This is a repository copy of *Utility of Magnetic Resonance Imaging in Diagnosis and Monitoring Enthesitis in Patients with Spondyloarthritis: An OMERACT Systematic Literature Review*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/141081/

Version: Accepted Version

Table 2: Psychometric properties and quality scores of selected studies (n = 38)

Author/Year of publication	Face validity*	Content validity*	Construct validity*	Criterion validity*	Reliab ility*	Responsiv eness*	QUALITY SCORING
Cross sectional studies							
McGonagle 2002	YES	YES	NO	NO	YES	NO	9
Olivieri 2002	YES	NO	NO	NO	NO	NO	7
Tan 2006	YES	NO	NO	NO	YES	NO	7
McQueen 2007	YES	YES	NO	NO	YES	NO	8
Tan 2007	NO	NO	NO	YES	NO	NO	5
Marzo-Ortega 2009	NO	NO	NO	NO	YES	NO	7
Maksymowicz 2010	NO	NO	NO	NO	NO	NO	6
Althoff 2013	YES	YES	NO	NO	YES	NO	7
Aydin 2013	YES	YES	YES	NO	YES	NO	8
Braum 2013	YES	YES	YES	NO	NO	NO	7
Ramirez 2014	YES	NO	NO	NO	NO	NO	8
Paramarta 2014	YES	YES	NO	NO	YES	NO	7
Poggenborg 2015	YES	YES	YES	NO	YES	NO	10
Tan 2015	YES	YES	NO	NO	YES	NO	7
Giraudo 2016	YES	NO	NO	NO	YES	NO	7
Agten 2016	NO	NO	YES	NO	YES	NO	8

Maldonado 2017	YES	YES	YES	NO	YES	NO	7
Aivazoglou 2017	YES	NO	YES	NO	NO	NO	7
Case control studies							
Olivieri 1998	YES	YES	YES	NO	NO	NO	6
Lambert 2004	YES	YES	NO	NO	YES	NO	6
Erdem 2005	YES	YES	NO	NO	YES	NO	5
Wiell 2007	YES	NO	NO	NO	NO	NO	8
Emad 2010	YES	NO	NO	NO	YES	NO	6
Weckbach 2011	YES	NO	YES	NO	YES	NO	7
Feydy 2012	YES	NO	NO	NO	YES	NO	8
Wiell 2013	YES	YES	YES	NO	YES	NO	9
Chen 2018	YES	YES	YES	NO	NO	NO	6
Cohort studies							
Godfrin 2004	YES	NO	YES	NO	YES	NO	6
Eshed 2008	YES	YES	NO	NO	YES	NO	6
Zhen-Guo 2013	YES	NO	YES	NO	YES	NO	9
Karpitschka 2013	YES	YES	YES	NO	YES	YES	9
Althoff 2016	YES	YES	NO	NO	YES	NO	8
Marzo-Ortega 2001	YES	YES	YES	NO	YES	YES	8

Tan 2004	YES	NO	NO	NO	NO	YES	8
de Hooge 2013	YES	NO	YES	NO	NO	YES	9
Randomised controlled trials							
Dougados 2010	NO	NO	NO	NO	NO	NO	7
Song 2011	YES	NO	NO	NO	YES	YES	9
Krabbe 2018	YES	YES	NO	NO	YES	YES	11

*Face validity was defined as expert opinion on the credibility of scoring system used in each article to measure enthesitis. Content validity estimated the reliability of the scoring system used in each study to measure the full spectrum of outcome - inflammatory and structural changes. Construct validity was achieved when MRI evaluation of enthesitis correlated with the following concepts of enthesitis: 1) clinical assessment of enthesitis using a validated enthesitis score (e.g., MASES), 2) ultrasound or radiographic assessment of enthesitis sites, and/or 3) comparison of different sequences of MRI in assessing enthesitis. Criterion validity was achieved when MRI evaluation of enthesitis correlated with a gold standard (e.g., histology). Reliability was defined in studies mentioning inter-rater reliability measures of scoring consistency between and within MRI readers, e.g. inter/intra-class correlation coefficients (ICCs) or kappa statistics. Responsiveness was achieved in studies documenting statistically significant changes in relation to treatment introduction or change.