Page 1

Appendix

TABLE E-1 Search Strategy

Embase (1,738 Papers Retrieved)	PubMed Central (1,928 Papers Retrieved)
1. reliability	([Reliability] OR [Agreement] OR [Interrater] OR [Intrarater] OR [Inter-rater] OR [Intra-rater] OR [Kappa coefficient] OR [Weighted kappa coefficient] OR [Intraclass correlation coefficient]) AND ([Radiographic] OR [Xray] OR [x-ray]) AND ([hip] OR [pelvis] OR [femoral] OR [acetabular] OR [acetabulum] OR [femur])
2. interrater	
3. intrarater	
4. 'intra rater'	
5. 'inter rater'	
6. radiograph	
7. radiographic	
8. xray	
9. X-ray	
10. kappa AND coefficient	
11. intraclass AND correlation AND coefficient	
12. agreement	
13. hip	
14. pelvis	
15. femoral	
16. acetabular	
17. acetabulum	
18. femur	
19. #1 OR #2 OR #3 OR #4 OR #5 OR #10 OR #11 OR #12	
20. #6 OR #7 OR #8 OR #9	
21. #13 OR #14 OR #15 OR #16 OR #17 OR #18	
22. #19 AND #20 AND #21	
23. #19 AND #20 AND #21 AND [english]/lim	

TABLE E-2 Metrological Reliability Data Extracted from Included Investigations*†

	Quantitative Test Used	Interrater Reliability (95% CI)	Intrarater Reliability (95% CI)
Lateral center-edge angle (Ogata)			
Cadet et al.[23] (2016)	ICC	ICC: 0.64	ICC: 0.69
Ratzlaff et al.[25] (2016)	ICC, Cohen kappa coefficient (κ): >40° (yes/no)	κ: 0.714 (0.50-0.92)	ICC: 0.87
Bouttier et al.[30] (2013)	ICC, Cohen kappa coefficient (κ): lateral dysplasia (≤20°) (yes/no)	ICC: 0.82 (0.71-0.89), κ: 0.73 (0.48-0.98)	ICC: 0.77-0.94
Engesæter et al.[63] (2012)	ICC	ICC: 0.93 (digital), 0.88 (manual)	ICC: 0.93-0.97 (digital), 0.85-0.96 (manual)
Carlisle et al.[34] (2011)	ICC	ICC: 0.64 (0.52-0.75)	ICC: 0.88 (0.85-0.91)
Hansen et al.[38] (2012)	ICC	ICC: 0.80	ICC: 0.90-0.91
Chadayammuri et al.[56] (2015)	ICC	ICC: 0.934 (0.850-0.971)	Not reported
Hanson et al.[57] (2015)	ICC	ICC: 0.92	Not reported
Anderson et al.[33] (2011)‡	ICC	ICC: 0.80 (0.68-0.88)	ICC: 0.85-0.96
Lateral center-edge angle (unspecified)			
Colvin et al.[58] (2011)	ICC	ICC: Not reported	ICC: 0.919-0.964
Schottel et al.[36] (2014)	ICC	ICC: 0.92 (0.90-0.94)	ICC: 0.75 (0.49-1.00)
Alpha angle (frog-leg lateral)			
Cadet et al.[23] (2016)	ICC	ICC: 0.75	ICC: 0.34
Carlisle et al.[34] (2011)	ICC	ICC: 0.21 (0.11-0.36)	ICC: 0.76 (0.70-0.81)
Khan et al.[47] (2015)	ICC	ICC: 0.35 (0.25-0.49)	ICC: 0.74-0.99
Wensaas et al.[55] (2012)#	ICC	ICC: 0.38-0.78	ICC: 0.86
Espié et al.[61] (2014)	ICC	ICC: 0.955 (0.943-0.965)	ICC: 0.987 (0.979-0.991)
Espié et al.[62] (2016)	ICC	ICC: 0.782 (0.607-0.868)	ICC: 0.920 (0.900-0.936)
Konan et al.[67] (2010)	ICC	ICC: 0.83 (0.69-0.89)	ICC: 0.88 (0.71-0.90)
Schottel et al.[36] (2014)	ICC	ICC: 0.90 (0.86-0.92)	ICC: 0.80 (0.56-1.00)
Clohisy et al.[65] (2007)	ICC	ICC: 0.83	ICC: 0.73
Alpha angle (Dunn)			
Barton et al.[59] (2011)	ICC	ICC: 0.90 (0.832-0.942)	ICC: 0.979 (0.952-0.991)
Alpha angle (modified Dunn)			
Ratzlaff et al.[25] (2016)	Cohen kappa coefficient (κ): >55° (yes/no), ICC	κ: 0.714 (0.50-0.92)	ICC: 0.97
Nepple et al.[31] (2014)	ICC, Cohen kappa coefficient (κ): >50° (yes/no), % agreement: >50° (yes/no)	ICC: 0.43 (0.35-0.56), κ: 0.37 (0.29-0.46), % agreement: 69.3	ICC: 0.92 (0.87-0.95), κ: 0.85 (0.73-0.97), % agreement: 92.5
Alpha angle (false-profile)			
Hellman et al.[40] (2015)	ICC	ICC: 0.81	Not reported
Alpha angle (cross-table lateral)			
Carlisle et al.[34] (2011)	ICC	ICC: −0.08 (−0.12 to −0.01)	ICC: 0.76 (0.70-0.81)
Pollard et al.[42] (2010)	ICC	ICC: 0.84	ICC: 0.91
Mast et al.[54] (2011)	ICC	ICC: 0.83 (0.70-0.91)	ICC: 0.96-0.98

