External Assessment Centre report

The purpose of the External Assessment Centre (EAC) report is to review and critically evaluate the sponsor's clinical and economic evidence and may include additional analysis of the submitted evidence or new clinical and/or economic evidence.

Title: VibraTip for testing vibration perception in the detection of diabetic peripheral neuropathy

Produced by: Newcastle upon Tyne Hospitals (NUTH) and York Health Economics Consortium (YHEC) External Assessment Centre (EAC).

Authors: Iain Willits

Roseanne Jones

Helen Cole

Mick Arber

Joyce Craig

Andrew Sims

Correspondence to: Iain Willits

Regional Medical Physics Department (RMPD)

NCCC, Freeman Hospital

High Heaton

Newcastle-upon-Tyne

NE7 7DN

Date completed: 28/04/2014

Declared interests of the authors

None

Acknowledgements

The authors would like to thank the following expert advisers who were appointed by NICE and provided advice about the detection of diabetic peripheral neuropathy to complement the clinical and economic evidence:

- Dr Adrian Wills, Consultant Neurologist, British Peripheral Nerve Society, Nottingham University Hospitals NHS Trust.
- Mr Allister Campbell, Deputy Podiatry Services Manager, Specialist Society, South Eastern Health and Social Care Trust, Lisburn Hospital, NI.

- Dr Frances Game, Consultant Diabetologist, Specialist Society, Derby Hospitals.
- Mr George Dunn, Specialist Podiatrist, Specialist Society, East Cheshire NHS Trust.
- Dr John Winer, Consultant Neurologist, Specialist Society, Queen Elizabeth Hospital, Birmingham.
- Dr Paul Chadwick, Principal Podiatrist, Specialist Society, Salford Royal Foundation Trust.
- Dr Umesh Dashora, Consultant Endocinology and Diabetes, Specialist Society, East Sussex Healthcare.
- Dr. Yusuf A. Rajabally, Consultant Neurologist, Honorary Senior Clinical Lecturer, Specialist Society, University Hospitals of Leicester, Leicester General Hospital.

We would also like to thank Dr John Allen (lead clinical scientist) at NUTH for his valuable insight into diabetic neuropathy and its diagnostic evaluation in the NHS.

Rider on responsibility for report

The views expressed in this report are those of the authors and not those of NICE. Any errors are the responsibility of the authors.

Contents

1 Summary 5
2 Background 11
2.1 Overview and critique of sponsor's description of clinical context 11
2.2 Overview of sponsor's description of ongoing studies 20
2.3 Critique of sponsor's definition of the decision problem 21
3 Clinical evidence 29
3.1 Critique of the sponsor's search strategy 29
3.2 Critique of the sponsor's study selection 31
3.3 Included and excluded studies 31
3.4 Overview of methodologies of all included studies 36
3.5 Overview and critique of the sponsor's critical appraisal 40
3.6 Results 57
3.7 Description of the adverse events reported by the sponsor 74
3.8 Description and critique of evidence synthesis and meta-analysis carried out by the sponsor 75
3.9 Additional work carried out by the External Assessment Centre in relation to clinical evidence 75
3.10 Conclusions on the clinical evidence 75
4 Economic evidence 80
4.1 Published economic evidence 80
4.2 De novo cost analysis 82
4.3 Results of de novo cost analysis 103
4.4 Interpretation of economic evidence 108
4.5 Additional work undertaken by the External Assessment Centre in relation to economic evidence 110
4.6 Conclusions on the economic evidence 113
5 Conclusions 116
6 Implications for research 118
References 120
Appendix 1. Supplementary information of literature searches performed by EAC. 125
Appendix 2. Technical studies submitted by sponsor 142
Appendix 3: EAC searches for economic studies 147
Appendix 4: Quality assessment of sponsor's *de novo* economic model 149

1 Summary

Scope of the sponsor's submission

The sponsor's submission matched the scope of the decision problem in the domains of population (people with diabetes being assessed for diabetic peripheral neuropathy [DPN]); intervention (VibraTip); comparators (10g monofilament, 128 Hz tuning fork, biothesiometer); cost analysis, and subgroups (people undergoing routine assessment in primary care and people undergoing specialised assessment in secondary care).

The sponsor's submission did not fully address the scope in terms of outcomes, with only 'Sensitivity and specificity in assessment of vibration perception' and 'Accuracy of risk assessment in ulcer formation' being reported in the clinical evidence submission, and 'Ulcer formation and amputation' being reported in the economic submission (although this was not supported by clinical evidence). The principal reason for this was that the identified studies did not report on the outcomes that were described in the scope.

Summary of clinical evidence submitted by the sponsor

The sponsor did not provide information on a literature search that could be replicated by the EAC. The sponsor provided the EAC with a total of 9 studies for consideration (1-9), one of which they claimed was out of scope (4) and three of which were unpublished at the time of writing the EAC assessment report (5, 7, 8).

The sponsor provided an adequate description of the characteristics of the studies, but did not critically appraise them appropriately. The sponsor reported the results from the studies as they appeared in the manuscripts. However, the sponsor did not attempt to appraise the author's conclusions or place the results in the context of the decision problem.

Summary critique of clinical evidence submitted by the sponsor

The EAC performed a literature search and identified 2 additional studies that were directly relevant to VibraTip. However, these were excluded on the grounds that one was a commentary that did not have primary efficacy data (10), and one was a conference abstract that reported on the same patients of a full study published at a later date (11). The EAC also excluded 3 studies identified by the sponsor. Two of these were technical papers by Horsfield and Levy (2012 and 2014) (4, 5). Whilst these were not directly relevant to the scope as they were not performed in human subjects, they did provide useful information pertaining to the decision problem. The other study was a randomised controlled trial. This was excluded as it did not match the scope in the population domain (7).

The six remaining studies were all cross-sectional diagnostic accuracy studies and were all considered to be relevant to the scope. These were the studies by Levy (2010) (6), Bowling *et al.* (2012) (1), Bracewell *et al.* (2012) (2), Urbancic-Rovan *et al.* (2012) (9), Garbas *et al.* (2013) (3) and Nizar *et al.* (2014) (8). The EAC critically appraised these studies using QUADAS-2 methodology. In terms of internal validity all the studies showed bias in the domain of patient selection because of evidence of selection bias and/or poor recruitment policy. All the studies had a high risk of bias in the index test domain, and all but one study (2) had an additional high risk of bias in the reference test domain. All the studies had a low risk of bias in the flow and timing domain where the reporting was adequate. In terms of applicability (external validity or generalisability), it was generally uncertain how applicable the populations were to the decision problem. The index test was compromised in all but one study (1), and the reference standard was only appropriate in one study (2).

The biggest limitation with 3 of the studies was the lack of a reference standard (3, 6, 9). Although the EAC fully appraised the results from these studies (which were all poorly reported), without the reference standard it was not possible to calculate true diagnostic accuracy. The study by Bowling *et al.* (2012) (1) was performed in a population who had had DPN pre-diagnosed, and reported on the target condition of 'at risk' feet. This study was considered to be of limited value because use of VibraTip or its comparators is not routinely used to detect DPN in this very high risk population. The study by Nizar *et al*. (2014) (8) used a modified reference standard, that of a neurothesiometer set to 20 V rather than the usual threshold of 25 V. This difference in vibration perception threshold is clinically significant, and for this reason, this study was judged not to be generalisable to the decision problem.

The study by Bracewell et al. (2012) was judged to be of reasonable quality and most relevant to the decision problem (2). This was because it recruited from the most appropriate population (diabetic patients), used an appropriate reference standard (neurothesiometer set at 25 V), and compared the diagnostic performance of VibraTip with the 10g monofilament and the 128 Hz (comparators specified in the scope). However, the study used ROC analysis to optimise the protocol for each device, and the results may not be generalisable to usual clinical practice. Additionally, although this was the second largest study identified on VibraTip (n = 141), it was likely to have been underpowered and prone to type 2 error.

The results of this study indicated that (relative to the neurothesiometer), VibraTip had a sensitivity of 0.79 (95% CI 0.69 to 0.90) and specificity of 0.82 (95% CI 0.74 to 0.90); 10g monofilament had a sensitivity of 0.84 (95% CI 0.75 to 0.94) and specificity of 0.83 (95% CI 0.75 to 0.91); and the 128 Hz tuning fork had a sensitivity of 0.69 (95% CI 0.57 to 0.81) and specificity of 0.90 (0.84 to 0.97). Thus the 10g monofilament was slightly superior to VibraTip in terms of both sensitivity and specificity, and the 128 Hz tuning fork had poorer sensitivity but superior specificity. Although it was not possible to perform hypothesis testing because of the way the

data was reported, it is unlikely that there was a statistically significant difference between the three devices used in this study.

Summary of economic evidence submitted by the sponsor

The sponsor did not provide a replicable literature search for economic studies concerning VibraTip, and did not provide the EAC with any published studies on the device. The sponsor provided the EAC with a *de novo* economic model written and executed in Microsoft Excel. The basis of the model was that of a decision tree covering a three year time horizon with a starting population of people with diabetes in the UK (estimated at 2.9 million). The model had two arms, which consisted of a current practice arm (10g monofilament or 128 Hz tuning fork) and an intervention arm, where patients were tested with VibraTip or retained practice (10g monofilament or 128 Hz tuning fork). The decision tree simulated patients on a pathway where they had chances of becoming high risk for ulceration, developing ulceration, having continued ulceration, or requiring amputation. Each stage in the model was associated with a transition probability, the number of patients in that state, and the cost of being in that state (device costs and management costs).

In the base case analysis, the sponsor assumed the uptake of VibraTip would be 40%. In addition the sponsors performed deterministic sensitivity analysis where the adoption uptake of VibraTip was assumed to be 20% or 100%, and this was combined with a second sensitivity analysis, whereby an assumption was made that the introduction of VibraTip was associated with a 1% reduction in ulcer formation.

