclude fractures
- Many articles about it
- Important clinical relevance
- Easy to use
The article

- **Goal:** To summarise the evidence on accuracy of the Ottawa ankle rules, a decision aid for excluding fractures of the ankle and mid-foot.

- **Design:** Systematic review.
Ottawa Ankle Rules

- Developed in 1992 by Stiell
- Goal: Questionnary to assess ankle and foot

Ottawa Ankle Rules

- Management of ankle sprains is a daily routine at emergency departments;
- Only 15% patients with ankle sprain in emergency department who undergo radiography present a fracture;
- Designed to avoid unnecessary radiography by ruling out fractures of the mallellus and mid-foot.
- **Ankle**: assess the hability to walk 4 steps (immediately after the injury or in the emergency department) and note localised tenderness of the posterior edge or tip of either malleolus (four spots).

- **Mid-foot**: assess the hability to walk and note tenderness of the navicular bone or the base of the fifth metatarsal.
Ankle (pain in the malleolar zone)
- Bone tenderness at the posterior edge/tip of lateral malleolus (6 cm) A
- Bone tenderness at the posterior edge/tip of medial malleolus (6 cm) B
- Inability to weight bear, both immediately and in the emergency department.

Mid-foot (pain in the mid-foot zone)
- Bone tenderness at the base of the fifth metatarsal C
- Bone tenderness at the navicular D
- Inability to weight bear, both immediately and in the emergency department.
Methods

- They focused on studies in which the Ottawa Ankle Rules was used to diagnose fractures of the ankle or mid-foot.

- Data bases:
  - Medline
  - Premedline
  - Embase
  - CINAHL
  - Cochrane Library
  - Science Citation Index Database
Abstracts and titles found by the data bases were independently examined by JS (professor) and LMB (senior research fellow).

Paper's eligibility was disputed.

The document was obtained and examined/scrutinised.

Use of a checklist to assess that criteria for inclusion was used.

Minimal requirements for inclusion were assessed and classified: construct at least a 2×2 table specifying the false positive rate and the false negative rate.

Disagreement in eligibility were resolved by consensus.

EK e LMB analised the methods of the data collections, patient selection, blinding and prevention of verification bias and description of the instrument and reference standart.

Calculated several pooled estimates of the negative likelihood ratio by successively increasing the number of methodological criteria required.
They calculated several pooled estimates of the negative likelihood ratio by successively increasing the number of methodological criteria required.

Calculated sensitivities, specificities, likelihood ratios, and other standard errors.

Searched for pooled sensitivity (using the bootstrap) and in the pooled likelihood ratio of a negative result (using a random effects model) how many times more likely it is to find a negative result among people with a fracture (sensitivity) than among those without (specificity).

Tested heterogeneity of sensitivities and specificities using $X^2$ tests, but the interpretation was hampered by small numbers of false negative results.

Analyses of groups and sub-groups as well as study design.
After inspection of the receiver operating characteristics plot they decided to pool sensitivities, but not specificities.
Results

- 1085 results from the data base research;
- 116 were obtained +15 from other references.

From that they analised 35 studies (criteria for inclusion);

32 studies investigate the accuracy of OAR (16 ankle; 11 mid-foot; 10 both; 6 of the 32 were in children);
Excluded studies that:
-collected data non-prospectively;
-unknown blinding of the radiologist;
-analysed data which was judged from other health professionals other than doctors;
-modified the OAR.
They were left with **27 studies** (12 ankle, 8 mid-foot, 10 both, 6 ankle or mid-foot in children);

**Population**
15 581 patients (47 patients (0.3%) had a false negative result: result that appears negative but it is actually positive).
Sensability and Specificity

Sensibility:
- 99.6% (studies under 48h hours after injury);
  Confidence interval: 98.2% - 100%
- 96.4% (studies with combined assessment);
  Confidence interval: 93.8% - 98.6%.

