

- 1 **Diagnostic accuracy of ultrasound and magnetic resonance imaging in**
- 2 **detecting Stener lesions of the thumb: Systematic review and meta-analysis**

ABSTRACT

3
4
5
6
7
8
9
10
11
12
13
14
15
16

This study aimed to assess diagnostic test accuracy of ultrasound and magnetic resonance imaging (MRI) in diagnosing thumb Stener lesions. MEDLINE, PubMed, Embase and Cochrane CENTRAL were searched for studies using ultrasound or MRI to detect Stener lesions following suspected thumb ulnar collateral ligament injuries. The reference standard was surgical exploration or clinical joint stability. Risk of bias was assessed using the QUADAS-2 tool. A random-effects bivariate meta-analysis was used to estimate pooled sensitivity and specificity. Forest plots were generated. Nine ultrasound ($n = 315$) and six MRI ($n = 107$) studies were included in meta-analysis (all high risk of bias). Pooled sensitivity and specificity for ultrasound was 95% (95% confidence interval 76-99) and 94% (88-97), and for MRI was 93% (69-99) and 98% (70-100). Both ultrasound and MRI demonstrate high diagnostic accuracy in detecting Stener lesions. Ultrasound is an appropriate first-line imaging modality.

Level of Evidence: II

17

INTRODUCTION

18 A Stener lesion is a full-thickness tear of the UCL, where the ligament is displaced proximal
19 and superficial to the adductor pollicis aponeurosis (Stener, 1962). Of full-thickness tears of
20 the distal UCL, 64-87% have been reported to result in a Stener lesion (Mahajan and
21 Rhemrev, 2013). Surgical intervention for Stener lesions is advocated as the interposed
22 aponeurosis prevents healing of the dislocated ligament. As such, early and correct diagnosis
23 of Stener lesions is therefore vital to ensure appropriate management and prevent long-term
24 sequelae of chronic pain, instability, or osteoarthritis (Christensen et al., 2016).

25

26 Clinical assessment of thumb ligamentous injuries in the acute phase can be difficult due to
27 excessive pain, swelling and muscle spasm. Imaging modalities such as ultrasound (US) and
28 magnetic resonance imaging (MRI) are frequently used to help identify a Stener lesion and
29 select patients for surgery (Figure 1). Whilst previous studies have investigated the role of US
30 and MRI (Beutel et al., 2019; Papandrea and Fowler, 2008), there is no clear consensus on
31 which imaging modality to utilise in clinical practice. The aim of this study was to perform a
32 systematic review and meta-analysis to assess the diagnostic test accuracy of US and MRI in
33 diagnosing Stener lesions of the thumb UCL.

34

35

METHODS

36 This systematic review and diagnostic test accuracy meta-analysis conformed to the Preferred
37 Reporting Items for Systematic reviews and Meta-Analysis (PRISMA) statement (Moher et
38 al., 2009) and the Synthesising Evidence from Diagnostic Accuracy Tests (SEDATe)
39 guidelines (Sotiriadis et al., 2016). The index test was US or MRI in patients with suspected
40 UCL injuries of the thumb. The reference standard for this review was either surgical

41 exploration, or joint stability of the thumb metacarpophalangeal joint during clinical
42 examination by a surgeon at follow-up, in a joint which was unstable at presentation.

43

44 **Literature search**

45 The literature search was conducted in MEDLINE, PubMed, Embase and Cochrane
46 CENTRAL electronic databases from inception until 23rd April 2019. The following keyword
47 search strategy was used: [“skier’s thumb” *or* “gamekeeper’s thumb” *or* “Stener” *or* “ulnar
48 collateral ligament”] *and* “ultrasound” *or* “ultrasonography”, *or* “magnetic resonance
49 imaging” *or* “MRI”. Database records were imported to Rayyan’s QCRI systematic review
50 tool for study screening and selection (Ouzzani et al., 2016).