The sponsors reported that in the base case, the introduction of VibraTip would lead to overall cost savings to the NHS of approximately £50,000 compared with the 10g monofilament and £40,000 compared with the 128 Hz tuning fork. From sensitivity analysis, the sponsor reported that increased adoption of VibraTip would lead to proportionate savings, as would be expected. It was calculated that if the use of VibraTip led to a 1% reduction in ulcer formation, then VibraTip would lead to savings of £6,430,000 compared with the 10g monofilament and £6,350,000 compared with the 128 Hz tuning fork, at the level of 40% adoption.

Summary critique of economic evidence submitted by the sponsor

The EAC performed a literature search for economic studies on VibraTip but did not identify any. The EAC critiqued the sponsor’s economic model and accompanying narrative. The model as it was presented to the EAC was not complete; for instance it did not have executable data for the 128 Hz tuning fork scenario. This, as well as other inconsistencies in the model’s structure and populated parameters, made it difficult to replicate the sponsor’s results, as reported in the narrative.

The EAC considered that a fundamental weakness of the model was that there was no diagnostic input into it, despite VibraTip being a diagnostic device. In the base case, this made the model a *de facto* cost minimisation study, with only the costs associated with device use per-examination contributing to the overall costs. Because of this, the overall cost savings reported by the sponsor were very low compared with the overall burden of the condition (calculated as £1,467,860,000 for VibraTip in the model).

The EAC considered the first sensitivity analysis, whereby the sponsor adjusted the adoption uptake of VibraTip which caused a corresponding change in cost savings in a linear manner, was uninformative. The EAC considered that the assumption made in the second sensitivity analysis, that use of VibraTip would lead to a 1% reduction in ulcer formation, had no evidence to support it, thus did not inform the decision problem.

The EAC also did not agree with the per-examination costs the sponsor had assumed for VibraTip and its comparators. The sponsor had calculated that VibraTip would cost 0.2 pence per-examination use, the 10g monofilament would cost 1.0 pence, and the tuning fork would cost 0.8 pence. Using technical papers on the useful life of VibraTip (4) and the durability of the Bailey's 10g monofilament (12), the EAC estimated the most likely cost per-examination using Vibratip would be between 4.0 pence and 10.0 pence. The corresponding figure for the 10g monofilament would be between 3.04 and 7.6 pence, with these ranges overlapping. The EAC could not estimate the per-examination cost for the 128 Hz tuning fork using this methodology because it has an indefinite useful life.

The EAC concluded that the economic case for VibraTip had not been made.

External Assessment Centre commentary on the robustness of evidence submitted by the sponsor

The EAC considered that the clinical evidence to support the use VibraTip for the detection of DPN was limited to 6 cross-sectional diagnostic accuracy studies (1-3, 6, 8, 9), with only 1 of these studies (2) being directly applicable to the decision problem. These studies did not address most of the outcomes described in the scope, and furthermore did not inform the *de novo* economic model. Therefore, the EAC considered the evidence submitted by the sponsor was not robust.

Summary of any additional work carried out by the External Assessment Centre

The EAC carried out the following additional work:

- Critical appraisal of the technical papers submitted by the sponsor (4, 5).

- Recalculation of the per-examination costs of VibraTip and the 10g monofilament.

Abbreviations

AUC	Area under the curve
CASP	Critical Appraisal Skills Programme
CG	Clinical guideline
DPN	Diabetic peripheral neuropathy
EAC	External Assessment Centre
EQ-5D	Euroquol 5 dimensions
FN	False negative
FP	False positive
NDS	Neuropathy disability score
NICE	National Institute for Health and Care Excellence
NLR	Negative likelihood ratio
NPV	Negative predictive value
PICO	Population, Intervention, Comparator, Outcomes
PLR	Positive likelihood ratio
PPV	Positive predictive value
PRISMA	Preferred Reporting Items for Systematic reviews and Meta-Analyses
QALY	Quality-adjusted life-year
QoF	Quality outcomes framework
QoL	Quality of life
QUADAS	Quality Assessment of Diagnostic Accuracy Studies
RCT	Randomised controlled trial
ROC	Receiver operator curve
TN	True negative
TP	True positive
VPT	Vibration perception threshold

2 Background

Throughout this report, the EAC makes reference to specific sections within the sponsor's submission as: (submission, section X.X).

2.1 Overview and critique of sponsor's description of clinical context

2.1.1 Critique of sponsor's description of the background condition

The sponsor provided a brief background description of diabetes and diabetic peripheral neuropathy (DPN). The EAC considered that the conditions described and facts reported were accurate. The EAC has provided a more comprehensive description in the following section.

2.1.2 EAC's overview of the conditions and technology

Diabetes Mellitus

Diabetes Mellitus is a chronic condition that affects around 3.19 million people within the United Kingdom (13). It is a complex metabolic disorder that affects a person's ability to metabolise glucose. There are two main types of diabetes mellitus. Type 1 diabetes is caused by a body's failure to produce insulin and accounts for about one tenth of all cases (14). Type 2 diabetes is caused by the body becoming resistant to insulin and resulting in the hormone working insufficiently. This accounts for about nine tenths of all cases, and is frequently linked with being overweight. Both forms of diabetes have the potential to cause serious complications if left untreated, including heart disease, stroke, blindness and nerve damage (15).

Diabetic Peripheral Neuropathy

Peripheral neuropathy is one of the most common complications of diabetes, affecting between 20 to 40% of people with the condition. About 2.5% (61,000 people with diabetes) are estimated to have a foot ulcer at any given time (16). It is responsible for a substantial increase in mortality and morbidity as well as deterioration in quality of life (QoL). Diabetic peripheral neuropathy (DPN) is characterised by damage to or degeneration of peripheral nerves of the extremities including the sensory, motor and autonomic nerves. Patients with DPN typically have one or more of the following symptoms: numbness, tingling, pain or weakness. The symptoms typically begin in the feet and spread proximally in a length dependent manner, with deterioration of sensory symptoms more prominent than loss of motor function (17).

If left untreated, DPN can cause further serious complications. The main risk is numbness, as minor injuries may go unnoticed and may then develop ulceration and secondary infection. In the worst cases, lower-extremity amputation may be required.

Approximately 5% of people with diabetes may develop a foot ulcer in any year, and amputation rates are around 0.5% per year (18). There is now evidence that improved glycaemic control can prevent the appearance and worsening of polyneuropathy in patients with insulin-dependent diabetes mellitus (19). However, to be as effective as possible, it should be instituted at an early stage, since the late stages of DPN are poorly reversible or irreversible (20).

Small- and large-fibre neuropathy

Peripheral Neuropathy may be either sensory or motor, and involve large fibres, small fibres, or both. Most patients with DPN have a combination of both large and small nerve fibre involvement (21). Large fibre neuropathies affect myelinated nerve fibres which have a large diameter and high conduction velocity. Collectively, these fibres carry sensory information related to muscle sense, touch, vibration, kinesthesia, pain and cold temperature. Large fibre neuropathies are more common in older adults, typically with type 2 diabetes. Signs and symptoms of neuropathy include reduced vibration (often the first objective evidence of neuropathy) and positional sense; weakness; muscle wasting; and depressed tendon reflexes (21). In comparison, small fibre neuropathies affect unmyelinated nerve fibres which have a small diameter and consequently have a low conduction velocity. These fibres carry sensory information relating to nociception, temperature, touch, pressure and itch. Small nerve fibre dysfunction usually occurs early and is more common in younger adults, typically with type 1 diabetes (21). It manifests first in the lower limbs with symptoms of pain and hyperalgesia, followed by a loss of thermal sensitivity and reduced light-touch and pin prick sensation (22).

Pathogenesis and vibration perception

The pathogenesis of DPN appears to be related to both vascular and nonvascular metabolic mechanisms. Hyperglycaemia causes accumulation and saturation of sugar in nerve fibres causing axonal thickening, structural damage, and impaired nerve conduction with progression to axonal loss. Neuronal microvasculature is also impaired in the presence of hyperglycaemia, and causes a decreased capillary blood flow to nerve fibres. Neuronal ischaemia is a well-established characteristic of peripheral neuropathy and is owing to the progressive nature of this disorder.

The loss of vibration sense in DPN is due to damage to the peripheral nerve receptors (mechanoreceptors) present in the superficial and deeper layers of skin and between the layers of muscle and periosteum. Typically, the sense of vibration results from the sinusoidal oscillation of objects placed against the skin. These mechanoreceptors in the skin respond to the oscillations by developing action potentials that are transmitted through their neural afferents, with a pulse code in which each action potential signals one cycle of a sinusoidal wave. The total number of sensory nerve fibres activated by a vibrating stimulus determines the intensity of

vibration; the frequency of firing determines the vibration frequency perceived. Humans are most responsive to vibration at frequencies of 200–250 Hz. If a patient is tested with a 250 Hz vibration near sensory threshold, only nerve receptors right under the contact point in the skin are activated. As the vibratory amplitude is increased, the more distant nerve receptors under the point of stimulation become activated (23). The nerve fibre loss and degeneration observed in peripheral neuropathy primarily affects these nerve fibres first, causing a loss of vibration and touch sense in the distal hands and feet**.**

The technology: VibraTip

VibraTip (McCallan Medical) is a device resembling a small key-ring fob that provides a near-silent vibration of consistent amplitude at a frequency similar to that of a calibrated tuning fork. It is intended for use when testing a person's vibration perception during routine checks for diabetic peripheral neuropathy (DPN) in people who have type 1 or type 2 diabetes.

