The reliability of the ankle brachial index: a systematic review

Sarah Casey1*, Sean Lanting1, Christopher Oldmeadow2 and Vivienne Chuter1

Abstract

Background: The ankle brachial index (ABI) is widely used in clinical practice as a non-invasive method to detect the presence and severity of peripheral arterial disease (PAD). Current guidelines suggest that it should be used to monitor potential progression of PAD in affected individuals. As such, it is important that the test is reliable when used for repeated measurements, by the same or different health practitioners. This systematic review aims to examine the literature to evaluate the inter- and intra-rater reliability of the ABI.

Methods: A systematic search of MEDLINE, EMBASE and CINAHL Complete was conducted to 20 January 2019. Two authors independently reviewed and selected relevant studies and extracted the data. Methodological quality was determined using the Quality Appraisal of Reliability (QAREL) Checklist.

Results: Fifteen studies of ABI reliability in a range of patient populations were identified as suitable for inclusion in the review: seven considered inter-rater reliability, four intra-rater reliability, and four studies evaluated both inter- and intra-rater reliability. Inter-rater reliability was found to be highly variable, with intraclass correlation coefficients (ICCs) ranging from poor to excellent (ICC 0.42–1.00), while intra-rater also demonstrated considerable variation, with ICCs from 0.42–0.98. Meta-analysis was not possible due to the lack of statistical information reported.

Conclusions: Results of included studies suggest the inter- and intra-tester reliability of the ABI is acceptable. However, inconsistencies in obtaining systolic pressure measurements, calculating ABI values, and incomplete reporting of methodologies and statistical analysis make it difficult to determine the validity of the results of included studies. Further research, with more consistent reliability methodology, statistical analysis and reporting conducted in populations at risk of PAD is needed to conclusively determine the ABI reliability.

Keywords: Ankle brachial index, Peripheral arterial disease, Lower extremity artery disease, Repeatability, Reproducibility, Reliability

Introduction/background

Peripheral arterial disease (PAD) describes the process of progressive atherosclerosis affecting arteries, most frequently in the lower limb. The prevalence in the general population has been estimated at up to 19% in people over the age of 55 years [1], with incidence increasing with advancing age and in the presence of smoking, inactivity and obesity [1, 2]. The presence of PAD is associated with increased risk of mortality and morbidity from cardiac atherosclerosis [2], and, in its advanced stages, can result in lower extremity ulceration and amputation [3]. Diabetes mellitus is an independent risk factor for the development of PAD [4], and in people with diabetes, atherosclerotic plaques tend to have a more distal and diffuse distribution and there is a more aggressive disease presentation [5].

Due to the high risk of concurrent cardiovascular morbidity, mortality and lower limb complications associated with PAD, accurate and reliable diagnostic testing methods are required for screening and ongoing monitoring [2, 6]. Early detection of PAD allows for intervention and management to reduce the risk of mortality and morbidity related to atherosclerosis (lifestyle modification, pharmacotherapy, e.g. statins, antiplatelets, and measures to address systemic risk factors such as hypertension or diabetes) [7]. Current recommendations for non-invasive lower limb vascular assessment include using the ankle-brachial index (ABI) as...
an objective measurement of peripheral blood flow [7, 8]. The ABI represents the ratio of ankle to brachial systolic pressure and is recommended to be calculated by dividing the higher systolic pressure of the dorsalis pedis and tibialis posterior vessels at the ankle with the higher of the systolic pressures measured in the brachial artery in both arms [7, 8].

The ABI is widely used to screen for PAD in different clinical settings and by different health professionals, from general medical practitioners to specialist vascular technicians [9, 10]. Reliability of the test for accurate ongoing monitoring of lower limb vascular status has the potential to be affected by a number of factors. As an operator-dependent test, this includes the experience and skills of the clinician, particularly as multiple clinicians are frequently involved in ongoing monitoring measurements [11, 12]. There are also a number of types of equipment (e.g., automated versus manual) and methods used to measure ankle and arm blood pressures (e.g., stethoscope, Doppler, photoplethysmography probe), with variable findings as to whether the results are interchangeable [13–16]. The pre-test protocol and test environment have also been demonstrated to affect the resting ABI at measurement, with variations in body position [17], recency of tobacco smoking, caffeine intake [18, 19] and exercise [20, 21], and pre-measurement rest time [22] all likely to introduce error to the measurement and affect the test-retest reliability.

Objectives
Given that the ABI is the recommended method for screening for the presence and progression of PAD, it is important that it is reliable. Therefore, the aim of this review was to systematically evaluate the literature to determine the inter- and intra-rater reliability of the ABI in adults.