CARREIRA ET AL.
THE RELIABILITY OF COMMONLY USED RADIOGRAPHIC PARAMETERS IN THE EVALUATION OF THE PRE-ARTHRITIC HIP. A SYSTEMATIC REVIEW

Barton et al.[59] (2011)	ICC	ICC: 0.85 (0.751-0.911)	ICC: 0.947 (0.88-0.977)
Clohisy et al.[65] (2007)	ICC	ICC: 0.56	ICC: 0.50
Alpha angle (anteroposterior)			
Wensaas et al.[55] (2012)#	ICC	ICC: 0.64-0.66	ICC 0.71
Barton et al.[59] (2011)	ICC	ICC: 0.947 (0.909-0.969)	ICC: 0.884 (0.746-0.949)
Clohisy et al.[65] (2007)	ICC	ICC: 0.85	ICC: 0.60
Alpha angle (frog-leg lateral and modified Dunn)			
Hooper et al.[49] (2016)	ICC	ICC: 0.56 (0.43-0.67)	Not reported
Alpha angle (highest alpha)			
Levy et al.[64] (2015)	ICC	ICC: 0.84	ICC: Not reported
Farkas et al.[66] (2015)	ICC	ICC: 0.96 (0.89-0.98)	ICC: 0.92 (0.80-0.97)
Neck-shaft angle			
Bouttier et al.[30] (2013)	ICC, Cohen kappa coefficient (κ): metaphyseal dysplasia (>140°) (yes/no)	ICC: 0.83 (0.72-0.90), κ: 1.00 (1.00-1.00)	ICC: 0.90
Nepple et al.[31] (2014)	ICC	ICC: 0.80 (0.75-0.84)	ICC: 0.88 (0.84-0.91)
Wilson et al.[43] (2011)	ICC	ICC: 0.62 (0.42-0.78)	ICC: 0.98 (0.95-0.99)
Park et al.[50] (2016)	ICC	ICC: 0.693 (0.593-0.772)	Not reported
Mast et al.[54] (2011)	ICC	ICC: 0.58 (0.31-0.76)	ICC: 0.94-0.95
Lequesne et al.[60] (2004)	ICC	ICC: 0.89	ICC: 0.82-0.94
Anterior center-edge angle			
Bouttier et al.[30] (2013)	ICC, Cohen kappa coefficient (κ): anterior dysplasia (≤20°) (yes/no)	ICC: 0.83 (0.72-0.90), κ: 0.41 (0.07-0.74)	ICC: 0.93
Zingg et al.[32] (2009)	ICC	ICC: 0.97	NR
Anderson et al.[33] (2011)‡	ICC	ICC: 0.73 (0.59-0.84)	ICC: 0.90-0.97
Carlisle et al.[34] (2011)	ICC	ICC: 0.38 (0.26-0.53)	ICC: 0.88 (0.85-0.91)
Tannast et al.[37] (2008)	ICC	ICC: 0.63 (0.50-0.71)	ICC: 0.54-0.69
Hanson et al.[57] (2015)	ICC	ICC: 0.89	ICC: Not reported
Ischial spine sign			
Nepple et al.[31] (2014)	Cohen kappa coefficient (κ): (present/absent), % agreement: (yes/no)	κ: 0.75 (0.71-0.80), % agreement: 87.9	κ: 0.74 (0.64-0.85), % agreement: 87.5
Kappe et al.[46] (2011)	Cohen kappa coefficient (κ): (present/absent), % agreement: (present/absent)	Fleiss κ: 0.543, % agreement: 77.5	Cohen κ: 0.375-0.899, % agreement: 80-95
Malviya et al.[48] (2016)	Free-marginal multirater kappa coefficient (κ):(present/absent)	κ: 0.50	Not reported
Mast et al.[54] (2011)	Adjusted kappa coefficient (κ) for case distribution and bias (present/absent), % agreement (present/absent)	κ: 1.00, % agreement: 100	κ: 1.00, % agreement: 100
Lateral center-edge angle (Wiberg)			

