The Action Research Arm Test (ARA or ARAT) is an observational test used to determine upper limb function. It was first described in 1981 as a modification of an earlier test, the Upper Extremity Function Test (UEFT) (Carroll, 1965) and was designed to assess recovery in the upper limb following cortical damage.

**Test procedure:** The test takes approximately 10 minutes to administer (de Weerdt and Harrison 1985) and while no special training is necessary it does require considerable non-standard equipment (various sized blocks of wood, cricket ball, stone, jug and glass, tube, washer and bolt, ball bearing, marble). The test consists of 19 items grouped in subtests (grasp, grip, pinch, and gross arm movement) and performance of each item is rated on a 4-point scale ranging from 0 (no movement possible) to 3 (movement performed normally). If subjects scores the maximum on the first, most difficult item of each subtest, they are credited with having scored 3 on all items of the subtest without having to be tested. If the patient scores less than 3, then the second item is tested. This is the easiest item, and if patients score 0 then they are unlikely to achieve a score above 0 for the remainder of the items and are credited with a zero for the other items and the assessor moves onto the next subtest. For example, in the Grasp subtest the first item is lifting a 10 cm³ block onto a shelf and the second item is lifting a 2.5 cm³ block. If the patient scores less than 3 for the first item and more than 0 for the second item then all items in the subtest should be assessed. The maximum obtainable score is 57.

**Reliability and validity:** Inter-rater and retest reliability have been shown to be high (ICC > 0.98) in studies involving patients with stroke (Van der Lee et al 2001). A small systematic difference was noted between two raters in one study (Van der Lee et al 2001) with a mean difference of 0.75 points and 95% CI 0.02 to 1.48. This same study also proposed a somewhat arbitrary value of 10% of the total range of the scale (i.e. 5.7 points) as the minimum clinically important difference, and then confirmed that a difference of this magnitude could be distinguished from measurement error. Concurrent validity has been confirmed by comparison with the upper limb component of the Fugl-Meyer Assessment and the Motor Assessment Scale (MAS) (Van der Lee et al 2001).

**Commentary**

It is equally important for clinicians and researchers to choose outcome measures that are valid, reliable and responsive to change. The ARAT is more frequently being used in both scenarios due to its ability to detect clinically-relevant changes in upper arm ability in the acute phase following stroke and in trials involving patients with a chronic condition. An advantage of the ARAT in the acute phase when upper limb function is limited is the ability to discontinue testing after failure of the least demanding items without sacrificing a valid score. During development, Lyle (1981) used Guttman scale analysis to ensure that items were truly hierarchical. This shortens by over 50% the time taken to complete the test. This is an advantage over an alternative outcome measure, the MAS (Hand Movements and Advanced Hand Activities Scales), where each item must be tested as the ordering of items is not truly hierarchical (Sabari et al 2005). Although the scoring of the ARAT appears complex, experience with the test confirms the comment by Lyle (p. 491) ‘This sounds complicated to explain, but is easy in practice’.

Another perceived limitation of the ARAT is that the scoring is subjective with respect to a score of 2 (‘can complete the test but takes abnormally long or has great difficulty’) or 3 (‘movement performed normally’). The original paper provided no operational definitions to elaborate on this, but subsequent studies have set time limits for each item as twice the standard deviation of the performance times of a sample of healthy adults (Van der Lee et al 2001) and comprehensive instructions have been published to ensure a standardised approach to performing the test (Yozbatiran et al 2008). An advantage of the ARAT is the possibility of videotaping assessment for scoring at a later date, or by another tester, without affecting reliability.

In summary, the ARAT is a responsive and valid measure of upper limb functional limitation and is a useful measure for use in upper limb rehabilitation and clinical research. A standardized approach to testing should be used to reduce variance between therapists and when conducting multisite research trials.

Michelle McDonnell
Research Centre for Human Movement Control, The University of Adelaide

**References**