Possibilities for Change

The Science of Youth Risk Screening

The differences between home-grown or combined tools and a validated, standardized screening – and why those differences really matter.

Introduction

Organizations who work with youth are often pressed for time and resources. Yet, they want to make a meaningful difference in the youth they serve by identifying and reducing the risks that cause avoidable illness and premature death.

This combination of issues often leads organizations to develop in-house or home-grown risk screenings in a good-faith effort to help the youth they work with. These screenings may be created from scratch – or a selectively compiled group of questions from existing screening tools. These in-house screenings are most commonly developed because the organization:

- Doesn't have the money to license an external tool
- Is seeking to screen for a specific set of risks (based on a trending risk, to fulfill a grant-funded intervention, etc.)
- Is concerned about time and are looking to implement screening with minimal impact to workflow

Whatever the reasons, the outcome of these internally developed screenings can be sub-optimal – and while well-meaning, they represent a missed opportunity.

Here's the science behind why.

Building an Evidence-Informed Framework for Risk Screening

Youth risk screening with a standardized, validated screening tool is a recommendation of leading healthcare organizations due to the evidence of need.

Just as there is evidence of the need for screening and risk reduction efforts, there is significant evidence for what makes a risk screening effective.

The U.S. Centers for Disease Control and Prevention (CDC) supported the development of an evidence-informed framework for risk screening in an effort to provide guidance to the Centers for Medicare and Medicaid Services (CMS), to healthcare providers, and to other professionals wishing to implement a risk screening process.[1]

"Researchers who create novel assessment instruments need to state the development process, reliability measures, pilot results, and any other information that may lend credibility to the use of homegrown instruments."[2]

Organization	Youth Health Risk Screening Recommendations
American Medical Association (AMA)	**Annual comprehensive screening** for risky behaviors
American Academy of Pediatrics (AAP)	Annual screening psychosocial/behavioral screening & drug/alcohol use screening
US Preventive Service Task Force: AHRQ	Screening for depression, **sexual activity and tobacco use/prevention**
American Academy of Family Physicians (AAFP)	Screening for **sexual activity,** depression, **tobacco use**
American College of Preventive Medecine	Annual **comprehensive screening** for risky behaviors – all visit types

The CDC framework recommends risk screenings meet the following ten criteria:

1	Balance comprehensiveness of screening with provider and patient burden
2	Build upon high priority questions
3	Use person-centered and culturally appropriate processess
4	Comply with all federal laws and regulations regarding access for persons with disabilities
5	Use a shared decision-making process
6	Offer training to health providers
7	Offer action-oriented information
8	Use principles of quality improvement
9	Incorporate information into secure electronic health records
10	Conduct research to quantify long term outcomes

Note: many organizations who have created in-house screeners focus extensively on the questions to be included in the screening. In youth risk reduction, questions are only the beginning. For more information on development considerations and criteria specific to youth risk reduction, download our evaluation guide.

Retrieved from and original report available at:
https://www.cdc.gov/policy/hst/hra/

Organizations developing an internal risk screening must focus on much more than the questions that will be asked. This is highlighted in the CDC framework – where only the first two criteria focus on the questions themselves. In the process of developing this framework, the CDC working group identified methodological concerns and limitations that are particularly relevant to risk screenings. They note that the potential lack of understanding of health risk questions by participants impact the effectiveness of a risk screening and must be considered in development. The CDC framework recommends screenings should be tailored for literacy, culture, and/or age groups, and **should be evaluated for validity and reliability.**[1]

Validity & Reliability…The Anatomy of Risk Screening

So how exactly are validity and reliability evaluated? The following is a brief introduction to some common measures of reliability and validity when evaluating an effective, evidence-based youth risk screening.

Reliability

Reliability refers to the **consistency** of a screening. If the results or findings from the use of the assessment can be replicated consistently with the same types of subjects (it gives the **same results every time**), it is considered to be reliable. Reliability is a part of the screening of validity.[2]

"While reliability is necessary, it alone is not sufficient. For a test to be reliable, it also needs to be valid. For example, if your scale is off by 5 lbs, it reads your weight every day with an excess of 5lbs. The scale is reliable because it consistently reports the same weight every day, but it is not valid because it adds 5 lbs to your true weight. It is not a valid measure of your weight."[3]

Inter-rater reliability

Inter-rater reliability is a measure of reliability used to ascertain the degree to which different individuals agree in their interpretation of the content being assessed.[4] Inter-rater reliability is particularly important in youth risk screening, because all youth will not necessarily understand or interpret the wording of questions in the same way. With youth, this interpretation of specific wording is also influenced by the rapid changes in brain development during adolescence. This means questions designed to assess risk in youth (and the associated testing of inter-rater reliability for those questions) should also account for their age and stage of brain development.

