



LJMU Research Online

Willems, L, Vermeulen, J, Wiegerinck, A, Fekkes, S, Reijnen, M, Warlé, M, De Korte, C and Thijssen, D

Construct Validity and Reproducibility of Handheld Ultrasound Devices n Carotid Artery Diameter Measurement.

<http://researchonline.ljmu.ac.uk/id/eprint/18558/>

Article

Citation (please note it is advisable to refer to the publisher's version if you intend to cite from this work)

Willems, L, Vermeulen, J, Wiegerinck, A, Fekkes, S, Reijnen, M, Warlé, M, De Korte, C and Thijssen, D Construct Validity and Reproducibility of Handheld Ultrasound Devices n Carotid Artery Diameter Measurement. Ultrasound in medicine & biooav. ISSN 0301-5629 (Accepted)

LJMU has developed [LJMU Research Online](#) for users to access the research output of the University more effectively. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LJMU Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain.

The version presented here may differ from the published version or from the version of the record. Please see the repository URL above for details on accessing the published version and note that access may require a subscription.

For more information please contact researchonline@ljmu.ac.uk

<http://researchonline.ljmu.ac.uk/>

1 **Construct validity and reproducibility of handheld ultrasound**
2 **devices in carotid artery diameter measurement**

3 LH Willems, MD[1]*; JJM Vermeulen, MD[2,3]*; AIP Wiegerinck[2]; S Fekkes[4]; MMPJ
4 Reijnen, MD, PhD [3,5]; MC Warlé, MD, PhD[1]; CL de Korte, PhD[4]; DHJ Thijssen,
5 PhD[2,6]

6 1) Department of Surgery, Radboud University Medical Centre, Nijmegen, The Netherlands;

7 2) Department of Physiology, Radboud Institute for Health Sciences, Radboud University

8 Medical Centre, Nijmegen, The Netherlands; 3) Department of Surgery, Rijnstate Hospital,

9 Arnhem, the Netherlands; 4) Department of radiology and nuclear medicine, Radboud

10 University Medical Centre, Nijmegen, The Netherlands; 5) Multi-Modality Medical Imaging

11 Group, TechMed Centre, University of Twente, Enschede, The Netherlands; 6) Research

12 Institute for Sport and Exercise Sciences, Liverpool John Moores University, Liverpool, United

13 Kingdom

14 *Shared first authors

15

16 **Corresponding author:** Loes Willems

17 Postbus 9101 (intern 618), 6500 HB Nijmegen

18 Phone number: +31 24 361 53 33. Fax number: +31 24 363 51 15

19 E-mail: loes.h.willems@radboudumc.nl

20

21

22 **Abstract**

23 The construct validity and reproducibility of three commonly used handheld ultrasound (US)
24 devices in measuring carotid arterial diameter was evaluated: Telemed MicrUs EXT-1H
25 (Telemed, Vilnius, Lithuania), Butterfly iQ (Butterfly Network, Inc, Guilford, CT) and Philips
26 Lumify (Philips Healthcare, Best, The Netherlands). An *in vitro* set-up was built to evaluate
27 construct validity, compared to high-end US, and intra-observer variability of handheld US.
28 Handheld devices showed a mean difference of 0.023 ± 0.030 cm, 0.012 ± 0.037 cm and
29 0.009 ± 0.046 cm for respectively Telemed, Butterfly and Lumify in comparison to high-end US.
30 Intraclass agreement with the high-end system as well as intra-observer variability for handheld
31 US devices was classified as excellent with all values above 0.95. Subsequently, interobserver
32 variability of handheld US was investigated in an *in vivo* set-up with 20 healthy volunteers.
33 Interobserver variability was classified as excellent for Telemed (0.901), good for Lumify
34 (0.827), and moderate for Butterfly (0.684) with a difference of respectively 0.005 ± 0.031 cm,
35 0.020 ± 0.050 cm and -0.003 ± 0.033 . In conclusion, handheld US demonstrated an excellent
36 construct validity and intra-observer variability. Additionally, excellent-to-good interobserver
37 variability for Telemed and Lumify was observed, where Butterfly demonstrated a moderate
38 interobserver agreement. These results indicate that handheld US devices are applicable in
39 measuring carotid arterial diameter.

40

41

42 **Keywords** Handheld ultrasound, ultrasonography, carotid arteries, diameter

43

44 **Introduction**

45 Endothelial dysfunction is one of the first signs of systemic atherosclerosis and contributes to
46 its progression by promoting coagulation, vasoconstriction, and deficient vascular repair,
47 ultimately leading to thickening of the arterial wall with narrowing of conduit arteries as result
48 (Lerman en Zeiher 2005; Bonetti 2003). Measuring arterial diameter changes in response to
49 physiological stimuli, such as shear stress (e.g., flow-mediated dilation) and sympathetic
50 stimulation (e.g., carotid artery reactivity), using ultrasound (US) has emerged useful to assess
51 endothelial dysfunction (Nabel 1988; van Mil 2017; van Mil 2018; Peace 2018; van Mil 2019).

52

53 Arterial diameter measurements during endothelial function testing currently depends on high-
54 end US machines. High costs and the static nature of these machines prevent the applicability
55 of these measurements at first- and second-line clinical centers. Past decades, an increasing
56 number of clinicians started using handheld US. (van den Heuvel 2018; Zieleskiewicz 2021)
57 Important advantages of handheld US include their lower costs in comparison to high-end US
58 devices and their simplicity of use, which makes handheld US applicable in outpatient clinics
59 and general practices. Moreover, handheld US may facilitate the implementation of the
60 assessment of artery diameters and diameter responses to physiological responses. To date, little
61 is known about the validity and reproducibility of contemporary handheld US to examine
62 arterial diameter.

63

64 The purpose of this study is to evaluate the construct validity and reproducibility of three
65 commonly used handheld US devices (Telemed MicrUs EXT-1H (Telemed, Vilnius,
66 Lithuania), Butterfly iQ (Butterfly Network, Inc, Guilford, CT) and Philips Lumify (Philips
67 Healthcare, Best, The Netherlands)) in measuring carotid arterial diameter. For this purpose,
68 first, *in vitro* evaluation of handheld US devices in a phantom set-up was performed to evaluate

69 the construct validity of handheld US devices in comparison to a high-end US device.
70 Subsequently, experiments were performed, comparing intra- and interobserver variability of
71 the handheld US devices within respectively an *in vitro* and *in vivo* set-up.