Philippon et al.[26] (2010)	ICC	ICC: 0.92 (0.87-0.95)	Not reported
Nepple et al.[31] (2014)	ICC, Cohen kappa coefficient (κ): <20° (yes/no), % agreement <20° (yes/no)	ICC: 0.88 (0.85-0.91), κ: 0.82 (0.78-0.87), % agreement: 93.2	ICC: 0.95 (0.93-0.96), κ: 0.83 (0.73-0.93), % agreement: 93.8
Goldman et al.[45] (2017)§	ICC, Cohen kappa coefficient (κ): <25° (yes/no)	ICC: 0.883, κ: 0.737	Not reported
Lee et al.[44] (2011)	ICC	ICC: 0.828 (0.724-0.801)	ICC: 0.844 (0.717-0.917)
Hooper et al.[49] (2016)	ICC	ICC: 0.73 (0.52-0.84)	Not reported
Stubbs et al.[51] (2011)	ICC	ICC: 0.669	Not reported
Domb et al.[53] (2013)	ICC	Not reported	ICC: 0.79
Mast et al.[54] (2011)	ICC	ICC: 0.73 (0.53-0.85)	ICC: 0.86-0.97
Hanson et al.[57] (2015)	ICC	ICC: 0.92	Not reported
Lequesne et al.[60] (2004)	ICC	ICC: 0.90	ICC: 0.93-0.98
Kutty et al.[39] (2012)	ICC	ICC: 0.960-0.979	ICC: 0.974-0.992
Engesæter et al.[63] (2012)	ICC	ICC: 0.91 (digital), 0.84 (manual)	ICC: 0.94-0.96 (digital), 0.84-0.97 (manual)
Tannast et al.[37] (2008)	ICC	ICC: 0.92 (0.89-0.94)	ICC: 0.97-0.98
Farkas et al.[66] (2015)	ICC	ICC: 0.96 (0.90-0.98)	ICC: 0.90 (0.76-0.96)
Crossover sign			
Cadet et al.[23] (2016)	Cohen kappa coefficient (κ): (present/absent)	κ: 0.20	κ: 0.41
Jamali et al.[24] (2007)	Cohen kappa coefficient (κ): (present/absent)	κ: 0.628	κ: 0.674-0.698
Ratzlaff et al.[25] (2016)	Cohen kappa coefficient (κ): (present/absent)	κ: 0.336 (−0.02-0.69)	κ: 0.58
Nepple et al.[31] (2014)	Cohen kappa coefficient (κ): (present/absent), % agreement: (present/absent)	κ: 0.60 (0.55-0.66), % agreement: 80.8	κ: 0.72 (0.62-0.83), % agreement: 86.3
Clohisy et al.[35] (2009)	Cohen kappa coefficient (κ): (present/absent)	κ: 0.39	κ: 0.46 (0.37-0.55)
Tannast et al.[37] (2008)	Cohen kappa coefficient (κ): (present/absent)	κ: 0.60 (0.50-0.68)	κ: 0.73-0.77
Schottel et al.[36] (2014)	Cohen kappa coefficient (κ): (present/absent)	κ: 0.40 (0.29-0.51)	κ: 0.54 (0.15-0.92)
Laborie et al.[41] (2011)	Cohen kappa coefficient (κ): (present/absent)	κ: 0.82	κ: 0.59-0.80
Kappe et al.[46] (2011)	Fleiss/Cohen kappa coefficient (κ): (present/absent), % agreement: (present/absent)	Fleiss κ: 0.514, % agreement: 77.1	Cohen κ: 0.325-1.00, % agreement: 65-100
Malviya et al.[48] (2016)	Free-marginal multirater kappa coefficient (κ): (present/absent)	κ: 0.60	Not reported
Mast et al.[54] (2011)	Adjusted kappa coefficient (κ) for case distribution and bias: (present/absent), % agreement >50°: (present/absent)	κ: 0.97, % agreement: 98	κ: 0.95-1.00, % agreement: 97-100
Tönnis angle			
Nepple et al.[31] (2014)	ICC, Cohen kappa coefficient (κ): >15° (yes/no), % agreement: >15° (yes/no)	ICC: 0.80 (0.75-0.84), κ: 0.76 (0.70-0.82), % agreement: 93.9	ICC: 0.94 (0.91-0.95), κ: 0.94 (0.86-1.00), % agreement: 98.8
Bouttier et al.[30] (2013)	ICC, Cohen kappa coefficient (κ): roof acclivity indicating dysplasia (≥12°) (yes/no)	ICC: 0.84 (0.73-0.90), κ: 0.71 (0.47-0.93)	ICC: 0.78-0.92, κ: not reported
Carlisle et al.[34] (2011)	ICC	ICC: 0.42 (0.29-0.56)	ICC: 0.83 (0.79-0.87)
Hansen et al.[38] (2012)	ICC	ICC: 0.87	ICC: 0.52-0.97