The VibraTip probe is applied to the patient's foot twice: once while non-vibrating and once while vibrating; the patient is asked to indicate when they feel the vibration. If the vibration is not detected this may suggest the development of DPN and further investigation can be initiated. VibraTip is intended as an alternative or replacement to the devices that are currently used in NHS clinical practice for testing foot sensory function such as the 10g monofilament (light touch/pressure sensation) and the calibrated tuning fork or biothesiometer (vibratory sensation). The device is hoped to provide a consistent application compared with the variable vibration and cold touch of the tuning fork and to offer continuous operation over the battery life compared with the 10g monofilament, which needs resting for 24 hours after every 10 full patient foot examinations (24).

2.1.2. Overview of relevant clinical guidelines

The sponsor correctly identified two clinical guidelines as being highly relevant to the decision problem, both published by the National Institute for Health and Care Excellence (NICE). The most relevant of these guidelines was *Type 2 diabetes foot problems: Prevention and management of foot problem*s (CG10) which was published in 2004 (18), and is due to be updated in 2015. The sponsor provided a brief description of the patient pathways described in this guideline, which the EAC considered accurately represented its content and was a satisfactory description of optimal clinical practice in the population of interest (people with diabetes undergoing foot assessment). Details of the CG10 patient pathway as interpreted by the EAC are reported in section 2.1.3.

The sponsor also identified NICE CG15, *Type 1 diabetes: diagnosis and management of type 1 diabetes in children, young people and adults* (2004) (25), but did not elaborate further on it. Type 1 diabetes usually arises in childhood and is

always insulin dependent, and it is also a significant cause of DPN. The following recommendations are made in CG15 that are relevant to the assessment of VibraTip:

> 1.3.5.3 Children and young people with type 1 diabetes should be offered: annual foot care reviews.
>
> 1.11.3.4 Use of a 10g monofilament plus non-traumatic pin prick is advised for detection of impairment of sensory nerve function sufficient to significantly raise risk of foot ulceration.

Thus in contrast to the recommendations given in CG10, the recommendations in CG15 do not specify the testing of vibratory sensation, for which VibraTip would be anticipated to replace existing methods such as tuning forks.

The sponsor did not discuss other guidelines or policies that might be relevant to the decision problem. The EAC has noted that there are several other guidelines and technology assessments that have been published by NICE which might directly or indirectly relate to the management of diabetic feet 'downstream' of detection of DPN, for instance improved glycaemic management. In addition, the Scottish Intercollegiate Guidelines Network (SIGN) has published *Management of diabetes: a national clinical guideline* (2010); which covers the management of both type 1 and type 2 diabetes (26), and in its foot care assessment and management recommendations is largely concordant with NICE CG10.

Finally, the sponsor did identify that assessment and care of diabetic feet forms an important element of the Quality and Outcomes Framework (QoF) for the General Medical Services contract (13). Specifically relevant to VibraTip is the following Diabetes Mellitus indicator:

> DM012. 'The percentage of patients with diabetes, on the register, with a record of a foot examination and risk classification: 1) low risk (normal sensation, palpable pulses), 2) increased risk (neuropathy or absent pulses), 3) high risk (neuropathy or absent pulses plus deformity or skin changes in previous ulcer) or 4) ulcerated foot within the preceding 12 months'. [score 4 points, target 50-90%].

In the rationale for this indicator, the authors state that foot inspection and assessment should include 'identifying the presence of sensory neuropathy (loss of ability to feel a monofilament, vibration or sharp touch)'.

This QoF indicator is relevant to the assessment of VibraTip for two reasons. Firstly, it provides a clear incentive for primary healthcare providers to perform annual foot assessments on the patients in their practice registered with diabetes mellitus. Secondly, the method of detecting loss of vibrational sense is not specified, meaning

VibraTip could be used as part of the assessment to satisfy this indicator. This has previously been raised as a potential bar to the introduction of VibraTip (10), but this concern appears to be unjustified.

2.1.3. EAC's interpretation of clinical care pathway

The population described in the scope of the decision problem, people (adults and children) with type 1 or 2 diabetes undergoing routine foot-care checks by healthcare workers in primary and secondary care settings, is fully covered by the NICE clinical guideline 10 *Type 2 diabetes foot problems: Prevention and management of foot problems* (CG10) (18). The patient pathways relevant to this guideline has been summarised as an algorithm in the accompanying Quick Reference Guide to CG10 (27), as well as the NICE pathway, illustrated in Figure 2.1. Although this guideline and clinical pathway was developed exclusively for people with type 2 diabetes, it can be reasonably extrapolated to cover people with type 1 diabetes, and some of the evidence base was derived from people with type 1 diabetes or case mixes of both. As discussed, the NICE clinical guidelines on *Type 1 diabetes: diagnosis and management of type 1 diabetes in children, young people and adults* (CG15) (25) does not recommend testing for vibratory sensation, and instead recommends the use of pressure with a monofilament or non-traumatic pin prick. However, NICE CG10 is largely concordant with guidelines published by SIGN, which covers the management of both type 1 and type 2 diabetes (26), as well as being consistent with the QoF indicator DM012, which are also not specific to type 1 or type 2 diabetes (13).

NICE CG10 stipulates that all patients who have been diagnosed with type 2 diabetes mellitus should undergo a full foot examination upon diagnosis and thereafter as part of an annual review. This annual review will usually be carried out in primary care at a GP or nurse-led surgery. The main purpose of the foot examination is to provide a stratified risk assessment for the development of foot ulcer. There are four elements to the examination recommended by NICE, which are:

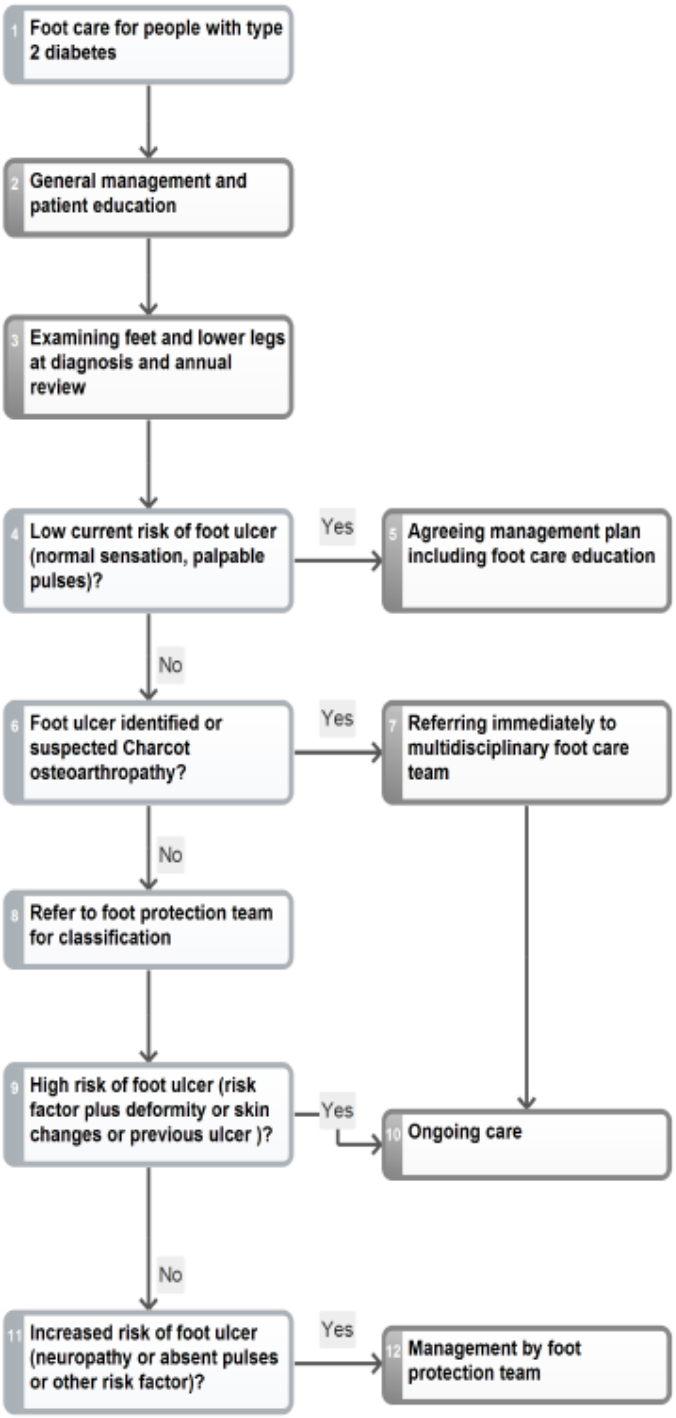
1. Testing of foot sensation [for neuropathy] using 10g monofilament or vibration [method unspecified].
2. Palpation of foot pulses.
3. Inspection for any foot deformity.
4. Inspection of footwear.

Assessment of risk of foot ulceration is based on the first three elements, and works on a clinical prediction rule based on the combination of risk factors such that a person is graded as low current risk if they have normal sensation *and* palpable

pulse *and* no foot deformity. If the person does not have current ulceration and is not considered to be low risk based on these criteria, NICE recommends referral to a foot protection team for further assessment (active ulceration requires urgent referral to multidisciplinary foot care team). Following referral, the patient is reassessed and graded as:

- Increased risk: neuropathy *or* absent pulses *or* other risk factor [such as skin changes or previous ulceration].
- High risk: neuropathy *or* absent pulses *and* deformity *or* skin changes *or* previous ulceration.

Figure 2.1. *NICE patient pathway, Foot care for people with type 2 diabetes (available online at* http://pathways.nice.org.uk/pathways/diabetes*).*



Following assessment, ongoing management is then based on the risk of ulcer formation such that:

- People at low risk are given ongoing foot care education and reviewed annually.

- People at increased risk require a review for the need for vascular assessment, an evaluation of footwear, enhanced foot care education, and more frequent assessments (at 3 to 6 month intervals) by a foot protection team.