51

52 **Eligibility criteria**

53 Inclusion criteria was as follows: (a) original studies which evaluated the diagnostic
54 accuracy of US or MRI for assessing thumb UCL injuries, (b) the reference standard for
55 studies was surgical exploration, in addition to clinical follow-up of joint stability (c) the
56 absolute number of true-positive (TP), false-positive (FP), false-negative (FN), true-negative
57 (TN) were reported, derivable or communicated by authors within two months upon request.
58 Non-English language, scientific abstract, and cadaver studies were excluded.

59

60 **Data extraction and risk of bias**

61 Identified studies were scrutinised in terms of publication authors, study institution,
62 recruitment period, and publication date to prevent inclusion of duplicate data. Data was
63 extracted on study design, total number of patients, number of patients who underwent
64 surgery or clinical follow-up, patient demographics, time interval between
65 injury/imaging/surgery/clinical follow-up, US technique, MRI protocol, and diagnostic test

66 accuracy data. Risk of bias was assessed using the Quality Assessment of Diagnostic
67 Accuracy Studies (QUADAS-2) tool (Whiting et al., 2011).

68

69 **Statistical analysis**

70 Diagnostic test accuracy data (TP, FP, FN, TN) was imputed into a 2 x 2 diagnostic matrix. A
71 random-effects meta-analysis, using a generalised linear mixed model, was performed to
72 calculate pooled sensitivity and specificity (Chu and Cole, 2006). Forest plots were
73 generated. Study heterogeneity was assessed by visual inspection of the forest plots. The
74 presence of threshold effect, when studies use different cut-offs to define a TP or TN result,
75 was investigated by a Spearman's rank correlation between sensitivity and sensitivity; a
76 significant negative correlation suggests threshold effect (Zamora et al., 2006). Publication
77 bias was not assessed because Deek's asymmetry test is not recommended when there are
78 less than ten studies per meta-analysis (Deeks et al., 2005; The Cochrane Collaboration,
79 2011). Statistical analysis was performed using Stata version 14.2 (StataCorp, 2015) and
80 Review Manager (RevMan) version 5.3 (The Cochrane collaboration, 2014).

81

82 **RESULTS**

83 **Study selection**

84 There were 929 titles identified from the search, of which 394 were duplicates. After title and
85 abstract screening, 26 articles proceeded to full text screening. Eligibility criteria was met by
86 eleven articles assessing US, five articles assessing MRI, and one article assessing both.
87 (Figure 2). From the eligible sixteen studies, fifteen studies (nine US and six MRI studies)
88 proceeded to meta-analysis. Specific considerations in study selection for meta-analysis are
89 outlined below.

90

91 The studies by Kohut and O'Callaghan (1993) and O'Callaghan et al (1994) shared the same
92 study location (Kohut and O'Callaghan, 1993; O'Callaghan et al., 1994). The former, smaller
93 study was excluded from US meta-analysis. Similarly, the study by Hergan et al (1995) of
94 both US and MRI patients was excluded from US meta-analysis, as it shared the same study
95 location as the larger US study by Hergan and Mittler (1995) (Hergan and Mittler, 1995;
96 Hergan et al., 1995). The study, however, is included in MRI meta-analysis.

97

98 In the study by Lohman et al (2001) there were four patients with equivocal Stener lesion
99 findings on either imaging or surgical exploration (Lohman et al., 2001). To permit inclusion
100 in the 2×2 diagnostic matrix, unclear cases were re-classified as positive findings. These
101 cases are likely to reflect high-grade injuries necessitating surgical repair, as with a Stener
102 lesion.

103

104 In the study by Harper et al (1996), fourteen patients who underwent MRI arthrography were
105 not included in meta-analysis, as this technique was not a consideration of this review
106 (Harper et al., 1996).