72

73 **Materials and Methods**

74 Design. In the first part of this study, the construct validity of handheld US devices was
75 evaluated using an *in vitro* setting to create a controlled environment with fixed parameters like
76 acoustic (speed of sound, acoustic impedance and attenuation, backscattering) (zell 20117) and
77 mechanical (tissue elasticity and viscosity) (Amador 2011) tissue properties for diameter
78 detection of the US devices. In total, 28 measurements were performed per US device, which
79 were compared against a contemporary high-end US machine. Measurements were repeated on
80 a second day to evaluate the intra-observer variability. In the second part of this study, repeated
81 measurements of the carotid artery diameter were performed within twenty healthy individuals.
82 The carotid artery was chosen for diameter assessment, because the carotid artery is easily
83 accessible by US and commonly used for the evaluation of atherosclerosis development.
84 (Podgorski 2016)

85

86 Handheld US devices. The following three commonly used handheld US devices were used to
87 evaluate construct validity and intra- and interobserver reproducibility: 1) Telemed MicrUs
88 EXT-1H (Telemed, Vilnius, Lithuania) with a linear array probe with frequency range 5-12
89 MHz, 2) Butterfly iQ (Butterfly Network, Inc, Guilford, Connecticut, United States) with single
90 probe emulating a linear and phased array probe by means of microsensors with frequency
91 range 1-10 MHz, and 3) Philips Lumify (Philips Healthcare, Best, The Netherlands) with a
92 linear array probe with frequency range 4-12 MHz. To evaluate construct validity using the *in*
93 *vitro* setting, handheld US machines were compared against a high-end US system with a linear

94 array probe with frequency range 5-14 MHz(Terason 3300, Terason Ultrasound, Burlington,
95 MA, USA).

96

97 *In vitro: Construct validity and intra-observer variability*

98 *Experimental setup.* An experimental setup was built to perform US measurements on a custom-
99 made flexible polyvinyl alcohol phantom mimicking an artery; Figure 1 shows a schematic
100 overview. The phantom artery was positioned in an US compatible box (water basin) and
101 connected to an in-house built circulatory system with physiological flow and pressure
102 conditions (Fekkes 2018). Different flow volumes were applied to simulate different phantom
103 diameters.

104

105 *Measurement protocol.* The gear pump, connected to the phantom artery circulation, was set at
106 a continuous flow of 0.3 L/min. The US transducer was longitudinally aligned with the phantom
107 artery and this position was maintained by use of a laboratory standard. Basic carotid
108 ultrasonography pre-sets were used. Gain and depth were adapted when considered necessary.
109 Consensus of the optimal position and settings was reached by two skilled sonographers (JV,
110 LW) and was kept the same for each device. The phantom artery was recorded during a 10
111 second interval. Hereafter, the flow was increased with 0.1 L/min, corresponding with
112 approximately 1 millimeter diameter increase per minute, and the phantom artery was recorded
113 again. These steps were repeated to a flow of 0.9 L/min. Subsequently, the pressure regulator
114 was set on a pulsatile flow of 0.3 – 0.9 L/min, with 60 pulses per minute, with the phantom
115 artery being recorded for 10 second periods. These procedures were repeated for all devices.

116

117 Measurements were repeated on a second day, which was performed within 30 days, to
118 determine the intra-observer variability. We ensured that all procedures were kept similar,
119 including the order of testing.

120

121 *In vivo: Interobserver variability*

122 *Participants.* A total of 20 volunteers were recruited. Inclusion criteria were age between 18
123 and 65 years and a body mass index of 18-30 kg/m². No subjects with previously diagnosed
124 carotid artery occlusive disease were included. Written informed consent was obtained prior to
125 participation from all volunteers. Approval of the local Medical Ethical Committee (study
126 number: CMO 2020-6700) and the local Institutional Review Board was obtained. This study
127 was conducted in accordance with the latest revision of the Helsinki Declaration of 1964.

128

129 *Procedures.* Data about gender, age, height, weight, smoking behavior, medical history, and
130 the familial occurrence of cardiovascular diseases were collected. Participants visited the
131 hospital once. During the visit, ultrasound measurements of the common carotid artery were
132 performed. Participants were in supine position with the neck extended and had rested at least
133 5 minutes before the start of US measurements. Room temperature was kept constant and only
134 one type of US gel was used. The left common carotid artery was longitudinally visualized
135 using the three handheld US devices and one high-end US device, which were applied in
136 randomized order. After image optimisation by the examiner (JV, LW), the carotid artery
137 diameter was recorded for 10 seconds. Subsequently, the probe was removed from the
138 participant and handed over to the second experienced examiner without adjusting ultrasound
139 settings. This was followed by the second 10-second recording of the carotid artery diameter.
140 The order of the two examiners was also randomized.

141

142 Diameter analysis

143 Dependent on US device, data were saved as or converted to an Audio Video Interleave (AVI)
144 file. US videos of the Butterfly device were converted using Movavi Video Converter 20
145 (Movavi Software, Wildwood, USA) using the original size (resolution 1696 x 1080) and
146 MPEG-4 codec. Additionally, US videos of the Lumify device were converted using MatLab
147 R2018b (Mathworks, Natick, MA, USA) using the VideoWriter function with quality index 90.
148 This resulted in a video resolution varying from 512 x 296 till 512 x 444 depending on the depth
149 setting during the measurement. For the Terason ultrasound videos, Camtasia (Camtasia
150 Softonic, Barcelona, Spain) was used to record the screen containing ultrasound images. This
151 was saved as an AVI file with a resolution of 1024 x 768. The Telemed ultrasound video was
152 directly saved as AVI file with a resolution of 1556 x 868.

153

154 Diameter analysis of the recorded US videos of the phantom and carotid arteries was performed
155 by a single-blinded investigator using BloodFlow Software (Version 4.0; National Instruments
156 LabVIEW, Austin, TX, USA) with semiautomated edge-detection and wall-tracking algorithm.
157 This software enables the identification of a region of interest (ROI) in the longitudinal plane
158 of an artery. ROIs were identified for each US video. Within the ROI, the lumen-arterial wall
159 interface was detected (Figure 2). The diameter was determined multiple times per frame
160 depending on the size of the ROI. Subsequently, a median diameter per frame was determined
161 and eventually a median diameter of all frames was determined for the resulting diameter per
162 measurement. For the resulting diameter, full cardiac cycles were included to minimize bias of
163 the average diameter. More details on this technique are described previously. (Thijssen 2009)
164 The software is largely independent of investigator bias. (Woodman 2001)

165

166 Statistical analysis

167 Phantom and carotid artery diameters were reported as mean + standard deviation (SD) for each
168 measurement. Baseline characteristics of the participants were reported as median with
169 interquartile range [Q1, Q3] and categorical variables are presented as percentages. Bland-
170 Altman plots were created to determine the agreement of measured diameters between the
171 handheld devices and the high-end US device and to determine the intra- and interobserver
172 variability of the three handheld US devices for *in vitro* and *in vivo* measurements. Differences
173 were plotted against the mean per comparison. Bland-Altman plots are visualized with one
174 continuous black line representing the mean and two dotted lines representing the limits of
175 agreement (1.96*standard deviation). (Altman and Bland 1983) Variability of measurements
176 was assessed using the intra- and interobserver variability by determining intraclass correlation
177 coefficient (ICC), which was presented for respectively the between-day comparison for the *in*
178 *vitro* set-up and between-observers comparison for the *in vivo* set-up. ICC were reported
179 according to the guideline of Koo and Li (2016), where a coefficient <0.50, between 0.50 and
180 0.75, between 0.75 and 0.90 and >0.90 represents respectively a poor, moderate, good and
181 excellent agreement. Additionally, coefficients of variation were calculated per participant, per
182 device, between observers, with the ratio of the standard deviation and the mean absolute
183 differences between observers. After Bonferroni correction, p-values <0.01 were considered as
184 significant. Statistical analysis was performed in IBM SPSS Statistics version 25 (IBM
185 Corporation, Armonk, NY, USA).

