Automated devices for identifying peripheral arterial disease in people with leg ulceration: an evidence synthesis and cost-effectiveness analysis

Dwayne Boyers,[1] Moira Cruickshank,[2] Lorna Aucott,[2] Charlotte Kennedy,[1] Paul Manson,[2] Paul Bachoo[3] and Miriam Brazzelli[1*]

[1]Health Economics Research Unit, University of Aberdeen, Aberdeen, UK
[2]Health Services Research Unit, University of Aberdeen, Aberdeen, UK
[3]NHS Grampian, Aberdeen, UK

[*]Corresponding author m.brazzelli@abdn.ac.uk

Published August 2024
DOI: 10.3310/TWCG3912

Scientific summary

Automated devices for identifying peripheral arterial disease in people with leg ulceration: an evidence synthesis and cost-effectiveness analysis

Health Technology Assessment 2024; Vol. 28: No. 37
DOI: 10.3310/TWCG3912

NIHR Journals Library www.journalslibrary.nihr.ac.uk

Scientific summary

Background

Peripheral artery disease (PAD) is a highly prevalent atherosclerotic condition characterised by the narrowing of the peripheral arteries resulting in restriction of blood supply to the affected limb. Although PAD is frequently asymptomatic, it can cause complications that can range from intermittent claudication (pain on walking which is relieved by rest) to critical limb ischaemia. Up to one-quarter of people with symptomatic PAD may require intervention, and amputation may be necessary if it is left untreated. Leg ulcers are defined as wounds that occur below the knee and either on or above the ankle (malleolus). Compression treatment (bandages or stockings) is recommended to treat venous leg ulcers, and there is a robust evidence base to support its effectiveness. However, compression therapy should be avoided in people with leg wounds and symptoms of arterial insufficiency, as compression may cause damage by impairing the arterial supply to the ulcerated leg. To improve PAD diagnosis and decide the most suitable treatment, people with leg ulcers are assessed using ankle–brachial pressure index (ABPI) measurements. ABPI is usually measured using a sphygmomanometer and manual Doppler device, which requires expertise from the relevant operator/healthcare professional. The procedure can be protracted and unpleasant for those with leg ulcers. Automated devices may be advantageous in reducing the length of time taken to assess ABPI and, thereby, any associated discomfort for the patient. In addition, automated devices may potentially be more accurate than manual processes in detecting PAD, thus conferring additional benefits such as reduced time to treatment and improved outcomes for people with leg ulcers.

Objectives

The specific objectives of this assessment were to:

- Determine the diagnostic performance and clinical utility of automated devices available in United Kingdom (UK) clinical practice [BlueDop Vascular Expert (BlueDop Medical), boso ABI-system 100 (BOSCH + SOHN), WatchBP Office ABI (Microlife), WatchBP Office Vascular (Microlife)], MESI ABPI MD (MESI), MESI mTABLET ABI (MESI), Dopplex Ability Automatic ABI System (Huntleigh Healthcare) for assessing the presence of PAD in people with leg ulcers.
- Develop an economic model to assess the cost-effectiveness of the automated devices available in UK clinical practice for assessing the presence of PAD in people with leg ulcers.

Methods

Clinical effectiveness

Comprehensive electronic searches of databases including MEDLINE, EMBASE, Cochrane Library Web of Science and CINAHL were conducted to identify relevant reports of published studies. Evidence was considered from studies of any design assessing the relevant automated devices versus standard clinical assessment using a manual Doppler device. Initially, the population of interest was people with leg ulcers requiring measurement of ABPI, but, due to the dearth of available evidence, it was broadened to any population receiving ABPI measurement. Data on the diagnostic performance of the automated devices including data on the level of agreement between ABPI readings from automated devices and those from the reference device were extracted from the included studies. Information on the use of the devices in clinical practice was also recorded. Risk of bias was assessed using the Quality Assessment of Diagnostic Accuracy Studies – version 2 (QUADAS-2), QUADAS-C and the Review Body for Interventional Procedures (ReBIP) checklists, according to the type of study design. For each device,

when sufficient data were available, we conducted random-effects meta-analyses using a Hierarchical Summary Receiving Operating Characteristic (HSROC) model.