107

108 **Study characteristics**

109 Nine single-centre studies assessed US (Hergan and Mittler, 1995; Hoglund et al., 1995;
110 Melville et al., 2013; Murphey et al., 1997; Noszian et al., 1995; O'Callaghan et al., 1994;
111 Schnur et al., 2002; Shinohara et al., 2007; Susic et al., 1999) (Table 1). All, but two studies
112 (Melville et al., 2013; Schnur et al., 2002) were prospective. The number of patients included
113 in each study ranged from 14 to 69. The most frequently used US probe was a 7.5MHz linear
114 array transducer. Six studies recorded the time interval from injury to US, with a range
115 between four hours to 203 days (Hergan and Mittler, 1995; Hoglund et al., 1995; Melville et

116 al., 2013; Noszian et al., 1995; O'Callaghan et al., 1994; Susic et al., 1999). Two studies
117 recorded the time from US (or injury) to surgery, with a range of 1 to 405 days (Hoglund et
118 al., 1995; Melville et al., 2013). In five studies, clinical follow-up ranged from 17 days to 5
119 years (Hoglund et al., 1995; Murphey et al., 1997; Noszian et al., 1995; O'Callaghan et al.,
120 1994; Schnur et al., 2002; Susic et al., 1999).

121

122 Six single-centre studies assessed MRI (Harper et al., 1996; Hergan et al., 1995; Hinke et al.,
123 1994; Lohman et al., 2001; Milner et al., 2015; Romano et al., 2003) (Table 2). All but one
124 (Milner et al., 2015) were prospective. The number of patients undergoing MRI ranged from
125 5 to 43. Three studies recorded the time interval to MRI and this ranged from three days to
126 two years (Harper et al., 1996; Hinke et al., 1994; Romano et al., 2003). One study recorded
127 the time interval from MRI to surgery and this was between 10 weeks and over one year
128 (Lohman et al., 2001). Magnetic strength ranged between 0.5-1.5 Tesla. Four of six studies
129 reported using a radiofrequency surface coil (Harper et al., 1996; Hinke et al., 1994; Lohman
130 et al., 2001; Romano et al., 2003).

131

132 **Risk of bias**

133 All fifteen studies included in meta-analysis were found to have a high risk of bias (Figure 3).
134 The 'Patient Selection' domain was of high risk of bias in all studies. All but two studies
135 (Schnur et al., 2002; Susic et al., 1999) did not state whether a consecutive or random sample
136 of patients was recruited. Susic et al (1999) and Schnur et al (2002) recruited random
137 samples of patients, but did not record any study exclusion criteria (Schnur et al., 2002; Susic
138 et al., 1999). O'Callaghan et al (1994), Hoglund et al (1995) and Nozian et al (1995) all
139 included participants under the age of 18 years old (Hoglund et al., 1995; Noszian et al.,
140 1995; O'Callaghan et al., 1994). The 'Reference Standard' domain was of high risk of bias in

141 all studies except Susic et al (1999), primarily because it was unclear whether the person
142 undertaking surgical repair or follow-up physical examination was blinded to the imaging
143 findings (Susic et al., 1999). Additionally, only five studies reported blinding of the imaging
144 operator to clinical findings (Hergan and Mittler, 1995; Høglund et al., 1995; Noszian et al.,
145 1995; Lohman et al., 2001; Romano et al., 2003). The 'Flow and Timing' domain was found
146 to be of high risk of bias in all studies. The time interval to the reference standard was either
147 unclear or considered to be prolonged.

148

149 The overall clinical applicability concerns to the review question were low in all studies but
150 four (Harper et al., 1996; Lohman et al., 2001; Noszian et al., 1995; Susic et al., 1999)
151 (Figure 3). The 'Patient Selection' domain was of high concern in the study by Nozian et al
152 (1995) as some patients were identified as developing a full-thickness tear secondary to stress
153 radiography whilst in the studies by Harper et al (1996) and Lohman et al (2001) the injuries
154 were conspicuously chronic (Harper et al., 1996; Lohman et al., 2001; Noszian et al., 1995).
155 The 'Index Test' domain was of high concern in the studies by Lohman et al (2001) and
156 Harper et al (1996) as consensus was used to determine the imaging result (Harper et al.,
157 1996; Lohman et al., 2001). The 'Reference Standard' domain was of high concern in the
158 study by Susic et al (1999), as one patient with a possible dislocated ligament refused surgery
159 (Susic et al., 1999).