Methods
Search strategy
A search of relevant biomedical journal databases from the University of Newcastle library website was performed to identify studies that consider the reliability of ABI measurement from database inception to January 2019 using MEDLINE (1946+), EMBASE (1947+), and CINAHL Complete. Truncated versions of some search terms were used to ensure that relevant studies were included (Table 1).

| 1 | Ankle brachial pressure |
|---|-------------------------|
| 2 | Ankle arm pressure       |
| 3 | Ankle brachial ind*      |
| 4 | Reliab*                 |
| 5 | Consistenc*              |
| 6 | Accura*                  |
| 7 | Reproduc*                |
| 8 | Repeat*                  |
| 9 | Agreement                |
| 10| Precision                |
| 11| 1 or 2 or 3 AND 4 or 5 or 6 or 7 or 8 or 9 or 10 |

Other sources
Hand searching of the reference list of appropriate articles was also conducted.

Data collection and analysis
All abstracts obtained were assessed independently by SC and SL for inclusion. There were no instances of disagreement between reviewers, so arbitration by a third person (VC) was not necessary. Data extraction was performed by SC and SL. It was pre-determined that a meta-analysis of reliability outcomes for inter- and intra-rater reliability would be conducted provided there were sufficient studies that report the estimator of interest, and a measure of uncertainty for this estimator (e.g., standard error, 95% confidence interval, non-truncated p-value). Given the expectation for a high degree of study heterogeneity, we believed a fixed effect meta-analysis would generally not be appropriate so we aimed to only pool estimates using a random effects approach provided there were at least 5 studies [24].

Methodological quality assessment
The studies that met the inclusion criteria were appraised for risk of bias using the Quality Appraisal of Reliability (QAREL) Checklist and qualitative methodological assessment [25]. All full-text papers were assessed for methodological quality independently by two reviewers (SC/SL), and as there were no disagreements arbitration by a third reviewer (VC) was not necessary.

Results
A total of 1703 articles were retrieved, of which 36 were identified as suitable for full-text review. Twenty-one papers were excluded based on the exclusion criteria (Fig. 1): 10 papers reported comparison of methods [26–35], five
studies did not report measures of reliability [36–40], two studies compared raters’ experience [41, 42] and one reported a novel trial design, for which the reliability results were duplicated in another included paper [43]. One paper used measures repeated at up to 365 days apart, with a mean time between measures of 228 days, which is long enough to encompass changes attributable to progression of PAD [44]. Two papers were conference abstracts, for which the full text could not be obtained as the authors did not respond to a request for further information [40, 45].

Of the included papers, seven measured inter-tester reliability [12, 46–51], four assessed intra-tester reliability [52–55], and four considered both inter- and intra-tester reliability [13, 16, 56, 57].

**Characteristics and overview of included studies**

The 15 studies in this review included a total of 916 participants, with data collected from a combination of one and both lower limbs (1396 limbs in total). Two studies did not state the number of limbs included [52, 53]. Eleven studies assessed inter-rater reliability [12, 13, 16, 46–50, 56, 57], and eight studies reported intra-rater reliability [13, 16, 52–57]. The characteristics of included studies are described in Table 2. Eleven studies reported participants’ gender, with more men (n = 416, 56.4%) overall than women, whilst gender was unreported in four studies [12, 46, 49, 50]. Most of the studies included predominantly older participants (age range 41–92 years) [12, 13, 16, 47–49, 51, 53–55, 57], however two studies recruited only younger adults (age range 22–30 years) [46, 56], one study included 18–80 year olds [52] and one study did not report participants’ ages [50]. The majority of studies [12, 47–51, 55, 57] included only participants with suspected PAD, or risk factors for atherosclerosis; three studied a mixed population including those without risk factors or clinical indicators of PAD.
Table 2: Participant characteristics and reliability measure