A two-stage, de novo decision analysis model was developed to assess cost-effectiveness. The first part was a decision tree model, which used a linked-evidence approach to capture the impact of test diagnostic accuracy on expected costs and quality-adjusted life-years (QALYs) for the first 24 weeks following test use. This included delayed venous ulcer healing due to false-positive (FP) test results (indicating PAD when the ulcer was venous) and increased risk of requiring invasive arterial treatment for inappropriately compressed arterial/mixed ulcers following a false-negative (FN) test result (indicating venous when underlying disease was arterial/mixed). It was assumed that any inaccurate tests would be identified within the 24-week time horizon of the decision tree.

The surviving proportion of the cohort then entered arterial, mixed or venous ulcer Markov models depending on their true underlying disease classification. The venous disease model included five mutually exclusive health states, centred around ulcer healing (healed index ulcer, unhealed index ulcer, recurrence, healed post recurrence and death). The arterial and mixed disease models included four health states, focusing on the long-term outcomes of the arterial component of disease [critical limb ischaemia (CLI), healed post CLI, amputation and death]. The decision to structure the mixed Markov model similarly to the arterial-only model was based on discussion with clinical experts who explained that, in clinical practice, the arterial component of disease is likely to take priority in the patient's care pathway.

Costs were based on National Health Service and Personal Social Service perspective costs (2021 values) and included:

- micro-costing of the automated and manual Doppler devices
- costs of applying compression for the unhealed duration of a venous ulcer
- costs of referral to vascular services for test-positive patients, including the additional costs of unnecessary referral for patients with a FP test result
- costs of treating arterial disease, including endovascular and bypass procedures as well as follow-up nursing care
- long-term follow-up costs in the Markov model included the cost of managing recurrent venous ulcers, recurrent CLI and long-term health and social care costs of amputation.

Health state utility values were obtained from the literature and were based on EuroQol-5 Dimensions data, valued using the UK value set where possible. Utilities were combined with mortality estimates for each health state to calculate QALYs. In the decision tree, utilities were dependent on the duration of ulcer healing time for venous ulcers, and whether patients had CLI for those with arterial/mixed disease. All utilities were adjusted for UK age- and sex-specific general population norms, allowing the cohort to experience reduced utility as they aged over subsequent model cycles.

Expected costs and QALYs were accumulated over a lifetime horizon, in 6-monthly cycles and an annual discount rate of 3.5% per annum was applied to future costs and QALYs. Probabilistic analyses (Monte Carlo simulation with 1000 draws for each parameter) were conducted for a range of pessimistic and optimistic alternative base-case scenarios. A full range of deterministic scenarios explored the impact of alternative sources of model inputs and assumptions on cost-effectiveness results.

Results

Nature, description and quality of the available evidence

The database searches identified 110 unique records, 79 records were supplied by the respective companies and 2 further studies were identified from reference lists. Twenty-four studies, published in

26 papers, were included in the systematic review of clinical effectiveness. Two studies enrolled specifically people with leg ulcers (167 participants in total) while the remaining studies (4258 participants in total) included people from primary care practices, cardiovascular risk services, vascular services and from epidemiological/general population-based studies. All studies used an ABPI threshold of 0.9. In healthy people, ABPI would be expected to be > 0.9. Most of the studies assessed the performance of a single automated device with only one study comparing two devices (WatchBP and MESI ABPI MD). Regarding the type of automated devices, two studies provided data on the BlueDop Vascular Expert device, four studies on the BOSO ABI-System 100, six studies on the Dopplex Ability, eight studies on the MESI ABPI MD and five studies on the WatchBP Office. No studies assessed the performance of the WatchBP Office Vascular and the MESI mTABLET ABI devices. Apart from one study conducted in New Zealand, all included studies were conducted in Europe (six in the UK). The risk of bias of included studies was assessed using the QUADAS-2 tool. Most studies were judged at low risk for the index test domain and at unclear risk for the patient selection, reference standard and flow and timing domains. The risk of applicability concerns was low in most studies.