186

187 **Results**

188 *In vitro*: Construct validity and intra-observer variability

189 The Bland-Altman plots for variability in *in vitro* measurements between handheld devices and
190 the high-end US device are shown in Figure 3. Compared to the high-end US device the
191 Telemed demonstrated a significantly larger diameter (0.023 ± 0.030 cm, $p < 0.001$, Table 1),

192 whilst no such difference was reported for the Butterfly (0.012 ± 0.037 cm) or Lumify
193 (0.009 ± 0.046 cm). Visually inspecting the Bland-Altman plots, we found comparable limits of
194 agreement across a large range of diameters between the three handheld US devices. ICC
195 comparing handheld US and high-end US was 0.996, 0.994 and 0.990 for Telemed, Butterfly
196 and Lumify, respectively.

197

198 No significant difference was found between measurement days for the Telemed
199 (0.013 ± 0.059 cm) and Butterfly (-0.012 ± 0.048 cm), whilst a small, but significant, difference
200 was found for the Lumify (0.008 ± 0.009 cm, $p=0.008$, Table 1). Bland-Altman plots (Figure 4)
201 reveal comparable limits of agreement across the three handheld US devices. ICC comparing
202 both measurements per handheld US device was 0.986, 0.990 and 1.000 for Telemed, Butterfly
203 and Lumify, respectively.

204

205 *In vivo: Interobserver variability*

206 Median age of the participants was 21.0 [IQR 20.0, 22.0] years and 40.0% was male.
207 Additionally, median BMI was 21.7 [IQR 20.4, 23.6], 10% was current smoker and 45% had a
208 family history of cardiovascular disease. Bland-Altman plots for *in vivo* measurements
209 comparing the interobserver variability of the handheld US devices are shown in Figure 5. No
210 significant difference in carotid artery diameter was found between operators for the Telemed
211 (0.005 ± 0.031 cm), Butterfly (0.020 ± 0.050 cm) or Lumify (-0.003 ± 0.033 cm, Table 1, Figure
212 5). Limits of agreement were smallest for the Lumify, with similar patterns and limits observed
213 for the Telemed and Butterfly. ICC for carotid artery diameter between the operators per device
214 was classified as excellent for the Telemed (0.901), good for Lumify (0.827), and moderate for
215 the Butterfly (0.684). Average coefficients of variation per participant, per device between
216 observers were 2.4%, 3.5% and 5.2% for Telemed, Lumify and Butterfly, respectively.

217

218 **Discussion**

219 This study has demonstrated that the three studied handheld devices show a good construct
220 validity and strong ICC compared with high-end US and excellent between-day intra-observer
221 variability using an *in vitro* setting for measuring arterial diameters. Between-observer
222 reproducibility of the handheld US devices within the *in-vivo* setting revealed an excellent-to-
223 good interobserver variability for the Telemed and Lumify, but a moderate variability for the
224 Butterfly.

225

226 Good consistency and excellent reliability were observed between handheld and high-end US
227 devices in an *in-vitro* setting, as all ICCs were well above 0.95. Nonetheless, a significant
228 difference between Telemed and the high-end US device was found, which may suggest limited
229 validity of the Telemed. One possible reason for this difference is (not) taking the intima-media
230 thickness into account when analyzing the diameter. Such consistent difference in determining
231 the diameter may result in structural difference between US devices. An example of this can be
232 seen in Figure 2, where the Lumify analyses detects the intima and the other devices detect the
233 outer wall. Furthermore, it is important to realize that Telemed demonstrated the smallest SD.
234 Taken this together, all three handheld US devices showed excellent construct validity.

235 Although *in vitro* set-ups are commonly used to determine validity of US devices, only
236 few studies focused on understanding (construct) validity using an *in vitro* set-up for handheld
237 US devices. Two studies were found comparing US devices. One study investigated carotid
238 strain assessment applying US speckle tracking using a clinical and high-end US device
239 (Larsson 2015), where the other study investigated optic nerve sheath diameters using a pocket-
240 sized US unit compared to a previously validated portable unit (Johnson 2016). Both studies
241 showed an ICC of respectively 0.73 at lowest for the clinical US device against 0.90 at lowest

242 for the high-frequency US device (Larsson 2015) and 0.75 for between device comparison and
243 0.83 for interobserver variability of the pocket-sized US device (Johnson 2016), which seems
244 slightly lower than the results presented in our study. Importantly, these previous studies
245 focused on other outcome measures. Other studies that evaluated the validity of handheld US
246 directly compared handheld US devices with each other (van den Heuvel 2018; Prekker 2013;
247 Niu 2019) or adopted other imaging modalities (Viadakovic 2007) using patients. A strength of
248 our study is therefore that the handheld US devices were both tested in *in vitro* set-up and
249 afterwards evaluated *in vivo* in volunteers.

250

251 In line with our results, other studies reporting on vascular US, have positively addressed the
252 use of handheld US devices (e.g. Acuson P10 (Stock 2015), Vscan (Mantella 2019) and
253 Butterfly (Alfuraih 2021)). Importantly, US devices were tested in relation to varying
254 pathological screening areas (e.g. size of liver, spleen and kidneys (Stock 2015), carotid artery
255 plaques (Mantella 2019) and abdominal aorta (Alfuraih 2021)). At least, these studies provide
256 further support that handheld US are feasible and reliable, with an ICC of ~0.8 with high-end
257 systems. (Stock 2015; Mantella 2019; Alfuraih 2021) However, the validity and reproducibility
258 must be considered within its specific use, which was related to the carotid artery diameter in
259 our study.

