Interpreting Studies of Fall Risk Tools
The diagnostic validity of a clinical tool is the extent to which the tool measures what it is designed to measure. With regard to fall risk tools, validity is the extent to which the tool correctly identifies patients at high risk for falls. A number of statistical measures are commonly used to assess the diagnostic validity of fall risk tools, including sensitivity and specificity. Because it is difficult, if not impossible, to create a tool that is 100% valid, diagnostic validity is measured by degrees. For most measures of tool validity, values close to 100% (or to 1, depending on the variable) reflect the optimal state of validity, but it is not realistic in the real world.

An effective fall risk assessment tool for use in the clinical setting should be sensitive enough to identify high-risk persons/patients and specific enough to identify those persons who are not at risk, thereby allowing for the targeted use of evidence-based interventions to minimize injurious falls. A tool should also be simple to use as holistic screening, easy to score, and quick to complete to minimize burden on care provider time. It should also support specific interventions aimed at reducing the underlying cause of the identified risk factors. For example, confusion is not a diagnosis, and yet etiologies of confusion are among the most underdiagnosed factors in many acute care facilities. Ideally, the presence of a risk factor like confusion should trigger an immediate evaluation of the patient; for patients with confusion, this evaluation should differentiate delirium from dementia or medication side effects, for example, and determine whether a more in-depth assessment is needed. In other words, reframing fall risk factors as a window into healthier aging and functionality, while reducing injurious falls, is the real opportunity of using a validated tool in today’s care continuum.

Aspects of Study Design Can Affect Validity of Fall Risk Tools
Considerations when interpreting studies of fall risk tools include the study design (eg, prospective vs. retrospective, case-control vs. uncontrolled, observational vs. interventional, etc.), patient population (eg, general acute care vs. specific populations), sample size (eg, large vs. small, number of fall events), duration (long vs. short), and methods of data retrieval and analysis (eg, chart review vs. prospective data collection). Larger and more diverse samples, particularly of the study event (ie, falls), may increase the reliability of findings and, possibly, foster broader applicability to real-world populations. For example, the initial validation study of the HIIFRM<sup>TM</sup> enrolled more than 1100 general acute-care patients and recorded 254 fall events over 2 years. This large study included the diverse patient population of a Level I Trauma center, with skilled nursing and rehabilitation onsite (increasing the generalizability of the findings), a large number of events (increasing the reliability of the findings), and a long duration (increasing capture of rare events).

In prospective studies, investigators determine before the study what variables will be assessed, which subjects will be enrolled, how study tools will be administered, the length and sample size of the study, and other parameters. This type of design fosters consistency of tool application and data acquisition. However, prospective studies in which providers administer the risk assessment but do not perform any intervention may suffer from the so-called Hawthorn effect. This term describes changes in care delivered by providers based on their awareness of risk. In other words, providers who identify patients who are high risk for falls will alter their approach to these patients, possibly reducing the incidence of falls. In contrast, retrospective studies collect data from existing sources (eg, Electronic Medical Records) and therefore cannot ensure consistency in how the tool is used, how the data are collected, and how subjects are enrolled, often with variability in assessments between providers.

Other aspects of study design that can affect measures of diagnostic validity include the use of fall prevention programs at the study sites. Effective fall reduction programs will skew validity analysis by preventing the outcome of interest (ie, falls), unless statistical analysis is performed to adjust for this confounder. The study population can also affect the evaluation of validity. An older or sicker population, for example, may have higher fall risk overall, possibly altering measures of tool accuracy and affecting the generalizability of the findings. Indeed, readers of the falls literature must use caution when interpreting the results of studies that incorporate falls prevention programs, specific patient populations (eg, elderly, pediatric), or settings (eg, rehabilitation, orthopedics, neurology) that may affect the overall incidence of falls.