| Reference                  | Number (n) | Gender (M:F) | Age (years) | Height (cm) | Weight (kg) | DM & control | DM duration | Medical history | PAD | Reliability measure |
|----------------------------|------------|--------------|-------------|-------------|-------------|--------------|-------------|------------------|-----|--------------------|
| Aboyans et al. (2008) [13] | 54         | 28, 26       | 52.8 ± 17.1 | NR          | NR          | DM 35.2%     | control NR  | NR               | HT 38.9%, DL 38.9%, CAD 33.3% CVD 11.1% TS 22.2% | ICC Doppler: Inter: 0.79 (0.70–0.85) |
| Alvaro-Afonso et al. (2018) [51] | 21         | 15, 6        | 67 ± 8.7    | NR          | NR          | All DM2: control NR | NR           | HT 19, DL 17, NEU 18, NEP 1, TS 6 | NR | Kappa coefficient Normal: 0.4 (p < 0.001), PAD 0.7 (p < 0.001), MAC 0.43 (p < 0.001) |
| Chesbro et al. (2011) [46]  | 20         | NR           | 22–30       | NR          | NR          | NR           | NR          | None             | ‘Healthy young adults’ | ICC Doppler: R 1.00 (95%CI 0.999–1.00, p < 0.001), L 0.99 (0.997–1.00, p < 0.001) |
| Chesbro et al. (2013) [56]  | 10         | 5, 5         | NR          | NR          | NR          | ‘Healthy volunteers only’ | None        | None             | None              | ICC Vascular Cuff Right: Trial 1: 0.853; Trial 2: 0.898 Left: Trial 1: 0.444; Trial 2: 0.938 Standard Cuff Right: Trial 1: 0.902; Trial 2: 0.817 Left: Trial 1: 0.826; Trial 2: 0.867 |
| de Graff et al. (2001) [57] | 54         | 31, 23       | 66 ± 12     | NR          | NR          | DM 36%       | control NR  | HT 43%, DI 35%, CAD 36%, CVD 20% | All PAD | Inter: 20 / 0.92 |
| Georgakarakos et al. (2013) [47] | 18         | 12, 6        | 54–74       | NR          | NR          | 9 DM: 3 oral, 6 insulin | NR          | HT 15, DL 9, TS 4 | All PAD | Inter: week 27 / 0.87 |
| Holland-Letz et al. (2007) [16] | 108        | 50, 58       | 68.1 ± 1.5  | NR          | BMI 29 ± 43 | DM 15.7%     | control NR  | HT 58.1% DL 54.8%, TS 9.2% current, 40.7% ex-, Hx vasc surg 6 subjects | 68.1 ± 1.5 | ICC for inter-observer: 0.423 |
| Jaffer et al. (2008) [50]   | 25         | NR           | NR          | NR          | NR          | NR           | NR          | NR               | All suspected PAD | Pearson CC r = 0.516 (p < 0.001) |
| Langen et al. (2009) [48]   | 20         | 11, 9        | 41–75       | NR          | NR          | NR           | NR          | NR               | All IC Symptoms | Interobserver variability: 10% (SD 0.8) |
| Mätzke et al. (2003) [12]   | 30         | 18, 15 (no of limbs) | 26 > 65 yrs, 7 ≤ 65 yrs | NR          | NR          | 8 DM, control NR | NR          | NR               | Ischaemic pressure lesion or rest pain | Coefficient of variation: 3.2 |
| Span et al. (2016) [49]     | 136        | NR           | 64 ± 78     | NR          | NR          | 19 (14%) control not reported | NR          | HT66, DL 58, TS 22 current, 39 ex- | RF or IC | Coefficient of variation Doppler: 5.9% R & L legs Auto: Right 3.2% Left 3.5% |

Intra-rater reliability (n = 8)
| Reference            | Number (n) | Gender (M:F) | Age (years) | Height (cm) | Weight (kg) | DM & control | DM duration | Medical history | PAD | Reliability measure |
|----------------------|------------|--------------|-------------|-------------|-------------|--------------|-------------|-----------------|-----|---------------------|
| Aboyans et al.       | 54         | 28:26        | 52.8 ± 17.1 | NR          | NR          | DM 35.2%    | NR          | HT 38.9%, DL 38.9%, CAD 33.3%, CVD 11.1%, TS 22.2% | 19 IC, 25 RF, 10 healthy | ICC Doppler: Intra: 0.89 (0.84–0.92), Pulse: Intra: 0.60 (0.44–0.73), ICC- Intra-rater Vascular Cuff Rater 1 R: 0.750; Rater 1 R: 0.696 Rater 2 R: 0.551; Rater 2 L: 0.869 Standard Cuff Rater 1 R: 0.628; Rater 1 L: 0.420 Rater 2 R: 0.620, Rater 2 L:0.585 |
| Chesbro et al.       | 10         | 5:5          | NR          | NR          | NR          | NR          | NR          | ‘Healthy young adults’                           | None            | Repeatability Coefficient/ICC Intra: day 9 / 0.08 Intra: week 22 / 0.89 |
| de Graff et al.      | 54         | 31:23        | 66 ± 12     | NR          | NR          | DM 36%      | NR          | HT 43%, DL 35%, CAD 36%, CVD 20%                  | Suspected       | ICC Single measurement 0.808 Mean: 0.927 |
| Demir et al.         | 161        | 87:74        | 52.03 ± 18.99 | 165.12 ± 8.88 | 75.61 ± 134 | DM NR       | NR          | HT 62.7%, DL 46.6%, TS 29.8%                     | Mixed population | |
| Faccenda et al.      | 36         | 28:8         | 56 ± 11     | NR          | NR          | All DM1     | NR          | No other hx reported                              | NR              | Coefficient of variation 8% |
| Holland-Letz et al.  | 108        | 50:58        | 681 ± 1.5   | BMI 29 ± 43 | DM 15.7%    | control NR | NR          | HT 58.1%, DL 54.8%, TS 9.2% current, 40.7% ex-, Hx vasc surg 6 subjects | 68.1 ± 1.5       | Intra-observer Variance: 0.008, SD 0.87 [0.081; 0.095] |
| Millen et al.        | 66         | 51:15        | 695 ± 12 yrs. (range 35–92) | NR          | NR          | 4 DM1, 14 DM2 | HT 79%, DL 68%, CAD 44%, CVD 17%, TS 15% current, 59% ex- | 36 IC, 4 rest pain | Coefficient of variation: Dopplex Ability: 9.65 ± 12% Parks Flo-lab: 4.95 ± 3% |
| Rosenbaum et al.     | 157        | 80:77        | 591 ± 13.2  | NR          | NR          | 35 DM, control NR | HT88, DL 103, CAD 14, CVD 2, TS 27 current, 49 ex- | 11 PAD, all RF or IC | Coefficient of variation: Dopplex Ability: 9.65 ± 12% Parks Flo-lab: 4.95 ± 3% |

