

Supplementary Table 1. Domains of the AQUA tool, signaling questions, and risk of bias judgments

Domain & Questions	Options	Risk of bias judgment
Domain 1: OBJECTIVE(S) AND SUBJECT CHARACTERISTICS		
Was (Were the objective(s) of the study clearly defined?	yes/no / unclear	low/high/unclear
Was (Were the chosen subject sample(s) and sample size appropriate for the objective(s)?		
Are the baseline and demographic characteristics of the subjects clearly defined?		
Could the method of subject selection have introduced bias?		
Domain 2: STUDY DESIGN		
Does the study design address the research question(s)?	yes/no / unclear	low/high/unclear
Were the materials used appropriate for the objective(s)?		
Were the methods used appropriate for the objective(s)?		
Could the study design have introduced bias?		
Domain 3: METHODOLOGY CHARACTERISATION		
Are the methods described in enough detail for reproducibility?	yes/no / unclear	low/high/unclear

Domain & Questions	Options	Risk of bias judgment
Was the experience of the individuals performing the study stated?		
Were materials and methods clearly described (manufacturers, suppliers, etc.)?		
Were measures taken to reduce inter- and intra-observer variability?		
Do the images accurately reflect the methods/techniques applied?		
Could the characterization of methods have introduced bias?		
Domain 4: DESCRIPTIVE ANATOMY	yes/no / unclear	low/high/unclear
Were anatomical definitions clearly and accurately described?		
Were the outcomes and parameters assessed appropriate and clearly defined?		
Were the images/illustrations clear and understandable?		
Were the ambiguous anatomical observations clearly described?		
Could the description of anatomy have introduced bias?		
Domain 5: REPORTING OF RESULTS	yes/no / unclear	low/high/unclear

Domain & Questions	Options	Risk of bias judgment
Was the statistical analysis appropriate?		
Are the results clear, comprehensible, and consistent?		
Do reported numbers correspond to the number of subjects? If not, explain?		
Are all persons/groups reported and evaluated?		
Could the reporting of results have introduced bias?		