Application of the risk assessment tools by clinicians during the studies may differ as well. For example, it is recommended that HIIFRM<sup>TM</sup> be administered repeatedly, including on admission, following changes in patient status, and after a fall event. However, many studies require assessment only on admission, with no re-evaluations during hospital stay, even following surgery or other changes in patient status. This use of the HIIFRM differs from recommended practice and may miss changes in fall risk, thereby introducing inaccuracies to study findings.
Studies of the HIIFRM: High Validity Across Populations

The HIIFRM has been evaluated in multiple studies from the United States, Italy, Portugal, Singapore, China, Korea, and other countries (Table 1). Study methodologies vary, but include case-control and prospective and retrospective designs. Validity metrics reported in these studies are relatively consistent: sensitivity in the range of 64.9% to 93.2% and specificity 51.3% to 89.3%, with positive predictive values (PPV) from 0.95% to 7.5% and negative predictive values (NPV) from 96% to 99.7%. These values reflect the percentage of fallers (PPV) and non-fallers (NPV) correctly identified by the tool.

It should be noted that lower sensitivities and specificities have been reported in some studies. For example, an Italian study reported a sensitivity of 45.8%, and two groups reported low specificities (43% and 35%), likely a result of the elderly patient populations enrolled in these studies. Two emergency department (ED) studies reported sensitivities of only 23.8% to 37.5%. In one of these studies, the HIIFRM was not administered during the study, but rather, assessed only through review of risk factors in patient charts, which would not allow for an accurate functional test. In the second study, only fall-related visits to the ED were counted as falls, and falls that occurred in the home that did not lead to an ED visit were not counted.

Summary

The HIIFRM compares very favorably to other tools in terms of sensitivity, specificity, and positive and negative predictive value in multiple clinical studies and diverse patient populations. Moreover, the HIIFRM is the only tool to incorporate independent risk factors identified through systematic evaluation and regression analysis of more than 600 variables reported in peer-reviewed literature. These qualities make the HIIFRM a truly predictive tool, and allow for the targeting of interventions to specific categories of injurious fall risk with a larger opportunity to build a care continuum approach that can reduce modifiable risk factors whenever possible to preserve health and function. Removing the underlying cause of risk can alter a person’s risk state, if the provider views risk factors as a holistic approach to care and management.