NR Not reported, DM Diabetes, HT Hypertension, DL Dyslipidaemia, CAD Coronary artery disease, CVD Cerebrovascular disease, NEU Neuropathy, NEP Nephropathy, TS smoking, IC Intermittent claudication, RF risk factors, PAD Peripheral arterial disease
Chesbro et al., [46] provided no details on the intra-rater inclusion in a meta-analysis. For example, the paper by Aboyans et al., [13, 56, 57], with ICC values ranging from poor (ICC: 0.42) to excellent (ICC: 1.0) [46]. Other estimates of reliability reported in included studies were coefficient of variation between raters [12, 49] (ranging from 3.2 to 5.9%), inter-observer reliability of 10% for raters [48], and a moderate Pearson’s correlation coefficient of 0.52 in a population with suspected PAD [50].

Of the remaining studies, one demonstrated statistically significant differences in ABI between raters in a population with severe PAD and in those with no disease, which did not occur in those participants with mild to moderate PAD [47], suggesting increased reliability with this disease state. In contrast, another paper reported Kappa coefficients of 0.4 (low agreement) for healthy limbs, 0.7 (good agreement) for limbs with PAD, and 0.43 (moderate agreement) for limbs with medial arterial calcification (MAC) (< 0.001 for all values) [51].

Intra-rater reliability

Intra-rater reliability results are included in Table 2. Various methods of calculating reliability were used. Of the eight included studies, four reported ICCs [13, 52, 56, 57], with ICC values ranging from poor (ICC: 0.42) to excellent (0.98) [57]. Interpretation of the results was limited again by the fact that not all studies reported 95% confidence intervals, with only two articles having done so. [13, 56]. Other estimates of reliability included coefficient of variation [53–55] (range 4.95% [54] – 15.8% [55]), and an intra-observer variance of 8% [16].

Discussion

The findings of this review are that the inter- and intra-tester reliability of the ABI across a number of mixed
| Item                                                                 | Aboyans et al. (2008) [13] | Alvaro-Afonso et al. (2018) [51] | Chesbro et al. (2011) [46] | Chesbro et al. (2013) [56] | de Graaff et al. (2001) [57] | Demir et al. (2016) [52] | Faccenda et al. (1988) [47] | Georgakarakos et al. (2013) | Holland-Letz et al. (2007) [16] | Jaffer et al. (2008) [50] | Langen et al. (2009) [48] | Matzke et al. (2003) [12] | Millen et al. (2018) [54] | Rosenbaum et al. (2012) | Span et al. (2016) [49] |
|----------------------------------------------------------------------|-----------------------------|-----------------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-------------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|
| 1. Was the test evaluated in a sample of subjects who were representative of those to whom the authors intended the results to be applied? | Y                           | Y                                 | Y                           | Y                           | Y                           | Y                           | Y                           | Y                           | Y                            | Y                           | Y                           | Y                           | Y                           | Y                           | Y                           |
| 2. Was the test performed by raters who were representative of those to whom the authors intended the results to be applied? | Y                           | U                                 | Y                           | Y                           | Y                           | U                           | Y                           | Y                           | Y                            | Y                           | Y                           | Y                           | Y                           | Y                           | Y                           |
| 3. Were raters blinded to the findings of other raters during the study? | Y                           | U                                 | Y                           | Y                           | NA                          | NA                          | NA                          | Y                           | U                            | U                           | Y                           | Y                           | U                           | Y                           | U                           |
| 4. Were raters blinded to their own prior findings of the test under evaluation? | Y                           | NA                                | U                           | U                           | U                           | U                           | U                           | Y                           | NA                           | NA                          | NA                          | NA                          | NA                          | NA                          | U                           |
| 5. Were raters blinded to the results of the reference standard for the target disorder (or variable) being evaluated? | NA                          | NA                                | NA                          | NA                          | NA                          | NA                          | NA                          | NA                          | NA                           | NA                          | NA                          | NA                          | NA                          | NA                          | NA                           |
| 6. Were raters blinded to clinical information that was not intended to be provided as part of the testing procedure or study design? | U                           | U                                 | U                           | U                           | U                           | U                           | U                           | U                           | U                            | U                           | U                           | U                           | U                           | U                           | U                           |
| 7. Were raters blinded to additional cues that were not part of the test? | U                           | U                                 | U                           | U                           | U                           | U                           | U                           | U                           | U                            | U                           | U                           | U                           | U                           | U                           | U                           |
| 8. Was the order of examination varied? | U                           | U                                 | Y                           | Y                           | U                           | N                           | U                           | U                           | N                            | U                           | U                           | Y                           | U                           | U                           | U                           |
| 9. Was the time interval between repeated measurements compatible with the stability (or theoretical stability) of the variable being measured? | U                           | Y                                 | Y                           | Y                           | Y                           | Y                           | Y                           | Y                           | Y                            | U                           | Y                           | U                           | Y                           | U                           | Y                           |
| 10. Was the test applied correctly and interpreted appropriately? | N                           | Y                                 | Y                           | Y                           | Y                           | N                           | Y                           | N                           | U                            | Y                           | Y                           | Y                           | N                           | Y                           | Y                           |
| 11. Were appropriate statistical measures of agreement used? | Y                           | Y                                 | Y                           | Y                           | Y                           | P                           | Y                           | Y                           | Y                            | U                           | N                           | Y                           | Y                           | Y                           | Y                           |