260 In contrast to the interobserver variability of the Telemed and Lumify, we found a
261 moderate variability of the Butterfly device. This latter observation may, at least in part, be
262 explained by the US transducer specifications of the Butterfly. Whilst Telemed and Lumify
263 utilize a classic linear array probe, the butterfly probe is differently shaped and emulates a linear
264 array probe by means of microsensors. The relatively small probe head of the Butterfly device,
265 allows for more variability in probe positioning, possibly resulting in some inter-operator
266 variability. Evaluation of arterial diameter is influenced by probe positioning (more proximal

267 or distal), but also artery shape, blood pressure variation, and tissue properties. (Beales 2011;
268 Mathiesen 2000; Triboulet 2006) Therefore, inter-operator variability *in vivo* can be
269 multifactorial and does not necessarily indicate lack of quality of the US device. Accordingly,
270 it is important to highlight that the Butterfly device has already proven to have a good
271 interobserver variability in assessing carotid artery plaque assessment (Alfuraih 2021). This
272 highlights the importance of (construct) validity studies for the large range of handheld US
273 devices, as device specifications may importantly determine the potential (clinical) application
274 of a specific US device.

275

276 A limitation of this study relates to analyzing standard B-mode images instead of using raw
277 radiofrequency data. The latter has a higher spatial resolution and might be preferred as golden
278 standard. Previous studies, however, have shown standard B-mode images to be robust for
279 measuring arterial characteristics with good precision and accuracy. (Steinbuch 2016; Dogan
280 2009). Using standard B-mode based analysis made it possible to make the analysis comparable
281 and consistent between the three handheld US devices. However, to optimize the US videos for
282 analysis, ultrasound settings were not standardized between devices or participants, which
283 could have had an impact on the final results.

284 B-mode images obtained from the various US machines had differences in format and quality.
285 Some US videos had to be converted to AVI files, which may have caused loss of quality of the
286 US videos (specifically affecting Lumify and Butterfly). The use of a reliable converter
287 software and converting packages effectively minimized loss of quality, further supported by
288 visual inspection of the US videos after conversion. Our software has proven to be reliable and
289 largely independent of investigator bias (Woodman 2001). Woodman et al. described the
290 method of analysis as well as some coefficient of variations for different determined parameters
291 with the software, with the largest conducted coefficient of variation being 6.7%. However, we

292 cannot fully exclude a bias caused by different types of videos obtained with the different US
293 machines. Nevertheless, since the quality of ultrasound devices have also been improved in the
294 last two decades just as converting software, the impact of this quality, e.g. image resolution
295 and video compression, on the analysis with this software is expected to be minimized.
296 Importantly, despite this possible bias, all devices showed excellent construct validity compared
297 to the high-end US device and excellent between-day reproducibility. Another limitation could
298 be the small sample size for Bland-Altman analysis (Lu 2016). Due to the explorative character
299 of the in-vivo part of this study, no sample size calculation was performed.

300

301 **Conclusion**

302 In conclusion, all handheld devices showed an excellent construct validity and intra-observer
303 variability *in vitro* and are therefore suitable to analyze carotid artery diameter. Interobserver
304 variability *in vivo* of the handheld devices were excellent-to-good for Telemed and Lumify,
305 where Butterfly showed a moderate variability. Although analysis software has proven to be
306 reliable, Butterfly and Lumify did not provide compatible US video, which could have caused
307 minor variation between the handheld devices. Nevertheless, this study demonstrated that
308 handheld US devices, especially Telemed and Lumify, are applicable in measuring carotid
309 arterial diameter.

310

311 **Acknowledgements**

312 None.

313

314 **Conflicts of interest statement**

315 The authors declare that there are no conflicts of interest.

316 **Figures Captions list**

317 *Figure 1.* Schematic overview of experimental setup of the in vitro experiment, where water
318 from the water reservoir was pumped around by the gear pump through the silicon tube which
319 was placed in a water basin. The probe of each ultrasound device was mounted in the laboratory
320 standard and positioned above the silicon tube such that a longitudinal plane was visualized

321

322 *Figure 2.* The detected borders of the lumen-arterial wall interface in participants for A)
323 Terason, B) Telemed, C) Butterfly and D) Lumify device, where the yellow square represents
324 the drawn ROI and the yellow lines represent the detected border.

325

326 *Figure 3.* Bland-Altman plots to compare assessment of the phantom diameters of the A)
327 Telemed, B) Butterfly and C) Lumify handheld US against the high-end US device (Terason),
328 where the continuous black line represents the mean difference and the dotted black lines
329 represent the limits of agreement per comparison

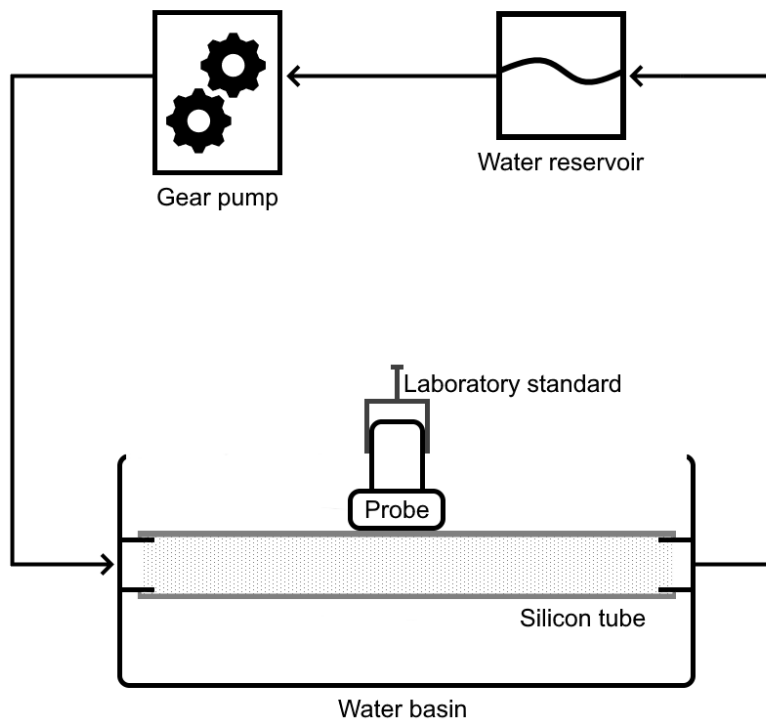
330

331 *Figure 4.* Comparison of the between-day variation of the in-vitro measurement of diameter for
332 A) Telemed, B) Butterfly and C) Lumify device, where the continuous black line represents the
333 mean difference and the dotted black lines represent the limits of agreement per comparison

334

335 *Figure 5.* Bland-Altman plots in-vivo measurements of the carotid diameter comparing both
336 operators using the A) Telemed, B) Butterfly and C) Lumify, where the continuous black line
337 represents the mean difference and the dotted black lines represent the limits of agreement per
338 comparison