- People at high risk should be reviewed by the foot protection team at 1 to 3 monthly intervals. Assessment should include an inspection of feet; need for vascular assessment; appropriate provision of foot care education, footwear and insoles; and skin and nail care.

- People who have active foot ulceration should be urgently referred to a multi-disciplinary foot care team. Treatment options include assessment for need for revascularisation; appropriate wound management; total contact casing; and optimisation of blood glucose levels. When the ulcer is healed, patients should be managed from then onwards as high risk patients.

2.1.4 NICE's rationale for clinical pathways described by CG10

The rationale behind the patient care pathway recommendation for assessment of feet in people with diabetes is given in the full guideline of NICE CG10 (28).

Regarding the need for regular foot examinations to assess the risk of ulcer formation, NICE made the following evidence statement:

> 'Evaluation of skin, soft tissue, musculoskeletal, vascular, and neurological condition on an annual basis is important for the detection of feet at raised risk of ulceration'.

This was based on the grade Ib evidence, which is 'evidence from at least one randomised controlled trial'. This was the prospective study by Pham *et al.* (2010) (29). Regarding the assessment of neuropathy, NICE made three evidence statements:

- 'Both vibration perception threshold measurement using a biothesiometer and sensation threshold measurement using a 10 gram monofilament accurately predict neuropathic patients at raised risk of ulceration. The 10 gram monofilament is convenient and easy to use'.

- 'Identification of neuropathy based on insensitivity to a 10 gram monofilament is convenient and appears cost-effective'.

- ‘Comparisons of different methods of measuring neuropathy reveal no dominant technique’.

These evidence statements were all classed as grade III meaning they were derived from ‘non-experimental descriptive studies, such as comparative studies, correlation studies and case-control studies’.

Moving from evidence statements to recommendations, NICE comment ‘No direct evidence was found to identify the optimal content or frequency of visual inspections and examinations, but regular surveillance at a minimum frequency of once annually was held by the guideline development group to be good clinical practice where no complications have previously been found’. Thus the recommendations for the use of simple, composite clinical prediction rule to assess risk of ulcer formation and therefore guide ongoing management was a pragmatic one guided on the expertise provided by the guideline development group.

2.1.5. Issues relating to current practice

In submission section 3.4, the sponsor reported the following issues regarding the current clinical practice concerning the assessment of diabetic feet [summarised by EAC]:

- If the 10g monofilament is used to test foot sensation, the monofilament is required to be ‘rested’ for 24 hours following use on 10 patients. This limitation is considered in the full version of CG10 (28), and is derived from a technical paper by Booth and Young (2000) (24).

- If vibration perception testing is performed with a 128 Hz tuning fork, there is considerable potential for user variability to negatively impact on the reproducibility of the test. The sponsor alludes to the fact that although the frequency of tuning forks are calibrated (typically to 128 Hz), the amplitude is dependent on the strength of the initial operator strike, which also rapidly decays. An additional problem raised by the sponsor was that the patient may be subject to cueing due to audible sound and temperature. The variability in tuning fork amplitude was described in an unpublished paper provided by the sponsor (5).

The EAC accepts that these are genuine issues with the methods currently recommended by CG10 to test for foot sensation, which are the two most commonly used comparators of the current assessment. However, the extent to which these practical issues impact on the clinical effectiveness of these devices is unclear.

2.1.6. Potential changes to pathway introduced by VibraTip

The sponsor stated that 'the adoption of VibraTip would not require existing pathways of care, as outlined in 3.3 above [brief description of clinical pathway], to be changed'. The sponsor continued to state that 'no changes [would be] required' with the introduction of VibraTip in terms of service organisation (section 3.6); introduction of additional tests or investigations (section 3.7); or additional facilities technologies and infrastructure (section 3.8). The sponsor states in section 3.9 that the introduction of VibraTip could lead to disinvestment in 10g monofilaments after their useful life had expired. Investment in tuning forks could be reduced but they would still be required for specialised neurological examination, auditory investigations, and fracture identification.

The EAC notes that, if introduced into clinical practice, VibraTip would probably directly replace either the 10g monofilament or the test for vibrational sense, which in clinical practice would usually be the 128 Hz tuning fork (or possibly a neurothesiometer or biothesiometer in specialised clinic). In CG10, NICE do not specify a preference for either method, although they note the convenience and ease of use of the 10g monofilament. The EAC has also noted during discussion with the sponsor the possibility that VibraTip could be used in conjunction with the 10g monofilament, to increase the diagnostic accuracy of the sensation testing. This possibility is also discussed in the Question and Answer (Q&A) section of the sponsor's website (www.vibratip.com), where it is stated 'As it takes only a few seconds, perhaps it would be worth adding VibraTip™ use to your current monofilament protocol to see how the results compare?'. If this practice was undertaken then VibraTip would represent an additional action (and cost) to the current care pathway.

2.2 Overview of sponsor's description of ongoing studies

The sponsor stated that VibraTip was a winner in the 'Smart Solutions for Healthcare competition' run by TrusTECH (the North West NHS Innovation Hub), and this secured a prize that procured funding for an ongoing study (30). This trial is listed on Clinicaltrials.gov (identifier number NCT01878682) and was also identified by the EAC literature search (section 3.3). However, when accessed on 12th March 2014, the trial was listed as 'not yet open for participant recruitment' and results are therefore not available for this assessment.

In its submission, the sponsor also stated there was a study expected from the Diabetes Research Unit at Le Pitie Salpetriere Hospital in Paris, although the design and publication intention for this study is not yet known, hence it is not anticipated that this study will inform the assessment.

The EAC did not identify any other studies currently being undertaken or planned from its literature search (see section 3.3).

2.3 Critique of sponsor's definition of the decision problem

In its submission, the sponsor did not expand on the nature of the decision problem compared with the final scope provided by NICE. No deviation from the scope was described by the sponsor (submission, Table A1). The EAC's interpretation of the key aspects of the decision problem is described in the following sections.

2.3.1. Population

The population described in the scope was 'People (adults and children) with type 1 or 2 diabetes undergoing routine foot-care checks by healthcare workers in primary and secondary care settings'.

The EAC can confirm that, in general, the population described in the scope matched the populations that were recruited in the diagnostic accuracy studies relevant to VibraTip identified by the sponsor. However, diabetes is a common disorder affecting a large, heterogeneous population, who have a varying underlying prevalence of DPN which, for instance, increases with age (31). Therefore caution is required when extrapolating results of these studies to the population described in the scope, which is broad.

The scope clearly states that the population to be included is those undergoing foot-care checks by a healthcare professional only (also see section 2.3.6). Therefore use of the device by a patient on themselves, or for any other purpose, is out of scope.

The population described in the sponsor's *de novo* economic model was people with diabetes. This matched the scope.

2.3.2. Intervention

The intervention is described in the scope and by the sponsor as simply 'VibraTip'. The sponsor has provided the EAC with the relevant CE (Conformité Européenne) documentation (CE certificate number No. G2130749367024), and the EAC is assured the product satisfies relevant mandatory regulatory requirements. However, although the description of the intervention as a physical entity is easily described, the mode in which VibraTip is used is not necessarily completely straight-forward.

The instructions for use for VibraTip, reported in the manufacturer's press pack (32) describe how, with the device held gently between thumb and index finger, VibraTip should be used to touch the patient's intact skin twice, each time for about 1 second, whilst the operator explains that 'this is touch one' and 'this is touch two' (to minimise audible cueing). VibraTip should be randomly activated on either the first or second touch, with the patient asked to confirm which of the two touches was associated with vibration.

However, the instructions for use do not specify which site of the foot should be used, whether multiple sites on the foot should be used, whether both feet should be assessed, and, if so, how many insensate readings would be required for a provisional diagnosis of DPN to be made. The EAC has found that there is considerable uncertainty and clinical variation in practice surrounding the use of sensation testing to diagnose DPN, and this may have a material impact on the generalisability of the submitted clinical studies (see section 3). For instance, although not specified in NICE CG10, there are various guidelines that recommend testing of specific sites of the foot to detect DPN. An example is the Foot Risk Awareness and Management Education (FRAME) project, where it is recommended that five sites (plantar aspect of 1st hallux, 2nd, 3rd and 5th metatarsophalangeal joints and plantar aspects of both heels) are tested (33). However, this is just one example, and feedback from clinical experts suggests it is not universally adopted, with other methods also being used, such as those recommended by the International Working Group for Diabetic Foot (IWGDF) (2012) (34) and the American Diabetes Association (2014) (35). This issue is being investigated in the update of CG10 due in 2015 (2).

The sponsor has confirmed with the EAC that VibraTip should be used as a direct replacement for the tuning fork or 10g monofilament, in such a way that it matches the use of its comparators. For instance, if an individual clinician's practice was to use the 10g monofilament on five sites of a foot, then, if introduced, VibraTip should be used on these same five sites too. However, whilst this clarifies the intended use of VibraTip as a direct replacement intervention, it does not clarify the uncertainty surrounding the diagnostic accuracy of differing approaches to physical examination.

Another complication that arises is when consideration of the purpose of the intervention is taken into account, from a procedural viewpoint, detection of sensory loss and thus diagnosis of DPN is not undertaken as an end in itself, but as part of a larger diagnostic pathway to assess for the risk of foot ulceration. According to the clinical algorithm reported in CG10 (18), if a patient is diagnosed with neuropathy, then they will at the very least be categorised as being at increased risk of ulcer formation, and require referral to a foot protection team. However, if a person has an absent pulse, they can also be categorised as being at increased or high risk of ulcer formation, even if neuropathy is absent. The EAC does not have data on the relative proportions of these discrete signs that are found to inform ulcer risk stratification (for instance the relative prevalence of DPN, absent pulse, and other risk factors), although it has been estimated in the past that around 50% of people presenting at dedicated diabetic foot clinics have neuropathic feet (36). Thus, if VibraTip was found to have greater sensitivity and/or specificity than its comparators for the detection of neuropathy, it may be possible to infer this would lead to more accurate risk stratification, but it would not be possible to accurately quantify this, as the proportional input of other risk factors would remain unknown.