Y: Yes, N: No, U: Unclear, P: Partly, NA: Not applicable
populations appears to be acceptable, however statistical tests of reliability in included papers were heterogeneous and levels of statistical reporting were inconsistent and incomplete. This makes interpretation of the reliability of the ABI in the context of clinical detection, evaluation and ongoing monitoring of peripheral arterial supply challenging, and prevented meta-analysis. For example, where studies lack 95% confidence intervals for ICCs, the validity of interpretation of the value is reduced as it fails to provide the lowest level of reliability that it represents. Similarly for coefficient or estimate of variation, values between 3.2 and 15.8% were reported. Whilst this is considered an acceptable level of variation for many clinical tests, for the ABI it can represent a range of values that may indicate both normal and pathological results; which could reduce the ability of ABI to reliably determine the presence and extent of PAD. For example, assuming a variation of 15%, an ABI of 1.0 (which is considered ‘borderline’ when ABI is used as a screening tool [6]) could represent a true value between 0.85 (indicative of PAD) and 1.15 (‘normal’).

Further complicating the interpretation and generalisability of the inter- and intra-rater reliability results of included studies was the heterogeneity of participant populations. Whilst the majority of studies included older people with PAD risk factors or suspected PAD, three studies also included healthy participants [13, 16, 52], and two used an exclusively young and healthy population [46, 56]. In clinical practice, ABI is used to evaluate peripheral arterial supply in people with risk factors for atherosclerosis, and in those with clinical signs and symptoms of PAD. The variation in the disease status of participants across the studies included in this review provides some difficulty in evaluating how the studies’ findings apply to the people in whom the ABI would clinically be used. The study that reported near-perfect inter- and intra-tester reliability included only healthy individuals under the age of 30 [56]. This population would not typically undergo vascular screening, and the results obtained do not indicate the ability of the ABI to perform reliably in the presence of pathology where the result is likely to be lower and therefore change in result indicative of worsening pathology is likely to be small. In contrast, inter-tester and intra-tester reliability was found to be poor in several populations in which this test is recommended including people with diabetes and without MAC, [51] and older people with risk factors for PAD [16].

Methodological differences between studies is also likely to have contributed to variable reliability outcomes, with automated oscillometric devices demonstrating marginally better reliability than manual assessment using Doppler [49, 55], while Doppler evaluation was found to be more reliable than the use of pulse palpation [13] or stethoscope [46]. Higher ABI reliability was found in more experienced raters [47]. Whilst most of the studies reported that participants rested for 5–15 min prior to testing [12, 13, 16, 46, 48, 51, 55–57], six studies did not describe any pre-test preparation [47, 49, 50, 52–54], and only one paper took steps to ensure that participants did not consume alcohol, caffeine or tobacco (which are known to affect blood pressure) in the two hours prior to testing which may have affected measurements, particularly when taken across two different testing sessions [55]. This lack of reporting of the methodology used to obtain systolic blood pressure measurements makes it difficult to compare results across the included studies as it is unknown how much external factors are likely to contribute measurement variability.