339

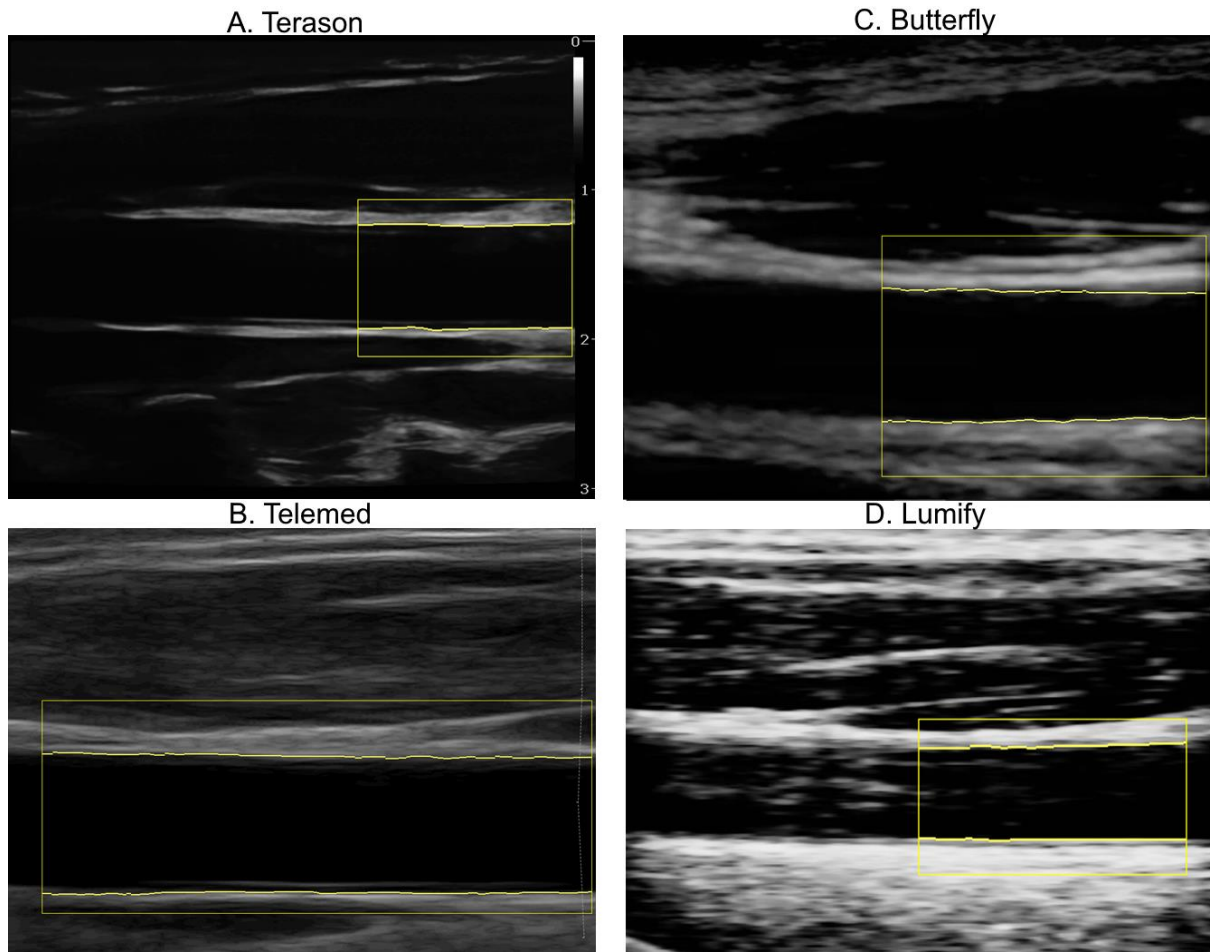


341

342 *Figure 1.* Schematic overview of experimental setup of the in vitro experiment, where water
343 from the water reservoir was pumped around by the gear pump through the silicon tube which
344 was placed in a water basin. The probe of each ultrasound device was mounted in the laboratory
345 standard and positioned above the silicon tube such that a longitudinal plane was visualized