160

161 **Meta-analysis of US**

162 Nine studies ($n = 315$) assessed the diagnostic test accuracy of US. The pooled sensitivity
163 was 95% (95% confidence interval 76-99%), and pooled specificity was 94% (88-97%)
164 (Figure 4). Visual inspection of the forest plot showed evidence of large inter-study

165 variability in sensitivity, with relatively homogenous specificity. There was no evidence of
166 threshold effect to explain underlying heterogeneity (Spearman's $R = 0.20$, $P = 0.60$).

167

168 **Meta-analysis of MRI**

169 Six studies ($n = 107$) assessed the diagnostic test accuracy of MRI. The pooled sensitivity
170 was 93% (69 to 99%) and pooled specificity was 98% (70 to 100%) (Figure 4). There was no
171 evidence of threshold effect to explain underlying heterogeneity (Spearman's $R = 0.89$, $P =$
172 0.015).

173

174 **DISCUSSION**

175 This systematic review and meta-analysis demonstrates that both US and MRI have high
176 diagnostic accuracy for the detection of Stener lesions, although the current evidence stems
177 from sparse heterogenous studies with evidently high risk of bias.

178

179 US demonstrated high sensitivity (95%) and specificity (94%) for the diagnosis of a Stener
180 lesion. The study by Susic et al (1999) showed an exceptionally low sensitivity of 40% (Susic
181 et al., 1999). The utilised imaging technique was not reported in this study and consideration
182 should be given to the operator-dependant nature of US. The aforementioned Susic et al
183 (1999) study and the Shinohara et al (2007) study reported the lowest specificities of 78%
184 and 71% respectively (Shinohara et al., 2007; Susic et al., 1999). The latter study illustrates
185 important Stener lesion mimics, such as a ruptured aponeurosis or dorsal capsule, that may
186 falsely depict a Stener lesion on US (Shinohara et al., 2007). Nevertheless, the high
187 diagnostic yield of US along with its advantageous low cost, accessibility, and dynamic
188 imaging capabilities confirm this is an excellent first-line imaging modality for assessing the
189 position of the UCL in suspected Stener lesions.

190

191 MRI also demonstrated high diagnostic accuracy, with a pooled sensitivity of 93% and
192 specificity of 98% despite uncertainty in summary estimates. In particular, sensitivity was
193 100% in three studies (Harper et al., 1996; Hergan and Mittler, 1995; Milner et al., 2015) and
194 specificity was 100% in four of six studies (Harper et al., 1996; Hergan and Mittler, 1995;
195 Milner et al., 2015; Romano et al., 2003). In the earliest study by Hinke et al (1994) the
196 diagnostic test accuracy was lower than that of other studies (Hinke et al., 1994). In this
197 study, the first three cases were misdiagnosed on MRI but were correctly diagnosed
198 retrospectively following surgical correlation, suggesting a learning curve in MRI
199 interpretation (Hinke et al., 1994). Similarly, the study by Lohman et al (2001) of
200 predominately chronic UCL injuries reported lower sensitivity and specificity than other MRI
201 studies (Lohman et al., 2001). Excessive scarring and the absence of high signal oedema,
202 often seen with an acute ligamentous injury, can make it difficult to reliably interpret the
203 position of the chronically injured UCL, likely accounting for their low accuracy.
204 Unfortunately, the paucity of reportable data and the wide time interval between injury and
205 imaging reported in the studies of this review made it impossible to distinguish acute and
206 chronic UCL injuries. Future research is required to assess the diagnostic accuracy of US and
207 MRI in identifying acute and chronic Stener lesions respectively.

208

209 The superior contrast resolution and three-dimensional capabilities of MRI should allow
210 excellent assessment of intricate features UCL injuries. Two MRI studies in this review
211 frequently encountered full-thickness tears of the UCL with a 3mm or greater separation of
212 the torn ends, despite the ligament remaining below the aponeurosis (Milner et al., 2015;
213 Romano et al., 2003). Termed the ‘quasi-Stener’ by Milner et al (2015) the study
214 demonstrated that 90% of these injuries fail conservative management and require surgical

215 repair (Milner et al., 2015). Interestingly, a recent systematic review from our institution has
216 shown limited evidence and consensus on the treatment of full-thickness UCL injuries
217 (Mikhail et al., 2018). Further research is required to determine if specific imaging features
218 on MRI could play a role in therapeutic stratification, beyond the position of the torn UCL
219 relative to the aponeurosis.