Two papers identified the presence of diabetes mellitus as a factor that may affect reliability of the ABI [12, 51], however only one study included a large enough sample of this cohort to perform statistical tests [51]. This study, which used only participants with diabetes, reported the Kappa coefficient for inter-tester measures for participants classed as having PAD or not, rather than performing ICCs on the measures obtained. The authors reported ‘good’ reproducibility of the ABI (K 0.7) in people classified by their ABI measurement as having PAD, but low reproducibility in those without PAD and in those with MAC. Previous research has also shown that people with diabetes demonstrate a different response to pre-measurement rest [22], and that brachial blood pressure measurement is also less reliable in these individuals [58]. Diabetes-related autonomic neuropathy has been shown to affect blood pressure regulation, with a lack of vasoconstriction arising from reduced sympathetic input, particularly in response to changes in temperature and position [59, 60].

Limitations
While the search methods employed in this study were designed to be robust, there may be some evidence that was not captured, for example unpublished data. Further limitations to this study are the inability to perform meta-analysis in order to obtain a quantitative analysis of the available reliability data for the ABI, and the inability perform any sub-analyses relating to individual populations such as those with diabetes, or methods of measurement such as automated or manual methods. Furthermore, there has been some disagreement in the literature about which pressure measurement should be used to calculate the ABI [61, 62], with no studies exploring the effect of calculation method on reliability. However, the method of calculation cannot be excluded as a factor affecting reliability that has not been considered by this review.

Conclusion
Results of included studies suggest the inter- and intra-tester reliability of the ABI is acceptable. However, inconsistencies in obtaining systolic pressure measurements,
calculating ABI values, and incomplete reporting of methodologies and statistical analysis make it difficult to determine the validity of the results of included studies. Further research of ABI reliability using a more consistent approach to study design and implementation and more detailed reporting of results in populations with vascular pathology and at risk of PAD is required. Based on current available data clinicians should ensure they interpret ABI results in the context of other vascular assessment findings, and patient management is not based upon this measurement alone.

Abbreviations
ABI: ankle brachial index; ICC: intraclass correlation coefficient; MAC: medial arterial calcification; PAD: peripheral arterial disease; QAREL: Quality Appraisal of Reliability

Acknowledgements
Not applicable.

Availability of data and material
All data generated or analysed during this study are available in this article.

Authors’ contributions
VC and SC conceived the study. Database searches and data extraction was conducted by SC and SL. Statistical evaluation was performed by CO. All authors contributed to the manuscript and approved the final manuscript.

Funding
Funding sources include the national Research Training Scheme (Australia).

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
1PO Box 127, Ourimbah, NSW 2258, Australia. 2CReDITSS – HMRI, School of Medicine and Public Health, Newcastle, Australia.