Equivalence Reliability

Equivalence ensures that the same test administered to two different people gives similar results.[5] Equivalence reliability for an youth risk screening can be measured by the correlation of scores between that screening and another validated instrument that measures the same risk constructs.[6]

Validity

The three common forms of validity include content, construct, and criterion-related validity. In order for screening to be considered valid it should meet all three of these types of validity.[2]

Construct or "Face" Validity

Construct validity can be described as "the degree to which a procedure, especially a psychological test or screening, appears effective in terms of its stated aims."

Because construct validity focuses on how effective something "appears" rather than an objective measurement of efficacy – it is also commonly referred to as "face" validity. **And for most home-grown screening tools, face validity is the only measure of validity applied.**

Content Validity

"Content validity refers to the extent to which the items of a measure reflect the content of the concept that is being measured." For example, a measure of loneliness might assess elements such as a feeling of isolation, the presence of close relationships, having a trusted person to talk to, etc.[7]

Content validity is initially determined using expert judgement. A screening with high content validity should have a high correlation in findings with other instruments measuring the same content domain.[8] In her guide to validity of screening instruments referenced earlier, Dr. Sullivan explains that because content validity is a prerequisite for other validity, it should receive the highest priority during instrument development.[2]

Criterion-related Validity

Criterion Validity is also called predictive validity – this measures how effective a screening is at accurately predicting or determining a behavior in a specified situation.[9, 10] To measure the criterion validity of a risk assessment it should be calibrated against a known standard or an established measure, this is known as concurrent validity.[8]

So What's Next? Putting Evidence Into Action

As professionals working with youth we have long accepted the evidence behind vaccines, monitoring developmental milestones, preventive care, and… well, it's a long list. With 75% of all morbidity and mortality in youth related directly to risk factors – making the determination to use a validated, reliable, age-specific risk screening that has been evaluated and proven effective with youth is quite literally a life or death decision. We believe that's reason enough to add validated, reliable youth risk assessment to your accepted evidence-based list.

For more about the validity and reliability of the RAAPS youth risk screening, check out our peer-reviewed study results at: http://www.possibilitiesforchange.com/publications/

REFERENCES

- Goetzel, RZ; Staley, P; Ogden, L; Stange, P; Fox, J; Spangler, J; Tabrizi, M; Beckowski, M; Kowlessar, N; Glasgow, RE, Taylor, MV. A framework for patient-centered health risk assessments – providing health promotion and disease prevention services to Medicare beneficiaries. Atlanta, GA: US Department of Health and Human Services, Centers for Disease Control and Prevention, 2011. Available at: http://www.cdc.gov/policy/opth/hra/.
- Sullivan, G.M.. A Primer on the Validity of Assessment Instruments. (2011). Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3184912/
- McLeod, S. A. (2007). What is reliability? Retrieved from www.simplypsychology.org/reliability.html
- Phelan, C.; Wren, J. Exploring Reliability in Academic Assessment. (2006). Retrieved from: https://chfasoa.uni.edu/reliabilityandvalidity.htm
- Shuttleworth, M. Instrument Reliability. (2009). Retrieved from: https://explorable.com/instrument-rliability
- Center for Postsecondary Research. Indiana University School of Education. National Survey of Student Engagement: Reliability. (2018). Retrieved from: http://nsse.indiana.edu/html/reliability.cfm
- Bell, B. Content Validity. Retrieved May 1, 2018 from: http://www.psychologyandsociety.com/contentvalidity.html
- Packer, M. Test Construction: Introduction and Overview. (2004). Retrieved from: http://www.mathcs.duq.edu/~packer/Courses/Psy624/test.html
- Shuttleworth, M. Criterion Validity. (2009). Retrieved from: https://explorable.com/criterion-validity
- Criterion Validity. Retrieved May 1, 2018 from: https://www.alleydog.com/glossary/definition.php?term=Criterion+Validity