2.3.3. Comparator(s)

Three comparators are listed in the scope which are listed as ‘a 10g monofilament, a calibrated tuning fork, [a] biothesiometer’. The EAC has assumed that the neurothesiometer, which has very similar functional capabilities to a biothesiometer, is also within scope.

The comparators described in the scope in general correctly match those that were investigated in the diagnostic accuracy studies supplied by the sponsor to assess the efficacy of VibraTip. Other comparators that have been investigated in these studies include the Ipswich touch test, NeuroTip, Tip Therm and Neuropad.

The EAC understands from the documented correspondence from the clinical experts that, in practice, the 10g monofilament and 128 Hz tuning fork are most likely to be used, with the former predominating in assessment of DPN in primary care. Biothesiometers are likely to be restricted to specialist practice where the patient is either at very high risk of DPN or is being assessed for the severity of extent of sensation loss. In the diagnostic accuracy studies submitted, biothesiometers or neurothesiometers were generally considered to be the reference standard (gold standard) for the detection of DPN.

There was sometimes considerable uncertainty regarding the nature of the comparators used in the diagnostic accuracy studies of VibraTip. Often, this related to matters of technique (as described in section 2.3.2.), concerning the number of sites the device should be used on (and number of insensate sites required for detection of DPN). At other times, the physical nature of the comparator was unclear, usually because of poor reporting quality (e.g. frequency of tuning fork not stated).

Further details of the comparators are given in the sections below.

10g Monofilament

Monofilaments, otherwise known as Semmes-Weinstein monofilaments, are nylon filaments that are constructed to buckle when a specified force is applied. They essentially assess the threshold of sensation for light touch or light pressure in a semi-quantitative manner. The loss of the ability to detect 10g of pressure at one or more anatomic sites on the plantar surface of the foot has been associated with loss of large-fibre nerve function.

Although monofilaments are calibrated to provide a specified force, inherent errors may be introduced during the measurement causing varying pressure and misdiagnosis. The approximated pressure varies in real-world conditions, such as the angle of the filament with the skin, the existence of any tip slip artefacts, ambient versus skin temperature, and humidity. With repetitive use, the stiffness and bending

properties of the monofilament are altered and clinicians are recommended to rest or change the monofilament after a certain period of use (37).

An Australian Health Technology Assessment (HTA) identified two studies which investigated the diagnostic accuracy of the device (38). Both these studies found that the 10g monofilament had a relatively high sensitivity (83% and 91%) but a low specificity (51% and 34%), meaning that it would be more suitable as a diagnostic device, but not a screening device (due to the potential for false positive results).

Systematic reviews have been published concerning the diagnostic accuracy (for DPN) (39, 40) and prognostic capabilities of the 10g monofilament (for the prediction of ulceration and amputation) (41). A common conclusion from the authors of these reviews is that the 10g monofilament (or indeed other forms of physical examination, including vibratory perception) alone is not adequate to screen for or diagnose DPN or assess for the risk of foot ulceration alone, and that other methods should be used in combination. This was also the conclusion of CG10 (18) and will be further considered in the update of this guideline.

128 Hz Tuning fork

The calibrated tuning fork is currently used in clinical practice and provides an easy and inexpensive test of vibratory sensation. Two types of calibrated tuning forks exist, the conventional calibrated tuning fork and the Reidel Seiffer calibrated tuning fork. The conventional calibrated tuning fork is a non-graduated instrument that measures the presence or absence of vibration, but does not quantitatively provide the degree of dysfunction of vibration sense. The Reidel Seiffer calibrated tuning fork, however, is a graduated tuning fork which is able to semi-quantitatively assess vibration sense.

Vibratory sensation should be examined over the tip of the great toe bilaterally by asking the patient to respond “yes” or “no” when asked whether they can feel a vibration. An abnormal response is typically defined as when the patient loses vibratory sensation and the examiner still perceives it while holding the fork. As with the 10g monofilament, there is no consensus amongst experts on which anatomical sites the tuning fork should be used on.

In general, the existing evidence to support the diagnostic accuracy of the tuning fork for the detection of DPN is lacking, in terms of both quality and quantity, which is often the case for older more established technologies. An Australian HTA examined the evidence for the diagnostic performance (for DPN) and predictive performance (for foot ulceration and amputation) of the 128 Hz tuning fork (38). The authors concluded, from three ‘average’ quality studies, that there was moderate evidence for the tuning fork in detecting DPN, and there was substantial evidence for the device being a good predictor of foot ulceration, but not amputation. One study

identified by the HTA found that the predictive value of the tuning fork for foot ulceration was dependent on the vibration perception threshold (VPT) of the patient, with the device being sensitive at detecting DPN if the patient was insensate at ≥25 V using a biothesiometer (42).

A systematic review (search date January 1966 to November 2009) identified only one study that fitted the inclusion criteria (use of nerve conduction studies as reference standard) (40). This was a prospective longitudinal study which assessed various patient characteristics for risk of ulceration. However, the sensitivity and specificity of the 128 Hz tuning fork were not stated in this study.

Biothesiometer / Neurothesiometer

The biothesiometer is an electrically-powered device that allows an examiner to raise the amplitude of vibration of a vibrating tip that oscillates at a constant frequency, typically 120 Hz. The neurothesiometer is a newer device than the biothesiometer, and has a self-contained battery allowing for greater portability, making it more suitable for diabetes screening programmes. The neurothesiometer has now effectively replaced the biothesiometer in clinical practice.

With the vibrating tip placed against the skin of a patient, the examiner adjusts the voltage dial to increase the amplitude of the vibration until the patient feels and reports the vibration. The vibration perception threshold recorded is the voltage (measured in V) at which the patient was able to identify the stimulation. The major limitation with the neurothesiometer and biothesiometer and similar devices such as the Vibrameter, Vibratron II and Maxivibrometer is that they produce a mechanical vibration wave, which progresses and spreads through superficial and deep tissues around and away from the point of contact between the vibrating tip and skin (37).

2.3.4. Outcomes

The outcomes listed in the scope, and their reporting in the included studies, are listed in Table 2.1. It can be seen that most of the outcomes listed in the scope were primarily diagnostic in nature, relating to the diagnostic accuracy of VibraTip and its comparators or other immediate outcomes related to its use. Two outcomes (ulcer formation and amputation, and QoL) were related to management events resulting downstream of the diagnostic pathway. Few of the outcomes specified matched those of the studies supplied.

Table 2.1. *List of outcomes in the scope and summary of studies reporting these outcomes.*

Outcome	**Diagnostic (D) or management (M)?**	**Number of studies outcome included (see section 3)**	**Comment**
Sensitivity and specificity in assessment of vibration perception and/or light touch	D	6	Sensitivity and specificity data had to be inferred or extrapolated from some studies, as no reference standard was used.
Sensitivity and specificity in assessment of grade of neuropathy	D	None	Neuropathy usually regarded as a dichotomous outcome (i.e. grading not appropriate).
Inter-rater agreement of assessment of grade of neuropathy	D	None	*Intra*-rater agreement assessed in 1 study.
Accuracy of risk assessment in ulcer formation	D	1	All patients in study had pre-diagnosed DPN.
Ulcer formation and amputation	M	None	Outcome in economic analysis.
Time taken for sensory testing	D	None	Outcome not used in submission.
Quality of life (QoL)	M	1	QoL related to home use of device by patient, so this paper out of scope.
Device-related adverse events	D/M	None	Not relevant for this device.

2.3.5. Cost analysis

The economic analysis provided by the sponsor, including the *de novo* model, matched that of the scope in all domains (see section 4). For the economic analysis, only the 10g monofilament and 128 Hz tuning forks were considered as comparators (in line with the scope), which the EAC considered was the correct approach. This was because the neurothesiometer (or biothesiometer), which can be used to test for vibration perception threshold (VPT), is used in specialist settings only, and would not be replaced by VibraTip or any other screening devices. In the submission, the sponsor gave details regarding the cost of the neurothesiometer, but did not include its use in the model.

2.3.6. Subgroups

Two subgroups were reported in the scope, which were ‘People with diabetes having routine annual checks in a primary care setting; [and] people with diabetes having more frequent checks in a secondary care setting’. These subgroups represented patients at different stages on the care pathway rather than clinically distinct cohorts. However, the description of these subgroups suggests the cohorts would have the characteristics described in Table 2.2.

Table 2.2. *Probable characteristics of subgroups described in the scope.*

Cohort	**Most likely comparator in practice**	**Relative prevalence of DPN**	**Risk of ulcer formation**
Annual checks primary care	10g monofilament	Low	Low
Frequent checks secondary care	10g monofilament 128 Hz tuning fork	High, or already diagnosed	Increased High

2.3.7. Special considerations, including issues related to equality

The scope states ‘Diabetic neuropathy is more common with increasing age and males may develop DPN earlier than females, but neuropathic pain causes more morbidity in females than in males. More secondary complications from diabetic neuropathy have been shown to occur in people of Hispanic or African American family origin. People with diabetes are considered as disabled under the Equality Act

2010'. No people with a protected disadvantage were considered to be disproportionately disadvantaged by this device.

The sponsor repeated the equality considerations included in the scope. The EAC is not aware of any other considerations that should be taken into account.

3 Clinical evidence

3.1 Critique of the sponsor's search strategy

3.1.1. Description of sponsor's search strategy

The sponsor provided very little information on the methods used to identify studies.

The description of the search for published studies is given as follows:

> 'Internet search facilities and specialised medical databases are the principal means by which information is gathered. Occasionally, an investigator will advise of planned publication of a study'.