References

### Table 1. Diagnostic validation studies of the Hendrich II Fall Risk Model (HIIFRM)

<table>
<thead>
<tr>
<th>Study</th>
<th>N</th>
<th>Setting</th>
<th>Study design</th>
<th>Sensitivity, %</th>
<th>Specificity, %</th>
<th>PPV/NPV, %</th>
<th>AUC</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>General Acute Care Inpatients</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hendrich et al, 1995*</td>
<td>102 falls 236 control</td>
<td>General acute care</td>
<td>Case-control</td>
<td>77</td>
<td>72</td>
<td>NR</td>
<td>NR</td>
<td>Cross validation with a 1987 data set showed a sensitivity of 83% and specificity of 66%</td>
</tr>
<tr>
<td>Hendrich et al, 2003²</td>
<td>355 falls 780 controls</td>
<td>General acute care</td>
<td>Case-control</td>
<td>74.9</td>
<td>73.9</td>
<td>NR</td>
<td>NR</td>
<td>HIIFRM validation study</td>
</tr>
<tr>
<td>Kim et al, 2007³</td>
<td>5489 total</td>
<td>Acute care</td>
<td>Descriptive, prospective</td>
<td>70</td>
<td>61.5</td>
<td>2/99.5</td>
<td>0.73</td>
<td>Comparative study at single hospital in Singapore</td>
</tr>
<tr>
<td>Lovallo et al, 2010⁴</td>
<td>1148 total 59 falls</td>
<td>Acute care, patients &gt; 50 years</td>
<td>Prospective, observational</td>
<td>45.8</td>
<td>71.0</td>
<td>6/96</td>
<td>NR</td>
<td>Comparative study at medical, surgical, and rehabilitation units at an Italian hospital</td>
</tr>
<tr>
<td>Chapman et al, 2011⁵</td>
<td>1540 total</td>
<td>Acute care</td>
<td>Descriptive, cross-sectional</td>
<td>64.9</td>
<td>69</td>
<td>7.5/98.1</td>
<td>NR</td>
<td>Comparative study of four fall-risk assessment tools in 17 units</td>
</tr>
<tr>
<td>Nassar et al, 2013⁶</td>
<td>1815 total</td>
<td>Acute care</td>
<td>Prospective</td>
<td>55.2</td>
<td>89.3</td>
<td>16.5/98.3</td>
<td>NR</td>
<td>Comparative study in Lebanese acute care hospital</td>
</tr>
<tr>
<td>Kim et al, 2013⁷</td>
<td>1026 total 32 falls</td>
<td>Acute care (neurological patients)</td>
<td>Prospective, descriptive</td>
<td>59.4</td>
<td>78.5</td>
<td>8.2/98.4</td>
<td>0.74</td>
<td>Comparative study of neurology, neurosurgery, and rehabilitation patients at a Korean hospital</td>
</tr>
<tr>
<td>Yip et al, 2016⁸</td>
<td>10381 total 64 falls</td>
<td>10 general adult wards, patients &gt; 21 years</td>
<td>Prospective observational</td>
<td>75</td>
<td>51.31</td>
<td>.95/99.7</td>
<td>0.67</td>
<td>Comparative study at single hospital in Singapore</td>
</tr>
<tr>
<td>Jung and Park, 2018⁹</td>
<td>15,170 controls 310 falls</td>
<td>Acute care</td>
<td>Retrospective case-control</td>
<td>67.4-80.0</td>
<td>59.5-64.0</td>
<td>4/99</td>
<td>0.70-0.74</td>
<td>Evaluated on neurology, neurosurgery, hematology, and oncology units (255 beds) in tertiary care hospital in Korea</td>
</tr>
<tr>
<td>Cho et al, 2018¹⁰</td>
<td>14,307 total 238 falls</td>
<td>Acute care</td>
<td>Retrospective</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>0.69</td>
<td>Modeling study conducted at tertiary care hospital in Korea</td>
</tr>
<tr>
<td>Study</td>
<td>N</td>
<td>Setting</td>
<td>Study design</td>
<td>Sensitivity, %</td>
<td>Specificity, %</td>
<td>PPV/NPV, %</td>
<td>AUC</td>
<td>Notes</td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------</td>
<td>----------------------------------</td>
<td>---------------</td>
<td>---------------</td>
<td>---------------</td>
<td>------------</td>
<td>-------</td>
<td>------------------------------</td>
</tr>
<tr>
<td>Ivziku et al, 2010</td>
<td>179 total 14 falls</td>
<td>Geriatric acute care, patients &gt; 65 years</td>
<td>Descriptive, prospective</td>
<td>86</td>
<td>43</td>
<td>11/97</td>
<td>0.72</td>
<td>Italian validation study</td>
</tr>
<tr>
<td>Caldevilla et al, 2013</td>
<td>586 total 104 falls</td>
<td>Acute care, patients &gt; 65 years</td>
<td>Prospective</td>
<td>93.2</td>
<td>35</td>
<td>17.2/97.3</td>
<td>0.65</td>
<td>Portuguese validation study</td>
</tr>
<tr>
<td>Zhang et al, 2015</td>
<td>989 total 32 falls</td>
<td>Acute care, patients &gt; 60 years with chronic diseases</td>
<td>Prospective cross-sectional</td>
<td>72.0</td>
<td>69.0</td>
<td>7/98</td>
<td>0.82</td>
<td>Chinese validation study</td>
</tr>
</tbody>
</table>

*Previous version of HIIFRM, which did not include Get-Up-and-Go test.
AUC: area under the receiver operating characteristic curve; PPV/NPV: positive predictive value/negative predictive value; NR: not reported; ED: emergency department; PCP: primary care physician

References