220

221 The studies in this review are hindered by significant drawbacks. The sparse number of
222 patients, non-consecutive recruitment, and retrospective nature of some studies implicate
223 spectrum bias. Blinding was largely unclear and suboptimal. Studies frequently failed to
224 report time interval between injury, imaging, and the reference standard which has clear
225 implications for diagnostic accuracy. A proportion of studies are arguably outdated by
226 technological advances in imaging, raising wider clinical applicability concerns. Modern US
227 examination advocates the use of high frequency 12-15MHz transducers, often in conjunction
228 with a small footprint (hockey stick) probe to improve resolution (Singh et al., 2016). High
229 magnetic strength MRI scanners, such as 3 Tesla, are becoming more widely available and
230 allow for improved signal-to-noise ratio, thinner slice acquisition, and shorter examination
231 time compared to lower strength scanners (Malone et al., 2016). There are several important
232 limitations relating to the review and meta-analysis. The reference standard for this meta-
233 analysis was surgery (the 'gold standard') or resolution of joint instability at clinical follow-
234 up. However, joint instability does not necessarily imply an underlying Stener lesion and vice
235 versa. Because not all patients with UCL injuries can be reasonably expected to undergo
236 surgery, the resolution of acute symptoms and absence of joint instability at follow-up was
237 considered reasonable means to exclude a Stener lesion. Due to study heterogeneity, a
238 random-effects model was used. Therefore, it is likely that the large uncertainty obtained in
239 summary estimates, particularly for MRI, could be explained by sampling variation rather

240 than heterogeneity. The high uncertainty in summary estimates of sensitivity compared to
241 specificity for US supports this notion, with far fewer patients with the target condition (TP)
242 than without (TN).

243

244 In conclusion, both US and MRI demonstrate high diagnostic accuracy for the diagnosis of
245 Stener lesions of the thumb UCL. US is a reliable first line imaging modality, although
246 institutions are best suited to utilise their choice of imaging based on local availability and
247 radiological expertise. The results of this meta-analysis are limited by sparse data and high
248 bias, and require confirmation through future robust multi-centre studies analysing US and
249 MRI accuracy.

250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274

REFERENCES

Beutel BG, Melamed E, Rettig ME. The stener lesion and complete ulnar collateral ligament injuries of the thumb a review. *Bull Hosp Jt Dis* (2013). 2019, 77: 11-20.

Christensen T, Sarfani S, Shin AY, Kakar S. Long-term outcomes of primary repair of chronic thumb ulnar collateral ligament injuries. *Hand (N Y)*. 2016, 11: 303-9.

Chu H, Cole SR. Bivariate meta-analysis of sensitivity and specificity with sparse data: A generalized linear mixed model approach. *J Clin Epidemiol*. 2006, 59: 1331-2.

Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. *J Clin Epidemiol*. 2005, 58: 882-93.

Harper MT, Chandnani VP, Spaeth J, Santangelo JR, Providence BC, Bagg MA. Gamekeeper thumb: Diagnosis of ulnar collateral ligament injury using magnetic resonance imaging, magnetic resonance arthrography and stress radiography. *J Magn Reson Imaging*. 1996, 6: 322-8.

Hergan K, Mittler C. Sonography of the injured ulnar collateral ligament of the thumb. *J Bone Joint Surg Br*. 1995, 77: 77-83.

Hergan K, Mittler C, Oser W. Ulnar collateral ligament: Differentiation of displaced and nondisplaced tears with us and mr imaging. *Radiology*. 1995, 194: 65-71.

Hinke DH, Erickson SJ, Chamoy L, Timins ME. Ulnar collateral ligament of the thumb: Mr findings in cadavers, volunteers, and patients with ligamentous injury (gamekeeper's thumb). *AJR Am J Roentgenol*. 1994, 163: 1431-4.