Summary of benefits and risks

The two studies assessing people with leg ulcers did not provide sensitivity and specificity estimates but reported that automated devices gave generally higher readings than manual Doppler. The results of the 22 studies assessing people without leg ulcers varied. Seventeen studies reported sensitivity and specificity estimates for the detection of PAD and showed that the automated devices had good sensitivity but only moderate sensitivity indicating that a proportion of people with PAD would be missed. Sensitivity of BlueDop Vascular Expert ranged from 66% to 95% and specificity from 90% to 94% in two studies; sensitivity of the BOSO ABI-System 100 ranged from 61% to 77% and specificity from 94% to 98% in three studies; sensitivity of Dopplex Ability_ranged from 20% to 79% and specificity from 86% to 96% in four studies; sensitivity of the MESI ABPI MD ranged from 57% to 75% and specificity from 67% to 99% in five studies; sensitivity of the WatchBP Office ABI ranged from 44% to 83% and specificity from 97% to 100% in four studies;

We were able to combine results across 12 studies (2004 participants in total) and 3 automated devices. The pooled sensitivity and specificity for PAD diagnosis using automated ABPI were 64% [95% confidence interval (CI) 57% to 71%] and 96% (95% CI 92% to 98%), respectively. Regarding the performance of individual devices, the pooled sensitivity for MESI ABPI MD was 67% (95% CI 59% to 74%) and the pooled specificity 94% (95% CI 83% to 98%); the pooled sensitivity for WatchBP Office ABI was 53% (95% CI 37% to 69%) and the pooled specificity 98% (95% CI 96% to 99%). For the remaining devices, we could not conduct meaningful meta-analyses due to the limited number of available studies.

Summary of cost-effectiveness, including sensitivity analyses

The uncertainties in the diagnostic accuracy evidence base and the unclear link between test results and patient management mean it is difficult to draw any firm conclusions on cost-effectiveness. A lack of evidence on the impact of the tests on important patient outcomes, the extent to which inaccurate test results would be identified in practice and the implications of acting on inaccurate test results contribute further uncertainty to the assessment of cost-effectiveness. Automated tests were less costly to deliver due to shorter testing times, but in most modelling scenarios, these cost savings were quickly offset by any additional risks and costs associated with withholding compression (FP) or inappropriately applying compression (FN). Given the current evidence base, it is unlikely that the automated tests would generate QALY gains or cost savings, unless a high proportion of FP and FN tests could be reliably identified in clinical practice through holistic patient assessment, and automated tests could deliver improvements in patient referral over manual Doppler testing.

Discussion

Strengths, limitations of the analyses and uncertainties

The methods used to conduct this assessment were detailed, thorough and in line with current methodological standards. We identified only two studies assessing the performance of automated devices in determining ABPI in people with leg ulcers. Given the current lack of evidence in people with leg ulcers, we decided to widen our target population to include studies assessing the use of automated devices for measuring ABPI in different settings. We identified and summarised 22 studies focusing on people without leg ulcers.

The main limitations of the clinical effectiveness assessment are summarised below.

- Lack of evidence on people with leg ulcers to draw any meaningful conclusion about this clinical population.
- Considerable clinical heterogeneity – in terms of characteristics of the patient population, setting and testing procedures – across studies that focused on people without leg ulcers.
- Suboptimal agreement between readings of the automated devices and those of the manual Doppler with a systematic tendency towards higher automated readings.
- Use of manual Doppler as the reference standard for detection of PAD.
- Variation in the prevalence of PAD across studies.
- Limited data on the performance of the automated devices in relevant subgroups of patients (e.g. diabetes patients).
- Uncertainty about the optimal threshold for automated ABPI measurement.
- Uncertainty about the potential role of automated devices in clinical practice (screening tool, alternative/adjuvant tool to current manual Doppler).
- Lack of data on the impact of the routine use of automated devices on health outcomes (e.g. the consequences of a delayed diagnosis because of FN results).
- No data on the WatchBP Office Vascular and MESI mTABLET ABI devices.