The description of the search for unpublished studies is given as follows:

> 'All sales (or requests for free samples) to research or specialist diabetic foot centres are noted and personally followed up. Proceedings for specialised conferences are screened for potential poster presentations that may in turn lead to study publications'.

No further details are given in section 7.1 or section 10, appendix 1.

3.1.2. Critique of the Sponsor's Search Strategy

The sponsor indicated that a search had been carried out for both published and unpublished studies. These searches were not clearly reported and the description given does not enable methods to be reproduced.

Insufficient information was provided to enable critique of the sponsor's search for published studies. The description of the search was minimal - search sources were not clearly described and details of search terms used were not provided.

The range of sources searched and search approaches used to identify unpublished literature were not stated. There was no indication that other relevant approaches were used, for example contacting topic experts or searching trial registries. Insufficient detail was provided to enable assessment of the choice of conferences, or the search terms used to screen for poster presentations.

3.1.3. EAC's literature search

The EAC conducted a search designed to retrieve primary literature on the use of VibraTip to detect diabetic peripheral neuropathy in people with type 1

or 2 diabetes, to identify the studies included by the sponsor and identify any evidence the sponsor may have missed.

Searches were limited to results published from 2007 because it is documented that the inventor of VibraTip, Professor Andy Levy, had his initial inspiration for the device in late 2006 (43). Searches were restricted to English language studies only. Translating foreign language studies was not possible within project resources. Animal studies were also excluded.

The selection of sources searched by the EAC included those recommended as the minimum within the NICE Sponsor submission template. In addition, the EAC searched three trial registers and the most relevant conferences and websites.

The strategy used in Ovid MEDLINE and MEDLINE is reported in Figure 3.1. Full details of all the search strategies used by the EAC search are provided in Appendix 1. The searches retrieved 1034 records in total. Following deduplication, 735 records were screened for relevance.

Figure 3.1. *EAC search strategy for Ovid MEDLINE and MEDLINE In-Process.*

1	exp Diabetes Mellitus/307249
2	diabet$.ti,ab,kf,jw. 427229
3	(dpn or iddm or dm1 or mody or dm2 or niddm or iidm or dm).ti,ab,kf. 42824
4	or/1-3 488471
5	Vibration/ 19511
6	(vibrat$ or vibrotactil$ or vibro-tactil$).ti,ab,kf. 47725
7	or/5-6 53175
8	4 and 7 983
9	(vibratip$ or vibra tip$).ti,ab,kf. 2
10	8 or 9 983
11	exp animals/ not humans/ 3898900
12	10 not 11 968
13	limit 12 to (english language and yr="2007 -Current") 289
14	remove duplicates from 13 288

Key to Ovid symbols and commands:

$	truncation symbol
ti,ab,kf,jw	searches are restricted to the Title, Abstract, Keyword Heading Word and Journal Word fields
/	searches are restricted to the subject heading field
exp	the subject heading is exploded
or/1-3	combine sets 1 to 3 using OR

3.2 Critique of the sponsor's study selection

The sponsor did not stipulate specific inclusion and exclusion criteria for the selection of studies. Instead, in sections 7.2.2 and sections 7.2.4, the sponsor stated 'Due to the very limited number of studies involving this device, we believe it is appropriate to include all available information'. The sponsor did not complete Table B1 (Selection criteria used for published studies) or Table B2. The sponsor did not follow the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) methodology to report on the studies identified and justify any inclusions and exclusions (44).

The EAC considered that, due to the paucity of available evidence, it was reasonable for the sponsor to maximise sensitivity by initially including all studies that referred to the intervention, VibraTip, and consider their relevance to the decision problem at a later stage. However, the rationale for this approach, together with explanatory text or tables explaining the PICO analysis (Population, Intervention, Comparator, Outcomes) should have been better described.

3.3 Included and excluded studies

3.3.1. Sponsor's included and excluded studies

The sponsor provided a list of nine studies that were considered to be relevant to the decision problem. These studies were listed numerically from 1 to 9, with the first seven being described as published and studies 8 and 9 being described as unpublished. Key information pertaining to studies 1 to 6 is reported in submission Table B3 and key information pertaining to the unpublished studies 8 to 9 is reported in submission Table B4. The sponsor did not include study 7 in either of these tables because they stated it should be excluded. This study, by Horsfield and Levy (2013), was a laboratory-based technical paper that focussed on the battery life and consistency of output of VibraTip (4). Thus, from the 9 studies identified, the sponsor included 8 studies which they judged to be directly relevant to the decision problem.

3.3.2. EAC's included and excluded studies

From the EAC's literature search (section 3.1.4), the EAC filtered retrieved studies according to the criteria described in Table 3.1. This filtering stage was carried out in duplicate by two researchers (IW and HC), and aimed to include all studies published in the previous 7 years that specifically mentioned VibraTip, or matched a description of this device. Filtering was

performed first through reading the study title, and then through reading abstracts where necessary. If the study could not be excluded from examination of the abstract, the full text article was acquired. The EAC's filtering and selection results are illustrated as a PRISMA flowchart in Figure 3.2.

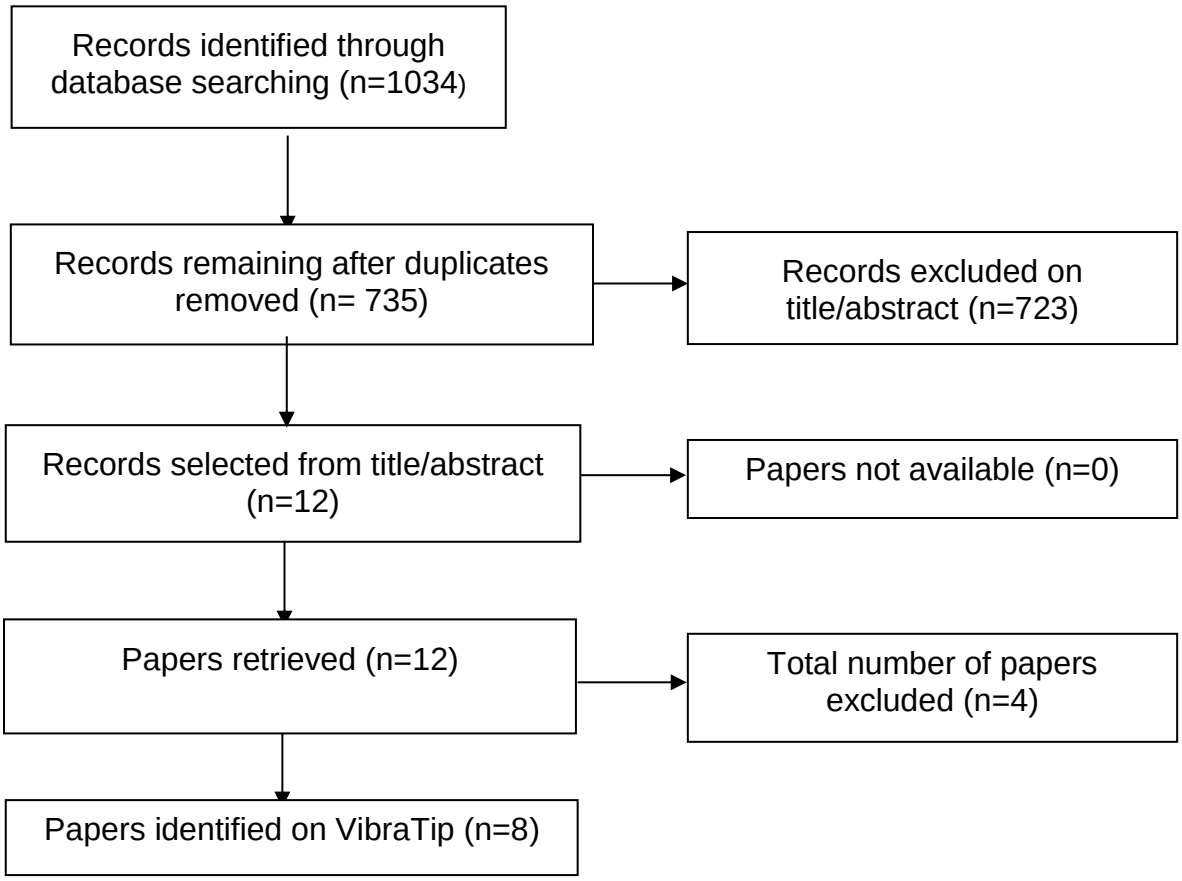
Table 3.1. *EAC Inclusion and exclusion criteria for selection of VibraTip studies.*

Criteria for selection of literature using EAC search strategy.	
Inclusion criteria	
Population	People with diabetes being tested for diabetic peripheral neuropathy (DPN).
Interventions	Vibratip
Comparator	Not specified to maximize sensitivity
Outcomes	Not specified to maximize sensitivity
Study design	Not specified to maximize sensitivity
Language restrictions	English only
Search dates	2007 onwards
Exclusion criteria	
Population	People being tested for vibration perception for any other clinical indication than diabetes, DPN screening, and assessment of risk of ulcer formation.
Interventions	Tuning forks; biothesiometers; neurothesiometers; any pressure / pain perception devices which do not utilize vibration perception.
Outcomes	None
Study design	None

Through independent searching and filtering, the EAC retrieved all the sponsor's included studies that were described as published, except for the study by Nizar *et al.* (2014) (8), which was 'in press' at the time the EAC performed its literature search (but is now published in the *British Journal of Diabetes & Vascular Disease*). The EAC did not identify the unpublished studies described by the sponsor, but were provided with full papers directly by the sponsor. The EAC considered that the paper excluded by the sponsor (Horsfield and Levy 2013) (4), whilst not providing direct comparative clinical evidence, nonetheless provided useful technical information concerning VibraTip.