Specificity:
- 47.9% (prevalence of fracture under 25%)
  Interquartile range: 42.3% - 77.1%);
- 26.3% (combined assessment).
  Interquartile range: (19.4% - 34.3%)
Accuracy in diagnosing ankle fractures

Probability of having a fracture after a negative result: 15%;

Probability of having a fracture after a negative result: 1.22% – Children;

Probability of having a fracture after a negative result: 2.91% - Combined Assessment.

Table 4: Pooled likelihood ratios (random effects) for negative result using Ottawa ankle rules in 27 studies (39 2×2 tables) on accuracy of the instrument in diagnosing ankle fractures. Probabilities of fracture after negative testing are calculated assuming 15% prevalence of fracture.

<table>
<thead>
<tr>
<th>Category</th>
<th>Negative likelihood ratio (95% CI)</th>
<th>P value for heterogeneity</th>
<th>Fracture probability (%) (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All (n=39)</td>
<td>0.10 (0.06 to 0.16)</td>
<td>&lt;0.001</td>
<td>1.73 (1.05 to 2.75)</td>
</tr>
<tr>
<td>Ankle assessment (n=15)*</td>
<td>0.08 (0.03 to 0.18)</td>
<td>&lt;0.001</td>
<td>1.39 (0.53 to 3.08)</td>
</tr>
<tr>
<td>Foot assessment (n=10)†</td>
<td>0.08 (0.03 to 0.20)</td>
<td>0.14</td>
<td>1.39 (0.53 to 3.41)</td>
</tr>
<tr>
<td>Combined assessment (n=14)‡</td>
<td>0.17 (0.10 to 0.30)</td>
<td>0.04</td>
<td>2.91 (1.73 to 5.03)</td>
</tr>
<tr>
<td>Children (n=7)</td>
<td>0.07 (0.03 to 0.18)</td>
<td>0.9</td>
<td>1.22 (0.53 to 3.08)</td>
</tr>
<tr>
<td>Adults (n=32)</td>
<td>0.11 (0.06 to 0.18)</td>
<td>&lt;0.001</td>
<td>1.90 (1.05 to 3.08)</td>
</tr>
</tbody>
</table>

Fracture prevalence§:
- Lower fourth (n=7): 0.04 (0.02 to 0.11) [0.97] 0.70 (0.35 to 1.90)
- Middle fourths (n=22): 0.09 (0.05 to 0.16) [0.001] 1.56 (0.87 to 2.75)
- Upper fourth (n=10): 0.22 (0.10 to 0.51) [0.007] 3.74 (1.73 to 8.26)
- Ottawa ankle rules applied ≤48 h (n=5): 0.06 (0.02 to 0.19) [0.65] 1.05 (0.35 to 3.24)
- Ottawa ankle rules applied >48 h (n=34): 0.11 (0.07 to 0.18) <0.001 1.90 (1.22 to 3.08)

*Two reports on children.
†One report on children.
‡Four reports on children.
§Median prevalence 7.9% in lower quartile, 12.7% in middle quartile, and 20.6% in upper quartile.
What could affect the accuracy of diagnosis?

<table>
<thead>
<tr>
<th>Table 5</th>
<th>Methodological criteria that could affect accuracy of diagnosis of ankle or mid-foot fracture. All studies were prospective</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Criterion</strong></td>
<td><strong>Negative likelihood ratio (95% CI)</strong></td>
</tr>
<tr>
<td>Type of entry to study:</td>
<td></td>
</tr>
<tr>
<td>Consecutive (n=16)</td>
<td>0.13 (0.06 to 0.26)</td>
</tr>
<tr>
<td>Arbitrary or unknown</td>
<td>0.09 (0.05 to 0.16)*</td>
</tr>
<tr>
<td>Gold standard applied:</td>
<td></td>
</tr>
<tr>
<td>All patients (n=17)</td>
<td>0.16 (0.09 to 0.26)</td>
</tr>
<tr>
<td>Not all patients (n=22)</td>
<td>0.08 (0.04 to 0.15)†</td>
</tr>
<tr>
<td>Blinding of radiologist:</td>
<td></td>
</tr>
<tr>
<td>Yes (n=27)</td>
<td>0.08 (0.05 to 0.15)</td>
</tr>
<tr>
<td>No or unknown (n=12)</td>
<td>0.15 (0.07 to 0.31)‡</td>
</tr>
</tbody>
</table>