Hoglund M, Tordai P, Muren C. Diagnosis by ultrasound of dislocated ulnar collateral ligament of the thumb. *Acta Radiol*. 1995, 36: 620-5.

Kohut G, O'Callaghan B. Gamekeeper's thumb. Ligament localisation by echography. *Ann Chir Main Memb Super*. 1993, 12: 257-61.

275 Lark ME, Maroukis BL, Chung KC. The stener lesion: Historical perspective and evolution of
276 diagnostic criteria. *Hand (N Y)*. 2017, 12: 283-9.

277 Lohman M, Vasenius J, Kivisaari A, Kivisaari L. Mr imaging in chronic rupture of the ulnar
278 collateral ligament of the thumb. *Acta Radiol*. 2001, 42: 10-4.

279 Lucerna A, Rehman UH. Stener lesion. *Statpearls*. Treasure Island (FL), StatPearls Publishing
280 StatPearls Publishing LLC., 2019.

281 Mahajan M, Rhemrev SJ. Rupture of the ulnar collateral ligament of the thumb - a review. *Int*
282 *J Emerg Med*. 2013, 6: 31.

283 Malone J, Preston K, Garner H, Bastic J, Wessell D, Jeffrey J. MRI of the hands and fingers.
284 *Appl Radiol*. 2016, 45: 9-21.

285 Mattox R, Welk AB, Battaglia PJ, Scali F, Nunez M, Kettner NW. Sonographic diagnosis of
286 an acute stener lesion: A case report. *J Ultrasound*. 2016, 19: 149-52.

287 Melville D, Jacobson JA, Haase S, Brandon C, Brigido MK, Fessell D. Ultrasound of displaced
288 ulnar collateral ligament tears of the thumb: The stener lesion revisited. *Skeletal Radiol*. 2013,
289 42: 667-73.

290 Mikhail M, Wormald JCR, Thurley N, Riley N, Dean B. Therapeutic interventions for acute
291 complete ruptures of the ulnar collateral ligament of the thumb: A systematic review.
292 *F1000Res*. 2018, 7: 714.

293 Milner CS, Manon-Matos Y, Thirkannad SM. Gamekeeper's thumb--a treatment-oriented
294 magnetic resonance imaging classification. *J Hand Surg Am*. 2015, 40: 90-5.

295 Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews
296 and meta-analyses: The prisma statement. *PLoS Med*. 2009, 6: e1000097.

297 Murphey SL, Hashimoto BE, Buckmiller J, Kramer D, Wiitala L. Ultrasonographic stress
298 testing of ulnar collateral ligament injuries of the thumb. *J Ultrasound Med*. 1997, 16: 201-7.

299 Noszian IM, Dinkhauser LM, Orthner E, Straub GM, Csanady M. Ulnar collateral ligament:
300 Differentiation of displaced and nondisplaced tears with us. *Radiology*. 1995, 194: 61-3.

301 O'Callaghan BI, Kohut G, Hoogewoud HM. Gamekeeper thumb: Identification of the stener
302 lesion with us. *Radiology*. 1994, 192: 477-80.

303 Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for
304 systematic reviews. *Syst Rev*. 2016, 5: 210.

305 Papandrea RF, Fowler T. Injury at the thumb ucl: Is there a stener lesion? *J Hand Surg Am*.
306 2008, 33: 1882-4.

307 Romano WM, Garvin G, Bhayana D, Chaudhary O. The spectrum of ulnar collateral ligament
308 injuries as viewed on magnetic resonance imaging of the metacarpophalangeal joint of the
309 thumb. *Can Assoc Radiol J*. 2003, 54: 243-8.

310 Schnur DP, DeLone FX, McClellan RM, Bonavita J, Witham RS. Ultrasound: A powerful tool
311 in the diagnosis of ulnar collateral ligament injuries of the thumb. *Ann Plast Surg*. 2002, 49:
312 19-22; discussion -3.

313 Shinohara T, Horii E, Majima M et al. Sonographic diagnosis of acute injuries of the ulnar
314 collateral ligament of the metacarpophalangeal joint of the thumb. *J Clin Ultrasound*. 2007, 35:
315 73-7.