The EAC's own literature search and filtering process is described using the PRISMA framework in Figure 3.2. The EAC identified two papers in this process that were not identified by the sponsor.

Figure 3.2. *PRISMA flow chart illustrating the EAC's inclusion and exclusion procedures from literature searching stage.*



In summary, the EAC identified and retrieved 12 full papers from the literature searching in total. Of these, 4 were excluded after reading the full paper as they were not relevant to VibraTip. This left 8 papers identified from the literature search that were about VibraTip. Three unpublished papers were supplied to the EAC from the sponsor, making a total of 11 papers where VibraTip was the principal subject. These papers and studies are reported in Table 3.2.

Table 3.2. *Identified studies by study type and in scope. Papers considered not directly relevant to the scope are shaded in grey.*

Author (year)	**Study type**	**In scope?**
Levy (2010) (6)	Diagnostic accuracy	In scope, included.
Bracewell *et al.* (2011) (11)	Diagnostic accuracy	Conference abstract of Bracewell et al. (2012), not discussed further.
Bowling *et al.* (2012) (1)	Diagnostic accuracy	In scope, included.
Bracewell *et al.* (2012) (2)	Diagnostic accuracy	In scope, included.
Baker (2012) (10)	Commentary	No new primary data, not discussed further.
Urbancic-Rovan (2012) (9)	Diagnostic accuracy	In scope, included.
Garbas *et al.* (2013) (3)	Diagnostic accuracy	In scope, included.
Horsfield and Levy (2013) (4)	Technical paper	Not in scope, but useful practical information
Nizar *et al.* (2014) (8)	Diagnostic accuracy	'In press'. In scope, included.
Levy and Greenwood (2014) (7)	Randomised controlled trial	Planned to be published. Not in scope
Horsfield and Levy (2014) (5)	Technical paper	Planned to be published. Not in scope, but useful information.

Studies included by the EAC

In total, six diagnostic accuracy studies were identified with unique patient data that were considered to be relevant to the scope of the decision problem. These studies formed the basis of the clinical effectiveness evidence submission for VibraTip, and were the only studies the EAC considered were directly relevant to the scope of the decision problem. Four of these studies were based in the UK (1, 2, 6, 8) and two were based in Slovenia (3, 9).

Studies excluded by the EAC

The EAC excluded 5 of the 11 papers that were retrieved. These were for the following reasons:

- The study by Bracewell *et al.* (2011) (11) was a conference abstract published in *Minutes of the 46th General Assembly of the European Association for the Study of Diabetes*, that reported on a diagnostic accuracy study of VibraTip. Upon retrieval, it was clear this study was the same as that reported in full by Bracewell *et al.* (2012) (2), therefore it was excluded in favour of the fully published study.

- The paper by Baker (2012) (10) was a commentary on two then recently published diagnostic accuracy studies; that of Bracewell *et al.* (2011) (11) and Bowling *et al.* (2012) (1). As it contained no original diagnostic data, it was not considered further.

- The studies by Horsfield and Levy (2013) (4) and Horsfield and Levy (2014, unpublished) (5) were technical papers. These papers did not fulfil the criteria of the scope in terms of population, comparators, or outcomes. However, the papers were of relevance to some important practical aspects of the assessment overall, and for these reasons the studies have been evaluated further in Appendix 2.

- The study by Levy and Greenwood (2014, unpublished) (7) was the only RCT identified. In this trial, patients with diabetic retinopathy (n=60) were randomised to take home either an educational package, or an enhanced educational package including VibraTip. Quality of Life (QoL) scores (measured using Euroqol 5-dimensions [EQ5D]) and perceptions of self-care ability were recorded at baseline and at 6 months. Whilst the EAC regarded this as a potentially interesting study, it was excluded on the grounds it did not match the scope in the population domain. The population described in the scope states ‘People (adults and children) with type 1 or 2 diabetes undergoing

routine foot-care checks by *healthcare workers* in primary and secondary care settings'. As the intervention was used by *patients* in their homes (self-measurement), it was judged this study should be excluded. Additionally, the presence of diabetic retinopathy in these patients suggests they would have been diabetic for a long time and therefore the risk of DPN would have been particularly high.

3.4 Overview of methodologies of all included studies

3.4.1 Sponsor's description of included studies

The sponsor identified a total of 8 papers that it considered were directly relevant to the scope of the decision problem, and provided more detailed information on 6 of these. The sponsor did not describe the 2 technical papers (4, 5) in further detail. The sponsor gave details of the RCT (7), considered out of scope by the EAC, in submission Table B5. This table was an accurate summary of the full study provided to the EAC. The sponsor gave details of the six diagnostic accuracy studies, considered to be in scope by the EAC, in Table B6. The EAC considered the sponsor had provided an adequate description of these studies.

3.4.2 EAC's description of included studies

The EAC has summarised the main characteristics of the 6 diagnostic accuracy studies in Table 3.3.

Table 3.3. *Characteristics of 6 diagnostic accuracy studies which addressed the scope of the decision problem.*

Study author	**Study Design**	**Patients and setting**	**Index test(s)**	**Reference test(s)**	**Target condition**
Levy (2010) (6)	Cross sectional 'agreement' study	Patients with diabetes undergoing review in hospital or podiatry clinic (n=100).	VibraTip 10g monofilament 128 Hz Tuning fork	None	N/A
Bowling *et al.* (2012) (1)	Cross sectional diagnostic accuracy study (Additionally measurement of 'intra-rater reliability)	Patients with peripheral diabetic neuropathy (varying severity) from community and hospital settings (n=83)	VibraTip (on hallux only) Ipswich touch test (45)	Neurothesiometer (≥25 V threshold) (NDS*, threshold ≥6)	'At risk' neuropathic feet.
Bracewell *et al.* (2012) (2)	Cross sectional diagnostic accuracy study	Patients with diabetes type 1 and 2 in secondary care (n=141)†	VibraTip NeuroTip 10g monofilament	Neurothesiometer (threshold ≥ 25 V)	Peripheral sensory neuropathy

Study author	Study Design	Patients and setting	Index test(s)	Reference test(s)	Target condition
			128 Hz tuning fork (hallux and medial malleolus only) Each performed in 5 sites on both feet		
Urbancic-Rovan *et al.* (2012) (9). Conference abstract.	Cross sectional 'agreement' study	Patients with diabetes (n=42)	VibraTip 10g monofilament 128 Hz tuning fork Tip Therm (46) Neuropad (47)	None	Diabetic sensory neuropathy
Garbas *et al.* (2013) (3). Conference poster.	Cross sectional 'agreement' study	Patients with diabetes mellitus (n=496)	VibraTip 128 Hz tuning fork	None	Described as 'sensory neuropathy' and 'vibration sensation impaired'
Nizar *et al.* (2014) (8)	Cross sectional diagnostic accuracy	Patients with type 1 and 2 diabetes	VibraTip	Neurothesiometer (20	Diabetic peripheral

Study author	Study Design	Patients and setting	Index test(s)	Reference test(s)	Target condition
	study (diagnostic case control study)	recruited from specialist diabetes clinic (n=100).	Tuning fork (oscillation frequency not specified)	V threshold)	neuropathy

*NDS: Neuropathy Disability Score (48).

† Plus 18 patients for inter-rater reliability study of whom 72% had active or previous ulceration. It is unclear if these patients were included in the diagnostic accuracy study.

3.5 Overview and critique of the sponsor's critical appraisal

3.5.1. Sponsor's critical appraisal of diagnostic accuracy studies

The sponsor did not perform a critical appraisal of the included diagnostic accuracy studies. In submission Table B8, the sponsor stated: 'All the studies in the various centres in Bristol, London, Manchester etc. were cross sectional. There were no confounding factors and the outcomes are essentially unbiased and generated by different expert centres who designed their own studies.' The sponsor then stated 'Not clear' or 'N/A' to all the questions set out in the CASP (Critical Appraisal Skills Programme) checklist. The wording the sponsor used in this section suggests their interpretation of 'bias' meant there was no commercial conflict of interests by the groups testing VibraTip, rather than the connotation of bias usually used in evidence based medicine meaning one-sided systematic error (49).

The EAC noted that the checklist reported by the sponsor was designed for cohort studies, which is not appropriate for diagnostic studies. Hence the question prompts were not always applicable (for instance 'Was the exposure accurately measured to minimise bias?'). The EAC notes that CASP does supply a checklist which is specific to diagnostic studies which could have been used instead (see www.caspinternational.org).

3.5.2. EAC's critical appraisal of diagnostic accuracy studies

Because the sponsor did not critically appraise the diagnostic accuracy studies, the EAC did so. Diagnostic accuracy studies of VibraTip were assessed for bias and applicability to the decision problem using the QUADAS-2 tool (revised Quality Assessment of Diagnostic Accuracy Studies) (50, 51). The QUADAS-2 tool assesses the risk of study bias (internal validity) in four domains (patient selection, index test, reference standard, flow and timing), and the applicability of the study to the decision problem (external validity or generalisability) in three domains (patient selection, index test, and reference standard). Domains are classed as low (risk of bias or applicability), high, or unclear, and no attempt is made to formally grade the strength of evidence the study provides. The QUADAS-2 tool is recommended for critically appraising diagnostic accuracy studies in both the development of NICE clinical guidelines (52) and NICE diagnostic assessments (53).

The sponsor provided 6 diagnostic studies to support the clinical effectiveness of VibraTip in the detection of DPN in patients with diabetes. Of these, 3 studies were true diagnostic accuracy studies (1, 2, 8), whilst the other 3 lacked a reference standard so could not provide true diagnostic accuracy information (3, 6, 9). This limited the usefulness of these studies in informing the decision problem.