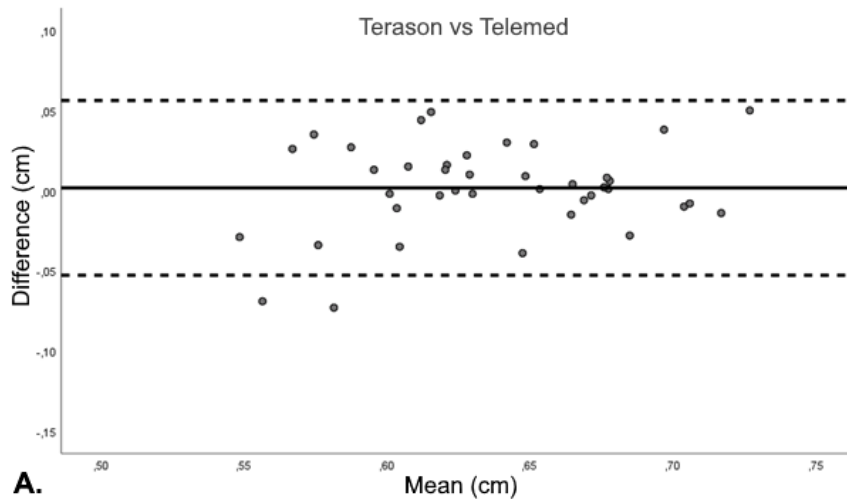
346

347

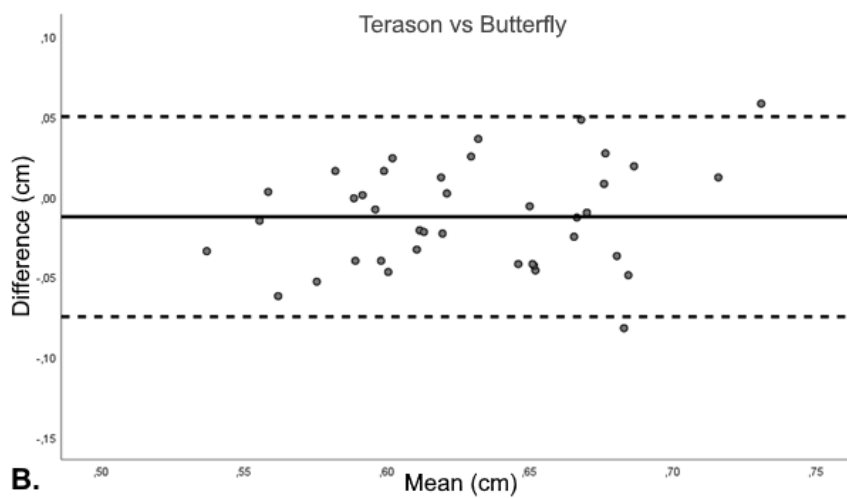


348

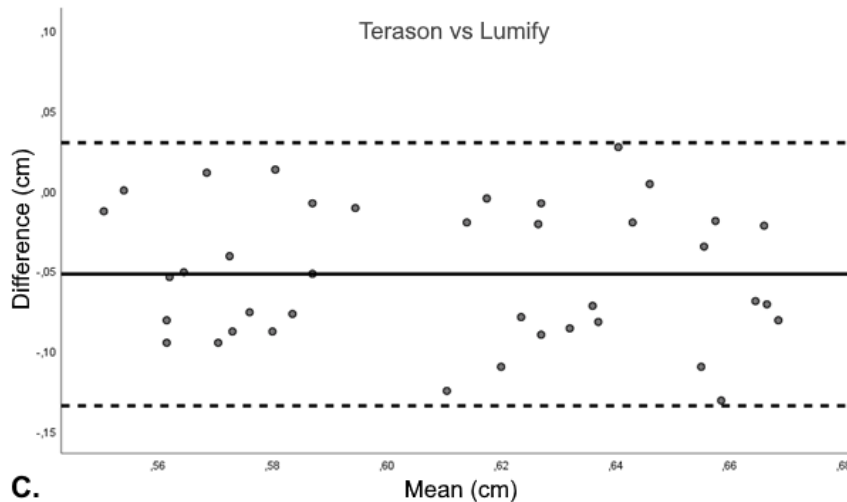
349 *Figure 2.* The detected borders of the lumen-arterial wall interface in participants for A)
 350 Terason, B) Telemed, C) Butterfly and D) Lumify device, where the yellow square represents
 351 the drawn ROI and the yellow lines represent the detected border.



A.



B.



C.

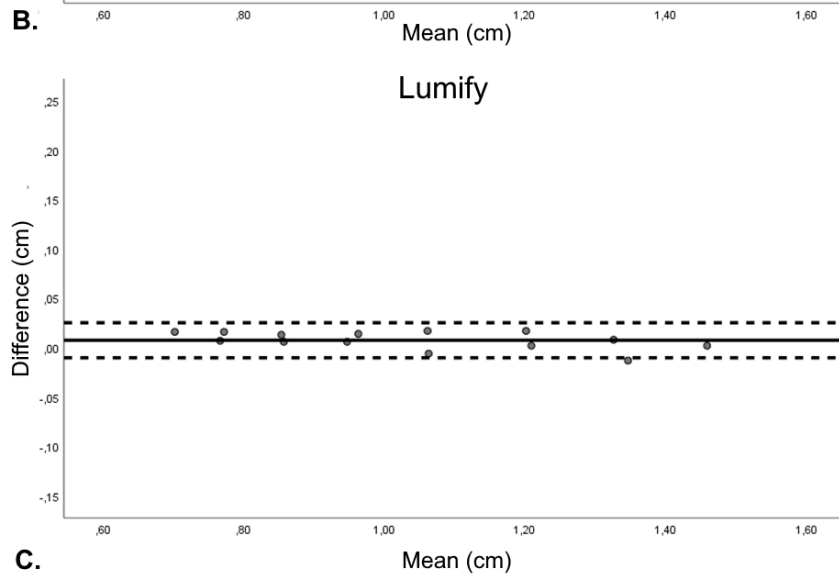
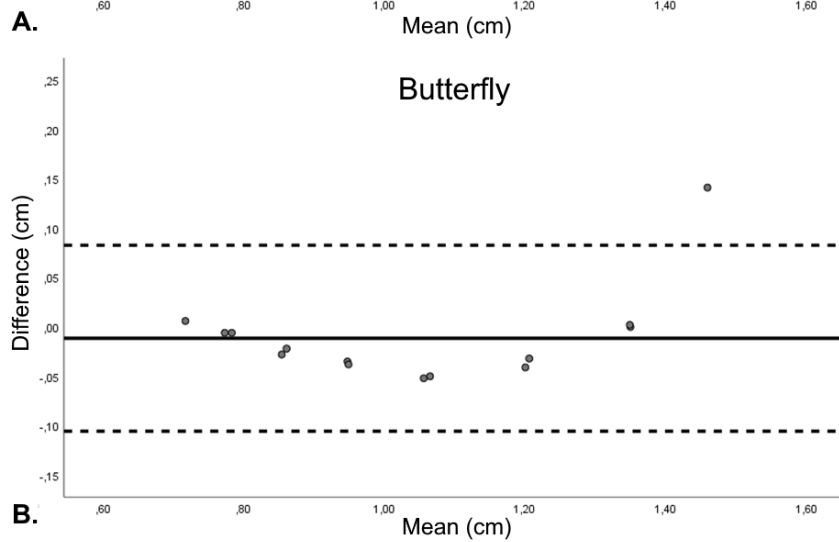
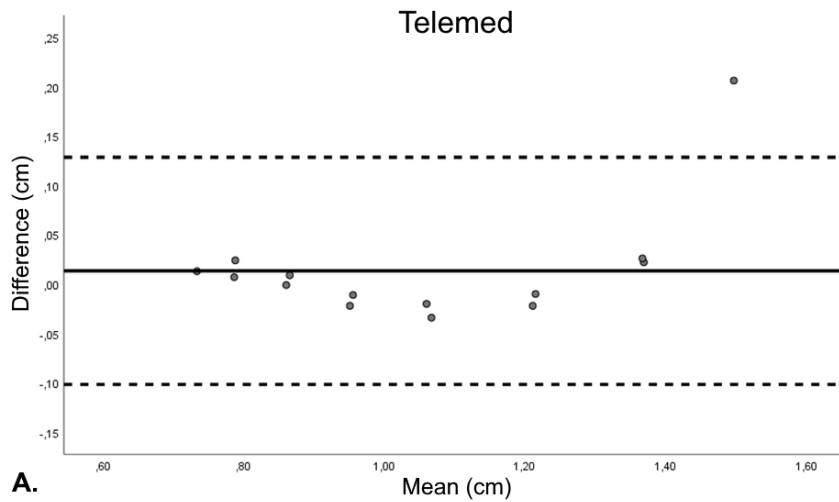
352

353 *Figure 3.* Bland-Altman plots to compare assessment of the phantom diameters of the A)

354 Telemed, B) Butterfly and C) Lumify handheld US against the high-end US device (Terason),

355 where the continuous black line represents the mean difference and the dotted black lines

