Psychometric and pragmatic properties of social screening tools

De Marchis EH, Brown E, Aceves BA, Loomba V, Molina M, Cartier Y, Wing H, Gottlieb LM
Summer 2022
### Defining psychometric properties

<table>
<thead>
<tr>
<th>Property</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internal consistency</td>
<td>Items that purport to measure the same construct produce similar scores in the same tool</td>
</tr>
<tr>
<td>Construct validity</td>
<td>Tool accurately measure a concept, trait, or other theoretical entity</td>
</tr>
<tr>
<td>Criterion validity</td>
<td>Tool can detect clinically important changes in the construct it measures over time</td>
</tr>
<tr>
<td>Structural validity</td>
<td>Tool correlates with an established standard of comparison</td>
</tr>
<tr>
<td>Responsiveness</td>
<td>Items rise and fall together, otherwise known as “test structure”</td>
</tr>
<tr>
<td>Norms</td>
<td>Tool can be considered generalizable, as assessed by indicators such as sample size and means</td>
</tr>
</tbody>
</table>
Defining pragmatic properties

- **Cost per use**: Per-use cost of using tool, ranging from poor (> $100 per use) to excellent (free)
- **Length of tool**: Number of items, ranging from poor (> 200 items) to excellent (< 10 items)
- **Language/reading accessibility**: Readability ranging from poor (content-level expertise required) to excellent (< 8th grade reading level)
- **Ease of training**: Ease in tool can be implemented, ranging from poor to excellent
- **Ease of interpretation**: Ease in which results can be interpreted, ranging from poor to excellent
Do social screening tools reliably measure social risks/assets across diverse settings and populations?


Key findings:

1. Screening tools had rarely been tested using gold standard tools to assess psychometric validity
2. Screening tools had generally favorable pragmatic properties
3. No information on if/how tool properties varied by patient race, ethnicity, or language

Aim of this study

The adoption of healthcare-based social screening has been steadily increasing since the prior review. To update their work, we conducted a systematic scoping review to describe the evidence on psychometric and pragmatic properties of multi-domain healthcare-based social screening tools.

Inclusion criteria:

- Published in the peer reviewed literature from 2018-2021
- Examined 2+ types of validity and/or the reliability of multi-domain social screening tools*

*Henrikson NB et al.’s original 2019 review did not require articles evaluate more than 1 type of validity or reliability testing
Results

Five articles met inclusion criteria.

Three articles evaluated new tools:
• Duke Population Health Profile
• SINCERE
• TLS-C

Two articles compared tools that had been described in the original review:
• WE CARE and CMS's AHC tool
• YCLS, AHC, and Children’s Health Watch housing questions
No tool development process followed every step of gold standard measure development, though all used at least one. The most commonly assessed psychometric constructs were:

- Internal consistency (N=4)
- Structural validity (N=4)
- Concurrent validity (N=3)

Most studies did not assess pragmatic validity.
Results

Racial, ethnic, and language differences are under-examined in the existing validity literature.

Only one study provided information on how psychometric properties varied by race and ethnicity, and these studies found reported no differences across studied populations.

No studies explored differences in tool properties by language.
Looking ahead

Research is needed on the validity and reliability of different screening tools. This research should include studies that examine differences in validity by sociodemographic factors (e.g., race/ethnicity and language).
Download the full SCREEN report and executive summary on the SIREN website.

Questions about this section?
Contact:
Erika.brown@ucsf.edu
Emilia.demarchis@ucsf.edu