Validity and reliability - What's it all about?
Part 1 Validity in quantitative studies

This is one of a series of short papers on aspects of research by Alison Twycross and Linda Shields.

Just because a research study has been published in a journal does not mean that it is good research or that the results are applicable to your area of clinical practice. When reading a paper, it is necessary to consider the validity and reliability of the study being described. This and the next article in this series will define validity and reliability, and will explain how you can ascertain whether a quantitative research study is both valid and reliable.

Validity means that a tool measures what it sets out to measure – for example, that a pain assessment tool measures pain intensity rather than anxiety. There are several measures of validity that provide evidence of the quality of a study.

Internal and external validity relate to the overall study design. Internal validity relates to the extent to which the design of a research study is a good test of the hypothesis or is appropriate for the research question (Carter and Porter 2000). External validity, meanwhile, relates to whether or not research findings can be generalised beyond the immediate study sample and setting.

Box 1: Measures to assess the validity of data collection tools

<table>
<thead>
<tr>
<th>Content validity</th>
<th>Criterion validity</th>
<th>Construct validity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whether a tool appears to others to be measuring what it says it does. Face validity is a simple form of content validity - the researcher asks a few people to check the tool covers all areas. A more rigorous way to assess content validity is to ask recognised experts in the area to give their opinion on the validity of the tool.</td>
<td>Concurrent or predictive validity are both measures of criterion validity. Concurrent validity uses an existing and well-accepted measure against which the new measure can be compared – for example, if you were developing a new pain assessment tool you would compare the ratings obtained from the new tools with those obtained using a previously validated tool. Predictive validity measures the extent to which a tool can predict a future event of interest – for example, does a tool developed to measure the risk of pressure sores in children in hospital in fact identify the children at risk? Criterion validity is usually measured using a correlation coefficient – when the correlation is high, the tool can be considered valid.</td>
<td>This tests the link between a measure and the underlying theory. If a test has construct validity, you would expect to see a reasonable correlation with tests measuring related areas. Evidence of construct validity can be provided by comparing the results obtained with the results obtained using other tests, other (related) characteristics of the individual or factors in the individual's environment which would be expected to affect test performance. Construct validity is usually measured using a correlation coefficient – when the correlation is high, the tool can be considered valid.</td>
</tr>
</tbody>
</table>

REFERENCES