356 represent the limits of agreement per comparison



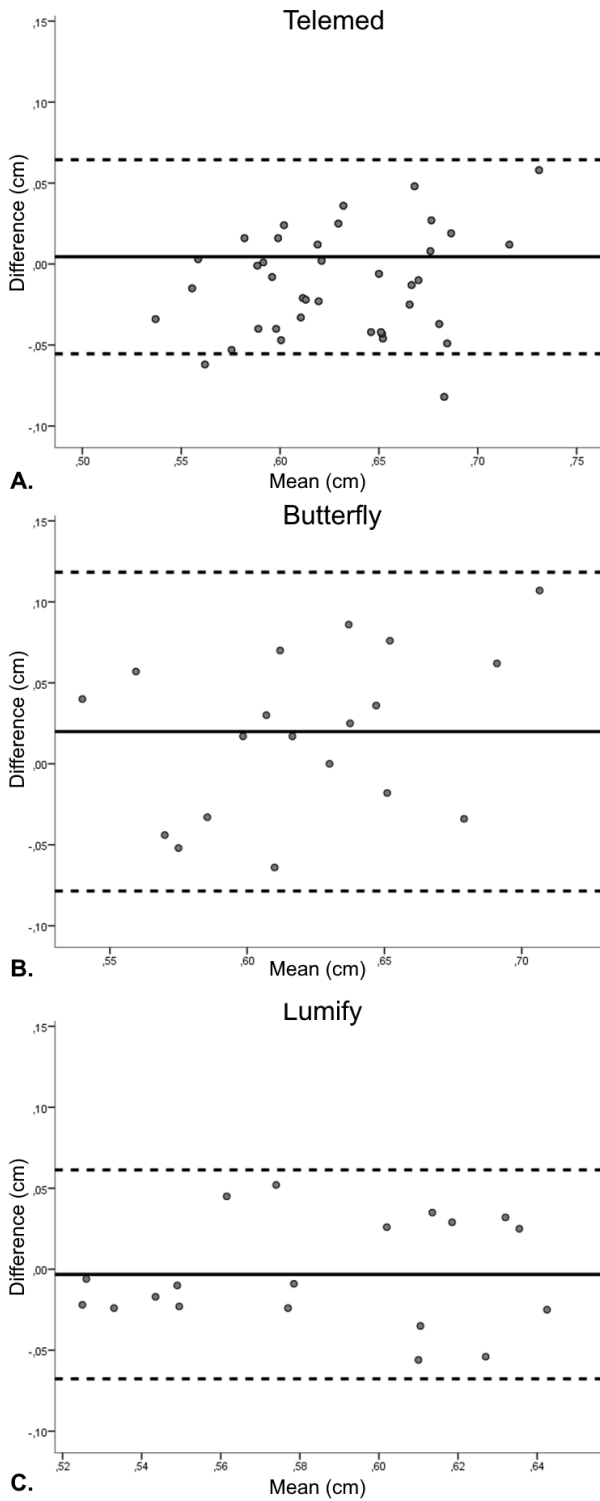
357

358 *Figure 4.* Comparison of the between-day variation of the in-vitro measurement of diameter for

359 A) Telemed, B) Butterfly and C) Lumify device, where the continuous black line represents the

360 mean difference and the dotted black lines represent the limits of agreement per comparison

361



363

364 *Figure 5.* Bland-Altman plots in-vivo measurements of the carotid diameter comparing both

365 operators using the A) Telemed, B) Butterfly and C) Lumify, where the continuous black line

366 represents the mean difference and the dotted black lines represent the limits of agreement per
367 comparison

368 **Tables**

369 *Table 1.* P-values for Bland-Altman analysis

	In-vitro validation	In-vitro variability	In-vivo variability
<i>Telemed</i>	<0.001	0.410	0.514
<i>Lumify</i>	0.0303	0.008	0.676
<i>Butterfly</i>	0.089	0.387	0.101

370

371

372

373 **References**

- 374 Alfuraih AM, Alrashed AI, Almazyad SO, Alsaadi MJ. Abdominal aorta measurements by a
375 handheld ultrasound device compared with a conventional cart-based ultrasound
376 machine. *Ann Saudi Med* 2021; 41: 376-382. DOI: 10.5144/0256-4947.2021.376
- 377 Altman DG, Bland JM. *Measurement in Medicine: the analysis of method comparison studies.*
378 JSTOR 1983; 32: 307-317.
- 379 Amador C, Urban MW, Chen S, Chen Q, An KN, Greenleaf JF. Shear elastic modulus
380 estimation from indentation and SDUV on gelatin phantoms. *IEEE Trans Biomed Eng*
381 2011; 58: 1706-1714. DOI: 10.1109/TBME.2011.2111419
- 382 Beales L, Wolstenhulme S, Evans JA, West R, Scott DJA. Reproducibility of ultrasound
383 measurement of the abdominal aorta. *Br J Surg* 2011; 98: 1517-1525. DOI:
384 10.1002/bjs.7628
- 385 Bonetti PO, Lerman LO, Lerman A. Endothelial dysfunction: a marker of atherosclerotic risk.
386 *Arterioscler Thromb Vasc Biol* 2003; 23: 168-175. DOI:
387 10.1161/01.atv.0000051384.43104.fc
- 388 Dogan S, Plantinga Y, Dijk JM, van der Graaf Y, Grobbee DE, Bots ML; SMART Study Group.
389 Manual B-mode versus automated radio-frequency carotid intima-media thickness
390 measurements. *J Am Soc Echocardiogr.* 2009 Oct;22(10):1137-44. doi:
391 10.1016/j.echo.2009.07.008. PMID: 19801303.
- 392 Fekkes S, Saris A, Nillesen MM, Menssen J, Hansen HHG, de Korte CL. Simultaneous
393 Vascular Strain and Blood Vector Velocity Imaging Using High-Frequency Versus
394 Conventional-Frequency Plane Wave Ultrasound: A Phantom Study. *IEEE Trans*
395 *Ultrason Ferroelectr Freq Control* 2018; 65: 1166-1181. DOI:
396 10.1109/TUFFC.2018.2834724

397 van den Heuvel TLA, de Bruijn D, Moens-van de Moesdijk D, Beverdam A, van Ginneken B,
398 de Korte CL. Comparison Study of Low-Cost Ultrasound Devices for Estimation of
399 Gestational Age in Resource-Limited Countries. *Ultrasound Med Biol* 2018; 44: 2250-
400 2260. DOI: 10.1016/j.ultrasmedbio.2018.05.023

401 Johnson GG, Zeiler FA, Unger B, Hansen G, Karakitsos D, Gillman LM. Estimating the
402 accuracy of optic nerve sheath diameter measurement using a pocket-sized, handheld
403 ultrasound on a simulation model. *Crit Ultrasound J* 2016; 8: 18. DOI: 10.1186/s13089-
404 016-0053-9

405 Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients
406 for Reliability Research. *J Chiropr Med* 2016; 15: 155-163. DOI:
407 10.1016/j.jcm.2016.02.012

408 Larsson M, Heyde B, Kremer F, Brodin LÅ, D'hooge J. Ultrasound speckle tracking for radial,
409 longitudinal and circumferential strain estimation of the carotid artery--an in vitro
410 validation via sonomicrometry using clinical and high-frequency ultrasound.
411 *Ultrasonics* 2015; 56: 399-408. DOI: 10.1016/j.ultras.2014.09.005