316 Singh JP, Kumar S, Kathiria AV, Harjai R, Jawed A, Gupta V. Thumb ultrasound: Technique
317 and pathologies. *Indian J Radiol Imaging*. 2016, 26: 386-96.

318 Sotiriadis A, Papatheodorou SI, Martins WP. Synthesizing evidence from diagnostic accuracy
319 tests: The sedate guideline. *Ultrasound Obstet Gynecol*. 2016, 47: 386-95.

320 StataCorp. *Stata statistical software: Release 14*. College Station, TX, StataCorp LP, 2015,
321 Vol. 14.2.

322 Stener A. Displacement of the ruptured ulnar collateral ligament of the metacarpo-phalangeal
323 joint of the thumb: A clinical and anatomical study. *J Bone Joint Surg B*. 1962, 6: 869-79.

324 Susic D, Hansen BR, Hansen TB. Ultrasonography may be misleading in the diagnosis of
325 ruptured and dislocated ulnar collateral ligaments of the thumb. *Scand J Plast Reconstr Surg*
326 *Hand Surg.* 1999, 33: 319-20.

327 The Cochrane Collaboration. Recommendations on testing for funnel plot asymmetry. In:
328 Higgins J, Green S (Eds.) *Cochrane handbook for systematic reviews of interventions*, 2011,
329 Vol. 5.1.0.

330 The Cochrane collaboration. Review manager (revman). 5.3 ed. Copenhagen: The Nordic
331 Cochrane Centre, 2014.

332 Tresley J, Singer AD, Ouellette EA, Blachman J, Clifford PD. Multimodality approach to a
333 stener lesion: Radiographic, ultrasound, magnetic resonance imaging, and surgical correlation.
334 *Am J Orthop* (Belle Mead NJ). 2017, 46: E195-e9.

335 Whiting PF, Rutjes AW, Westwood ME et al. Quadas-2: A revised tool for the quality
336 assessment of diagnostic accuracy studies. *Ann Intern Med.* 2011, 155: 529-36.

337 Zamora J, Abaira V, Muriel A, Khan K, Coomarasamy A. Meta-disc: A software for meta-
338 analysis of test accuracy data. *BMC Med Res Methodol.* 2006, 6: 31.

339

TABLE AND FIGURE LEGENDS

Table 1: Characteristics of studies investigating the diagnostic accuracy of ultrasound in diagnosing Stener lesions compared to surgery or clinical follow up of joint stability

Table 2: Characteristics of studies investigating the diagnostic accuracy of magnetic resonance imaging (MRI) in diagnosing Stener lesions compared to surgery or clinical follow up of joint stability

Figure 1: Stener lesion of the thumb ulnar collateral ligament (A) Illustration of the ruptured ulnar collateral ligament (arrow) dislocated superficial and proximal to the adductor aponeurosis (arrowhead) (B) Coronal short tau inversion recovery (STIR) magnetic resonance imaging (MRI) of the thumb with a Stener lesion (C) Longitudinal axis ultrasound at the level of the first metacarpophalangeal joint with a Stener lesion. MC = Metacarpus; PP = Proximal Phalanx

Figure 2: Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram for identification and screening of studies using ultrasound or magnetic resonance imaging (MRI) for the detection of Stener Lesions. Nine ultrasound and six MRI studies were included in the diagnostic meta-analysis.

Figure 3. The quality of ultrasound ($n = 9$) and MRI ($n = 6$) studies included in meta-analysis as assessed using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool; Risk of bias is indicated as high (red) or low (green).

Figure 4. Forest plots of sensitivity and specificity for individual ultrasound studies ($n = 9$) and magnetic resonance imaging (MRI) studies ($n = 6$) in detecting Stener lesions. The reference standard is surgical exploration or joint stability on clinical follow-up.

SUPPLEMENTARY MATERIAL (not for online publication)

Supplementary material 1: PRISMA 2009 checklist

Supplementary material 2: Data (xlsx file)

Supplementary material 3: Data (csv file)

Supplementary material 4: Data analysing and coding (STATA file)