*Consecutive vs arbitrary or unknown, P=0.49 (meta-regression analysis testing).
†All patients vs not all patients, P<0.001 (meta-regression analysis testing).
‡Yes vs no or unknown, P=0.29 (meta-regression analysis testing).
Discussion

- Less than 2% of patients (in most subgroups) who were negative for fracture actually had a fracture.
- Because OAR is calibrated towards high sensitivity, they were interested in the pooled sensitivity and the pooled likelihood ratio of a negative result.

It means that they were interested in how much the probability of having a fracture decrease when a test (OAR) is negative.
Specificity 10% - 79%
(represented the number of the total group of patients without the condition who have a negative test, based on a definitive standard)

Still need a radiography to confirm that they don’t have a fracture
What could influence specificity?

Difference between:
- Clinical skills;
- Test interpretation;
- Experience;
- Patient’s pain expression;
- Palpation.
Limitations

The type of practitioner and the years of experience are not specified in the review.

The inclusion conditions are also not specific.
Good results with sensibility, but the goal is to decrease the number of fractures undetected early, so the specificity needs to be higher.

The instrument has a sensitivity of almost 100% and a modest specificity, and its use should reduce the number of unnecessary radiographs by 30 to 40%.

This review proved that the OAR has a high sensibility but not enough specificity, which means that when the results is negative the odds of having an ankle fracture are low, but if the OAR tests are positive it doesn’t mean that you have an ankle fracture.
What about the clinical usefulness of the OAR?

Exclusion criteria for the use of the Ottawa Ankle Rules were (1992):

- chronic injury (more than 10 days),
- pregnancy,
- the presence of isolated injuries to the skin (e.g., lacerations, abrasions, burns),
- patients under 18 years of age.
In the case of the Ottawa Ankle Rules, with a sensitivity range of 96.4% to 99.0%, a negative test finding is a reasonable indicator that no fracture is present.

A negative likelihood ratio is between 0 and 1; ratios closer to 0 increase the odds that the condition will truly be absent with a negative test.

Negative likelihood ratios of these sizes represent a large and nearly conclusive shift in the probability that the condition is not present.
The Buffalo Rule

- Modification of OAR

- The Buffalo Rule means that the point tenderness criterion directed to the crest or midportion of the malleoli (distal 6 cm of the fibula and tibia), reducing the likelihood of palpating over injured ligament structures.

- In the 2 studies to date that assessed the diagnostic accuracy of the Buffalo Rule, sensitivity for malleolar pain was reported to be 100% (that is, all patients with malleolar pain had fractures; 95 confidence interval [CI] = 59%, 100%) and 100% (95% CI = 78%, 100%), and specificity for malleolar pain was reported to be 59% (95% CI = 47%, 71%) and 45% (95% CI = 43%, 46%).

- The Buffalo Rule is reported to result in a 54% reduction in radiography costs.

The Ottawa Ankle Rules are reported to result in a 19% to 38% reduction in radiography costs associated with excluding ankle fractures after sprain injury.
Limitations

- Only a few of the authors reported on the characteristics of the clinicians;
- Assessment reported to date has been conducted in the emergency departments hospitals;
- There is no finding directed to sports injury population.
Conclusion/Implications for practice

- With physical therapy range of practice it becomes imperative to remain current with clinical decision aids, which are developed for the purpose of enhancing patient care through informed decision making.

- Based on the current research, it is recommended that the Ottawa Ankle Rules be included in both athletic training clinical practice and educational programs.

- It’s a way to reduce cost in the health department and increased awareness of unnecessary tests and procedures.