With regard to the economic modelling, we identified the following areas of uncertainties that complement those identified for the review of clinical effectiveness evidence and raise doubt about the robustness of the cost-effectiveness results:

- A lack of data regarding the impact of different tests on patient-relevant outcomes such as ulcer healing.
- It is unclear whether automated tests could achieve tangible benefits in terms of a reduced time to compression therapy in patients with venous disease. Any benefits would rely on a lack of skills to complete manual Doppler assessment in the community, and it is unclear how widespread such a skill shortage might be.
- Uncertainty around whether inaccurate test results might be identified during clinical evaluation of patients during a testing appointment, and thus the extent to which inaccurate results would be acted upon in clinical practice [i.e. if tests would lead to inappropriate compression of arterial ulcers (FNs), or delayed time to compression (FPs)].
- Limited data regarding the costs and outcomes specifically for mixed ulcer disease.

Generalisability of the findings

It is unclear how the results of studies assessing the accuracy of automated devices for measuring APBI in people without leg ulcers could be generalised to people with leg ulcers.

Conclusions

Future research is needed to evaluate the use of automated devices within specific populations (people with leg ulcers) and relevant settings. For the broader use of automated devices in clinical practice, more robust evidence is required to establish whether the use of automated devices is appropriate and cost-effective for the general screening of clinical populations with any vascular concerns. In addition, evidence is needed to support the use of automated devices as an alternative or adjunct to manual Doppler in people with symptoms of PAD.

Study registration

This study is registered as PROSPERO CRD42022327588.

Funding

This award was funded by the National Institute for Health and Care Research (NIHR) Evidence Synthesis programme (NIHR award ref: NIHR135478) and is published in full in *Health Technology Assessment*; Vol. 28, No. 37. See the NIHR Funding and Awards website for further award information.

Health Technology Assessment

NICE TAR and DAR

ISSN 2046-4924 (Online)

Impact factor: 3.6

A list of Journals Library editors can be found on the NIHR Journals Library website

Launched in 1997, *Health Technology Assessment* (HTA) has an impact factor of 3.6 and is ranked 32nd (out of 105 titles) in the 'Health Care Sciences & Services' category of the Clarivate 2022 Journal Citation Reports (Science Edition). It is also indexed by MEDLINE, CINAHL (EBSCO Information Services, Ipswich, MA, USA), EMBASE (Elsevier, Amsterdam, the Netherlands), NCBI Bookshelf, DOAJ, Europe PMC, the Cochrane Library (John Wiley & Sons, Inc., Hoboken, NJ, USA), INAHTA, the British Nursing Index (ProQuest LLC, Ann Arbor, MI, USA), Ulrichsweb™ (ProQuest LLC, Ann Arbor, MI, USA) and the Science Citation Index Expanded™ (Clarivate™, Philadelphia, PA, USA).

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: journals.library@nihr.ac.uk

The full HTA archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hta.

Criteria for inclusion in the *Health Technology Assessment* journal
Manuscripts are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

HTA programme

Health Technology Assessment (HTA) research is undertaken where some evidence already exists to show that a technology can be effective and this needs to be compared to the current standard intervention to see which works best. Research can evaluate any intervention used in the treatment, prevention or diagnosis of disease, provided the study outcomes lead to findings that have the potential to be of direct benefit to NHS patients. Technologies in this context mean any method used to promote health; prevent and treat disease; and improve rehabilitation or long-term care. They are not confined to new drugs and include any intervention used in the treatment, prevention or diagnosis of disease.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

This article

The research reported in this issue of the journal was commissioned and funded by the Evidence Synthesis Programme on behalf of NICE as award number NIHR135478. The protocol was agreed in April 2022. The draft manuscript began editorial review in October 2022 and was accepted for publication in August 2023. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' manuscript and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this article.

This article presents independent research funded by the National Institute for Health and Care Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, the HTA programme or the Department of Health and Social Care. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, the HTA programme or the Department of Health and Social Care.

This article was published based on current knowledge at the time and date of publication. NIHR is committed to being inclusive and will continually monitor best practice and guidance in relation to terminology and language to ensure that we remain relevant to our stakeholders.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Newgen Digitalworks Pvt Ltd, Chennai, India (www.newgen.co).