Received: 30 May 2019 Accepted: 29 July 2019 Published online: 02 August 2019

References
1. Meijer WT, Hoes AW, Rutgers D, Bots ML, Hofman A, Grobbee DE. Peripheral arterial disease in the elderly. Rotterdam Stud. 1998;18(2):185–92.
2. Norman PE, Eikelboom JW, Hanky GJ. Peripheral arterial disease: prognostic significance and prevention of atherothrombotic complications. Med J Aust. 2004;181(3):150–4.
3. Bonham PA. Get the LEAD out: noninvasive assessment for lower extremity arterial disease using ankle brachial index and toe brachial index measurements. J Wound Ostomy and Contience Nurs. 2006;33(1):30–41.
4. American Diabetes Association. Peripheral arterial disease in people with diabetes. Diabetes Care. 2003;26(12):3333–41.
5. Jude EB, Oyibo SO, Chalmers N, Boulton AJM. Peripheral arterial disease in diabetic and nondiabetic patients: a comparison of severity and outcome. Diabetes Care. 2001;24(8):1433–7.
6. Aboyans V, Criqui MH, Abraham P, Allison MA, Creager MA, Diehm C, et al. Measurement and interpretation of the ankle-brachial index. A Scientific Statement From the American Heart Association. Circulation. 2012;126(24):2890–909.
7. Gerhard-Herman MD, Gornik HL, Barrett C, Banches NR, Corriere MA, Drachman DE, et al. 2016 AHA/ACC guideline on the management of patients with lower extremity peripheral artery disease: executive summary.
8. A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. 2016.
9. Norgren L, Hiatt WR, Dormandy JA, Nehler MR, Harris KA, Fowkes FG. International consensus for the Management of Peripheral Arterial Disease (TASC II). J Vasc Surg. 2005;41(3):1395–458.
10. Haigh K, Bingley J, Goldedge J, Walker P. Peripheral arterial disease - screening in general practice. Aust Fam Physician. 2013;42:391–5.
11. Al-Qaissi M, Nott DM, King DR, Kaddouss S. Ankle brachial pressure index (ABPI): an update for practitioners. Vasc Health Risk Manag. 2009;5:833–41.
12. Nicolai SP, Krudenier LM, Rouwet EV, Bartelink ML, Prins MH, Teijink JA. Ankle brachial index measurement in primary care: are we doing it right? Br J Gen Pract. 2009;59(563):422–7.
13. Matze S, Frankena M, Alback A, Riallo M, Lepantalo M. Ankle brachial index measurement in critical leg ischaemia - the influence of reproducibility. Scand J Surg. 2003;92:144–7.
14. Aboyans V, Lacroix P, Doucet S, Preus PM, Cirqiu MH, Laskar M. Diagnosis of peripheral arterial disease in general practice: can the ankle-brachial index be measured either by pulse palpation or an automatic blood pressure device? Int J Clin Pract. 2003;57(1):29–33.
15. Stoesen JA, Li Y, Wang J-G. Comparison of three measures of the ankle-brachial blood pressure index in a general population. Hypertens Res. 2007;30:555.
16. Ramanathan A, Conaghan PJ, Jenkinson AD, Bishop CR. Comparison of ankle-brachial pressure index measurements using an automated oscillimetric device with the standard doppler ultrasound technique. ANZ J Surg. 2003;73(3):105–8.
17. Holland-Letz T, Endres HG, Biedermann S, Mahn M, Kunert J, Goh S, et al. Reproducibility and reliability of the ankle—brachial index as assessed by vascular experts, family physicians and nurses. Vasc Med. 2007;12(2):105–12.
18. Pollak EW, Chavis P, Wolfman EF. The effect of postural changes upon the ankle arterial perfusion pressure. Vasc Surg. 1976;10(4):219–22.
19. Pilli R, Naidu M, Pinguli U, Takakkapally RK. Study of cardiovascular effects of caffeine in healthy human subjects, with special reference to pulse wave velocity using photoplethysmography. Int J Nutr Pharmacol Neurol Dis. 2012;2(3):924–30.
20. Waring WS, Goudsmit J, Manwijk J, Webb DJ, Maxwell SR. Acute caffeine intake influences central more than peripheral blood pressure in young adults. Am J Hypertens. 2003;16(11 Pt 1):919–24.
21. Yaraco AR, Gardner AW. Acute reduction in ankle-brachial index following smoking in chronic smokers with peripheral arterial occlusive disease. Angiology. 1999;50(5):355–60.
22. Carter SA. Response of ankle systolic pressure to leg exercise in mild or questionable arterial disease. N Engl J Med. 1972;287(12):578–82.
23. Chuter VH, Casey SL. Effect of premeasurement rest time on systolic ankle pressure. J Am Heart Assoc. 2013;2(4):e000203.
24. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Grissche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS Med. 2009;6(6):e1000097.
25. Borenstein M, Hedges LV, Higgins JP, Rothstein HR. A basic introduction to fixed-effect and random-effects models for meta-analysis. Res Synth Methods. 2010;1(2):197–111.
26. Lucas NP, Macaskill P, Irwig L, Bogduk N. The development of a quality appraisal tool for studies of diagnostic reliability. J Clin Epidemiol. 2001;54:854–61.
27. Jvd S, Verbogt NP, Mulder PG, Steunenberg SL, Steunenberg BE, Lvd L. The clinical applicability of an automated photoplethysmographic determination of the ankle-brachial index after vascular surgery. Vascular. 2016;24:5:545–8.
28. Harrison ML, Lin H-F, Blakely DW, Tanaka H. Preliminary assessment of an automatic screening device for peripheral arterial disease using ankle–brachial and toe–brachial indices. Blood Press Monit. 2011;16(3):138–41.
29. Khandanpour N, Armon MG, Jennings B, Clark A, Meyer FJ. Photoplethysmography, an easy and accurate method for measuring ankle brachial pressure index: can photoplethysmography replace Doppler? Vasc Endovasc Surg. 2005;49(6):578–82.
30. Bonham PA, Cappuccio M, Hulsey T, Michel Y, Kelechi T, Jenkins C, et al. Are ankle and toe brachial indices (ABI-TBI) obtained by a pocket Doppler interchangeable with those obtained by standard laboratory equipment? J Wound Ostomy Contience Nurs. 2007;34(1):135–44.
31. Raines JK, Farrar J, Noclely K, Pena J, Davis WW, Willems HJ, et al. Ankle/brachial index in the primary care setting. Vasc Endovasc Surg. 2004;38(2):131–6.
31. Gestin S, Delluc A, Salou AH, Colas A, Guiguen F, Gladu G, et al. Ankle brachial pressure index (ABPI): color-Doppler versus ultrasound Doppler correlation study in 98 patients after analysis of interobserver reproducibility. J Mal Vasc. 2012;37(4):186–94.

32. Kollias A, Xilomenos A, Protogerou A, Dimakakos E, Stergiou GS. Automated determination of the ankle-brachial index using an oscillometric blood pressure monitor: validation vs. Doppler measurement and cardiovascular risk factor profile. Hypertens Res. 2011;34(7):825–30.