412 Lerman A, Zeiher AM. Endothelial function: cardiac events. *Circulation* 2005; 111: 363-368.
413 DOI: 10.1161/01.CIR.0000153339.27064.14

414 Lu M, Zhong W-H, Liu Y-X et al. Sample size for assessing agreement between two methods
415 of measurement by Bland-Altman method. *The international journal of biostatistics*
416 2016. DOI: 10.1515/ijb-2015-0039

417 Mantella LE, Colledanchise K, Bullen M, Héту MF, Day AG, McLellan CS, Johri AM.
418 Handheld versus conventional vascular ultrasound for assessing carotid artery plaque.
419 *Int J Cardiol* 2019; 278: 295-299. DOI: 10.1016/j.ijcard.2018.12.014

420 Mathiesen EB, Joakimsen O, Bonna KH. Intersonographer Reproducibility and Intermethod
421 Variability of Ultrasound Measurements of Carotid Artery Stenosis: The Tromso Study.
422 *Cerebrovascular Diseases* 2000; 10: 207-213.

423 van Mil ACCM, Hartman Y, van Oorschot F, Heemels A, Bax N, Dawson EA, Hopkins N,
424 Hopman MTE, Green DJ, Oxborough DL, Thijssen DHJ. Correlation of carotid artery
425 reactivity with cardiovascular risk factors and coronary artery vasodilator responses in
426 asymptomatic, healthy volunteers. *J Hypertens* 2017; 35: 1026-1034. DOI:
427 10.1097/HJH.0000000000001274

428 van Mil ACCM, Tymko MM, Kerstens TP, Stembridge M, Green DJ, Ainslie PN, Thijssen
429 DHJ. Similarity between carotid and coronary artery responses to sympathetic
430 stimulation and the role of alpha1-receptors in humans. *J Appl Physiol (1985)* 2018;
431 125: 409-418. DOI: 10.1152/jappphysiol.00386.2017

432 van Mil ACCM, Pouwels S, Wilbrink J, Warlé MC, Thijssen DHJ. Carotid Artery Reactivity
433 Predicts Events in Peripheral Arterial Disease Patients. *Ann Surg* 2019; 269: 767-773.
434 DOI: 10.1097/SLA.0000000000002558

435 Nabel EG, Ganz P, Gordon JB, Alexander RW, Selwyn AP. Dilation of normal and constriction
436 of atherosclerotic coronary arteries caused by the cold pressor test. *Circulation* 1988;
437 77: 43-52.

438 Niu L, Bao L, Zhu L, Tan Y, Xu X, Shan Y, Liu J, Zhu Q, Jiang C, Shen Y. Diagnostic
439 Performance of Automated Breast Ultrasound in Differentiating Benign and Malignant
440 Breast Masses in Asymptomatic Women: A Comparison Study With Handheld
441 Ultrasound. *J Ultrasound Med* 2019; 38: 2871-2880. DOI: 10.1002/jum.14991

442 Peace A, Van Mil A, Jones H, Thijssen DHJ. Similarities and Differences Between Carotid
443 Artery and Coronary Artery Function. *Curr Cardiol Rev* 2018; 14: 254-263. DOI:
444 10.2174/1573403X14666180910125638

445 Podgorski M, Winnicka M, Polguy M, Grzelak P, Łukaszewski M, Stefańczyk L. Does the
446 internal jugular vein affect the elasticity of the common carotid artery? *Cardiovasc*
447 *Ultrasound* 2016; 14: 40. DOI: 10.1186/s12947-016-0084-1

448 Prekker ME, Scott NL, Hart D, Sprengle MD, Leatherman JW. Point-of-care ultrasound to
449 estimate central venous pressure: a comparison of three techniques. *Crit Care Med* 2013;
450 41: 833-841. DOI: 10.1097/CCM.0b013e31827466b7

451 Steinbuch J, Hoeks AP, Hermeling E, Truijman MT, Schreuder FH, Mess WH. Standard B-
452 Mode Ultrasound Measures Local Carotid Artery Characteristics as Reliably as
453 Radiofrequency Phase Tracking in Symptomatic Carotid Artery Patients. *Ultrasound*
454 *Med Biol.* 2016 Feb;42(2):586-95. doi: 10.1016/j.ultrasmedbio.2015.07.030

455 Stock KF, Klein B, Steubl D, Lersch C, Heemann U, Wagenpfeil S, Eyer F, Clevert DA.
456 Comparison of a pocket-size ultrasound device with a premium ultrasound machine:
457 diagnostic value and time required in bedside ultrasound examination. *Abdom Imaging*
458 2015; 40: 2861-2866. DOI: 10.1007/s00261-015-0406-z

459 Thijssen DH, Dawson EA, Tinken TM, Cable NT, Green DJ. Retrograde flow and shear rate
460 acutely impair endothelial function in humans. *Hypertension* 2009; 53: 986-992. DOI:
461 10.1161/HYPERTENSIONAHA.109.131508

462 Triboulet J, Nasr E, Poignet P, Dombre E, Dautzat M. Evaluation of the influence of Probe
463 Pressure on the B-mode Ultrasound Measurement of Arterial Diameter. *Conf Proc IEEE*
464 *Eng Med Biol Soc* 2006.

465 Vidakovic R, Feringa HH, Kuiper RJ, Karagiannis SE, Schouten O, Dunkelgrun M, Hoeks SE,
466 Bom N, Bax JJ, Neskovic AN, Poldermans D. Comparison with computed tomography
467 of two ultrasound devices for diagnosis of abdominal aortic aneurysm. *Am J Cardiol*
468 2007; 100: 1786-1791. DOI: 10.1016/j.amjcard.2007.07.032

469 Woodman RJ, Playford DA, Watts GF, Reed C, Taylor RR, Puddey IB, Beilin LJ, Burke V,
470 Mori TA, Green D. Improved analysis of brachial artery ultrasound using a novel edge-
471 detection software system. *J Appl Physiol* 2001; 91: 929-937.

472 Zell K, Sperl JJ, Vogel MW, Niessner R, Haisch C. Acoustical properties of selected tissue
473 phantom materials for ultrasound imaging. *Phys Med Biol* 2007; 52: N475-484. DOI:
474 10.1088/0031-9155/52/20/N02

475 Zieleskiewicz L, Lopez A, Hraiech S, Baumstarck K, Pastene B, Di Bisceglie M, Coiffard B,
476 Duclos G, Boussuges A, Bobbia X, Einav S, Papazian L, Leone M. Bedside POCUS
477 during ward emergencies is associated with improved diagnosis and outcome: an
478 observational, prospective, controlled study. *Crit Care* 2021; 25: 34. DOI:
479 10.1186/s13054-021-03466-z