Lee et al.[44] (2011)	ICC	ICC: 0.697 (0.543-0.818)	ICC: 0.502 (0.218-0.710)
Hooper et al.[49] (2016)	ICC	ICC: 0.63 (0.55-0.70)	Not reported
Goldman et al.[45] (2017)§	ICC, Cohen kappa coefficient (κ): <10° (yes/no)	ICC: 0.916, κ: 1.0	Not reported
Stubbs et al.[51] (2011)	ICC	ICC: 0.752	Not reported
Bittersohl et al.[52] (2013)	ICC	ICC: 0.830	ICC: 0.809
Domb et al.[53] (2013)	ICC	Not reported	ICC: 0.91
Mast et al.[54] (2011)	ICC	ICC: 0.45 (0.15-0.67)	ICC: 0.88-0.95
Lequesne et al.[60] (2004)	ICC	ICC: 0.90	ICC: 0.90-0.97
Engesæter et al.[63] (2012)	ICC	ICC: 0.86 (digital), 0.83 (manual)	ICC: 0.84-0.90 (digital), 0.80-0.95 (manual)
Tannast et al.[37] (2008)	ICC	ICC: 0.61 (0.52-0.70)	ICC: 0.74-0.89
Tönnis grade			
Valera et al.[29] (2016)	Cohen kappa coefficient (κ): (0, 1, 2, 3), % agreement: (0, 1, 2, 3)	κ: 0.173-0.397, % agreement: 50.4-65.0	κ: 0.364-0.397, % agreement: 60.6-66.6
Nepple et al.[31] (2014)	Weighted Cohen kappa coefficient (κ): (0, 1, 2 or 3), % agreement : (0, 1, 2 or 3)	κ: 0.22 (0.16-0.28), % agreement: 60.9	κ: 0.53 (0.40-0.66), % agreement: 75.6
Carlisle et al.[34] (2011)	Weighted Cohen kappa coefficient (κ): (0, 1, 2, 3)	κ: 0.17	κ: 0.57 (0.47-0.66)
Clohisy et al.[35] (2009)	Weighted Cohen kappa coefficient (κ): (0, 1, 2, 3)	κ: 0.59	κ: 0.60 (0.54-0.66)
Schottel et al.[36] (2014)	Weighted Cohen kappa coefficient (κ): (0, 1, 2, 3)	κ: 0.22 (0.12-0.31)	κ: 0.48 (0.17-0.79)
Malviya et al.[48] (2016)	Free-marginal multirater kappa coefficient (κ): (yes/no)	κ: 0.40	Not reported
Mast et al.[54] (2011)	Adjusted kappa coefficient (κ) for case distribution and bias: (0, 1, 2, 3), % agreement: (0, 1, 2, 3)	κ: 0.97, % agreement: 98	κ: 0.95-1.00, % agreement: 97-100
Posterior wall sign			
Nepple et al.[31] (2014)	Cohen kappa coefficient (κ): (present/absent), % agreement: (present/absent)	κ: 0.73 (0.68-0.77), % agreement: 87.5	κ: 0.84 (0.75-0.93), % agreement: 92.5
Laborie et al.[41] (2011)	Cohen kappa coefficient (κ): (present/absent)	κ: 0.83	κ: 0.55-0.73
Kappe et al.[46] (2011)	Fleiss/Cohen kappa coefficient (κ): (present/absent), % agreement: (present/absent)	Fleiss κ: 0.633, % agreement: 82.9	Cohen κ: 0.483-0.946, % agreement: 80-98
Malviya et al.[48] (2016)	Free-marginal multirater kappa coefficient (κ): (present/absent)	κ: 0.23	NR
Tannast et al.[37] (2008)	Cohen kappa coefficient (κ): (present/absent)	κ: 0.62 (0.50-0.73)	κ: 0.68-0.70