33. Hamel J-F, Foucaud D, Fanello S. Comparison of the automated oscillometric method with the gold standard Doppler ultrasound method to access the ankle-brachial pressure index. Angiology. 2010;61(5):487–91.

34. Davies JH, Williams EM. Automated plethysmographic measurement of the ankle-brachial index: a comparison with the doppler ultrasound method. Hypertens Res. 2016;39(2):100–6.

35. Bundó M, Urea M, Muñoz-Oníz L, Pérez C, Llusà J, Forés R, et al. Measurement of the ankle brachial index with a non-mercury sphygmomanometer in diabetic patients: a concordance study. BMC Cardiovasc Disord. 2013;13(1):1–7.

36. Fowkes FG, Housley E, Macintyre CC, Prescott RJ, Ruckley CV. Variability of ankle and brachial systolic pressures in the measurement of atherosclerotic peripheral arterial disease. J Epidemiol Community Health. 1988;42(2):128–33.

37. Ena J, Lozano T, Verdú G, Argente CR, González VL. Accuracy of ankle-brachial index using the SCVL(®), an arm and ankle automated device with synchronized cuffs, in a population with increased cardiovascular risk. Vasc Health Risk Manag. 2012;8:239–46.

38. Chongthawonsatid S, Dutsadeevettakul S. Validity and reliability of the ankle-brachial index by oscillometric blood pressure and automated ankle-brachial index. J Res Med Sci. 2017;22:224.

39. Alzamora MT, Baena-Diez JM, Sorribes M, Forés R, Toran P, Vicheto M. Peripheral Arterial Disease study (PERART): prevalence and predictive values of asymptomatic peripheral arterial occlusive disease related to cardiovascular morbidity and mortality. BMC Publ Health. 2007;7. https://doi.org/10.1186/1471-2458-7-348.

40. Bakalakou K, Marinakos A, Nouli A, Taxiarchou E, Margariti C, Papanikitas K, et al. Efficacy of automatic blood pressure device to determine reliably the ankle brachial index (ABI). Eur J Intern Med. 2011;22:59.

41. Monti M, Calanca L, Alatari A, Mazolli L. Accuracy of in-patients ankle-brachial index measurement by medical students. Vasa. 2016;45(1):43–8.

42. Navee J, Damsbo B, Lund JO, Munck A. Measurement of blood pressure, ankle brachial pressure and calculation of ankle brachial index in general practice. Fam Pract. 2012;29(3):345–51.

43. Endres H, Hucke C, Holland-Letz T, Trampisch H-J. A new efficient trial design for assessing reliability of ankle-brachial index measures by three different observer groups. BMC Cardiovasc Disord. 2006;6(1):33.

44. Weatherley BD, Chambliss LE, Heiss G, Catellier DJ, Ellison CR. The reliability of the ankle-brachial index in the atherosclerosis risk in communities (ARC) study and the NHLBI family heart study (FHS). BMC Cardiovasc Disord. 2006;6(1):11.

45. Balkanoy OO, Arapi B, Arslan C, Gurel SA. Reliability of ankle brachial index measurement by oscillometric blood pressure monitoring device. Heart Surg Forum. 2012;15:566.

46. Chesbro SB, Asongwed ET, Brown J, John EB. Reliability of Doppler and stethoscope methods of determining systolic blood pressures: considerations for calculating an ankle-brachial index. J Natl Med Assoc. 2011;103(9):863–9.

47. Georgakarakos E, Papadaki E, Vamvakouropou V, Lytras D, Tsiokani A, Trolakaki O, et al. Training to measure ankle–brachial index at the undergraduate level: can it be successful? Int J Low Extrem Wounds. 2013;12(2):167–71.

48. Hy L, Jv G, Rubbens L. Interoobserver variability of ankle-brachial index measurements at rest and post exercise in patients with intermittent claudication. Vasc Med. 2009;14(3):221–6.

49. Špan M, Geršak G, Willassek SC, Moža M, Kolir A. Detection of peripheral arterial disease with an improved automated device: comparison of a new oscillometric device and the standard Doppler method. Vasc Health Risk Manag. 2016;12:305–11.

50. Jaffer U, Elmagrabi AW, Cameron A, Osman I. Agreement of community-performed ankle brachial pressure indices (ABPI) with vascular laboratory performed assessment. Vasc Dis Manag. 2008;5(2):271–2.

51. Alvaro-Afonso FJ, García-Morales E, Molines-Barroso RJ, García-Alvarez Y, Sanz-Corbalan I, Lazaro-Martinez J. Interobserver reliability of the ankle-brachial index, toe-brachial index and distal pulse palpation in patients with diabetes. Diab Vasc Dis Res. 2018;15(4):344–7.