*When a range of ICCs is provided for intrarater reliability, a combined calculation was not available in the given study, so a range of values for all raters was reported. 95% CI = 95% confidence interval. †If ICCs were tabulated at multiple time points for interrater reliability, data from the first trial were reported[38,39]. ‡ICCs provided for the "templated" technique rather than the "estimated" technique. §Reliability was assessed before and after a consensus-building session; the data following the consensus-building session are reported here. #Interobserver reliability tested between 2 pairs of 3 observers rather than all 3 pairs; range provided for the 2 pairs.

Page 6

TABLE E-3 Review Inclusion and Exclusion Criteria

Inclusion Criteria	Exclusion Criteria
1. Peer-reviewed analysis	1. Study population consisted of pediatric patients with cerebral palsy or patients with Legg-Calvé-Perthes disease or advanced osteoarthrosis (Tönnis grade >2)
2. Full-text in English language	2. Mean age of study population <16 years
3. Radiographs made no earlier than 2001	3. Reliability of making a diagnosis (i.e., cam impingement, pincer impingement, dysplasia) was analyzed without defining the reliability of a specific radiographic measurement/sign
4. Interrater and/or intrarater reliability of at least 1 of the predetermined parameters was assessed on radiographs	4. Study tested the reliability of a number of different radiographic measurements/signs but provided a range of the values calculated without assigning a quantity to each individual measurement
5. Reliability of continuous parameters was assessed using an intraclass correlation coefficient (ICC)	5. The intermodality correlation of a measurement was evaluated without reporting on the reproducibility of the measurement within a single imaging modality (i.e., computed tomography vs. radiography)