On appraisal, all the studies were at high risk of bias in at least one domain. The results of the QUADAS-2 critical appraisal are reported in Table 3.4 and Table 3.5, and displayed graphically in Figure 3.3. A pictorial summary of results is reported in Table 3.6.

Table 3.4. *Critical Appraisal of internal validity of diagnostic accuracy studies for VibraTip, developed from QUADAS 2.*

Study	**Domain and risk of bias**			
	Patient selection	**Index test**	**Reference standard**	**Flow and timing**
Levy (2010) (6)	Patients with diabetes recruited consecutively, but patient characteristics not described. **HIGH RISK**	Order of VibraTip, 10g monofilament, and tuning fork randomised, outcomes interpreted by operator. Five sites tested but not reported what threshold for positive test was. **HIGH RISK**	No reference test or target condition reported. Outcomes consisted only of agreement between index tests. **HIGH RISK**	All tests performed immediately sequentially in all recruited patients. **LOW RISK**
Bowling *et al.* (2012) (1)	Limited information on patient characteristics, recruitment probably not consecutive. **HIGH RISK**	VibraTip (one site) and Ipswich touch test interpreted by operator with possible knowledge of reference test. No claim for randomisation.	NDS and neurothesiometer performed by operator with possible knowledge of index test. Pre-specified thresholds described.	All tests performed immediately sequentially in all recruited patients. **LOW RISK**

Study	Domain and risk of bias			
	Patient selection	**Index test**	**Reference standard**	**Flow and timing**
		HIGH RISK	Randomisation not stated. **HIGH RISK**	
Bracewell *et al.* (2012) (2)	Patients with diabetes from secondary care, but recruitment strategy not described (probably not consecutive) and patient characteristics not clearly described. **HIGH RISK**	VibraTip, Neurotip, 10g monofilament and tuning fork performed in random order by operator. Thresholds derived from ROC analysis (on same dataset). **HIGH RISK**	Neurothesiometer (≥25 V) performed by operator. Order of tests randomised. **LOW RISK**	All tests performed immediately sequentially in all recruited patients. **LOW RISK**

Study	Domain and risk of bias			
	Patient selection	**Index test**	**Reference standard**	**Flow and timing**
Urbancic-Rovan *et al.* (2012) (9), Conference abstract.	Patients with diabetes. Setting and recruitment method not described. **HIGH RISK**	VibraTip, 10g monofilament, 128 Hz tuning fork, Tip Therm, Neuropad performed but order not specified and no reference test described. **HIGH RISK**	No reference test. Outcomes restricted to reporting of 'normal test'. **HIGH RISK**	Tests presumed to be performed sequentially in one episode but reporting not adequate to establish this with certainty. **UNCLEAR RISK**
Garbas *et al.* (2013) (3). Conference poster.	Patients with diabetes mellitus. Patient characteristics and clinical setting not described. **HIGH RISK**	VibraTip used as per product instrunctions. 128 Hz tuning fork used once inter-phalangeal joint of big toe, use of damped tuning fork control test not described. Order not described. No reference test described. **HIGH RISK**	No reference test. Outcomes restricted to reporting of test positive for 'vibration sensation impaired'. **HIGH RISK**	Index tests performed immediately sequentially in all recruited patients. **LOW RISK**

Study	Domain and risk of bias			
	Patient selection	**Index test**	**Reference standard**	**Flow and timing**
Nizar *et al.* (2014) (8)	Patients with type 1 and 2 diabetes from specialist diabetes clinic. Patients allocated as cases or controls (50% diagnosed with diabetic peripheral neuropathy), but recruitment strategy not defined and no matching attempted. Patient characteristics not equal at baseline. **HIGH RISK**	Order of VibraTip and tuning fork not randomised. Disease status of patient known. Frequency of tuning fork not reported. **HIGH RISK**	Neurothesiometer performed prior to index tests. Threshold of 20 V used, which differs from other studies using neurothesiometer. **HIGH RISK**	Index tests performed immediately sequentially in all recruited patients. **LOW RISK**

Table 3.5. *Critical appraisal of external validity (applicability or generalisability) of diagnostic studies for VibraTip, developed from QUADAS-2.*

Study	Domain and applicability concerns			
	Patient selection	**Index test**	**Reference standard**	**Comment**
Levy (2010) (6)	Patients with diabetes were attending annual reviews but this included podiatry clinics and these may have been at high risk of diabetic peripheral neuropathy (spectrum bias). **UNCLEAR RISK**	VibraTip used on 5 sites but unclear how many insensate sites required for a positive result, and how this accords with standard practice. **HIGH RISK**	No reference standard was used. **HIGH RISK**	As no reference test stated, diagnostic accuracy data not available.
Bowling *et al.* (2012) (1)	All patients had pre-existing DPN of varying severity. Possibly out of the 'main' scope. **UNCLEAR RISK**	VibraTip used on one site, which seems to be consistent with instructions for use. **LOW RISK**	The target condition is 'at risk' neuropathic feet, which matches the subgroup analysis of the scope only. **UNCLEAR RISK**	Both patient selection and the target condition diagnosed by the reference standard are in scope for subgroup analysis only.

Study	Domain and applicability concerns			
	Patient selection	**Index test**	**Reference standard**	**Comment**
				No direct comparison with 10g monofilament, the most appropriate comparator.
Bracewell *et al.* (2012) (2)	Patients from secondary care diabetes service, but otherwise poorly described. **UNCLEAR RISK**	Following ROC analysis, two insensate sites used for positive diabetic neuropathy test. Currently not standard practice. **HIGH RISK**	Neurothesiometer 'gold standard' for detection of diabetic neuropathy (the target condition). **LOW RISK**	Diagnostic accuracy results (using 2/10 insensate sites) clearly presented, including data from 10g monofilament.
Urbancic-Rovan *et al.* (2012) (9), Conference abstract.	Patient characteristics and setting inadequately described. **UNCLEAR RISK**	No details of how index tests were performed. **HIGH RISK**	No reference test described. **HIGH RISK**	As no reference test stated, diagnostic accuracy data not available.
Garbas *et al.* (2013) (3).	Patient characteristics and setting inadequately	Tests performed once on each foot, unclear	No reference test	As no reference test stated, diagnostic

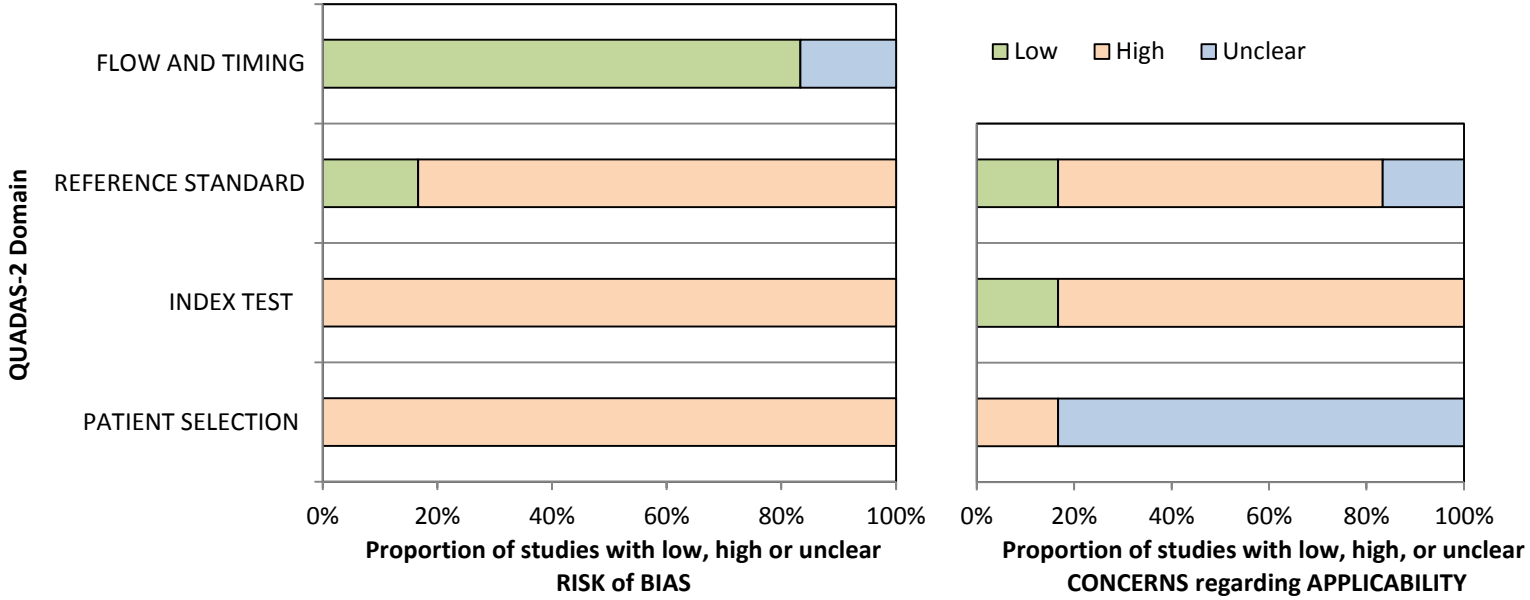
Study	Domain and applicability concerns			
	Patient selection	**Index test**	**Reference standard**	**Comment**
Conference poster.	described. **UNCLEAR RISK**	techniques and sites match usual UK practice. **HIGH RISK**	described. **HIGH RISK**	accuracy data not available.
Nizar *et al.* (2014) (8)	Patient characteristics well described but patient groups (with and without DPN) not equivalent. Case control design. **HIGH RISK**	Index tests applied once only to medial malleoli (where possible). Not current standard practice. Frequency of tuning fork not stated. **HIGH RISK**	Diabetic peripheral neuropathy diagnosed by gold standard neurothesiometer, but threshold set low at 20 V. **UNCLEAR RISK**	Case control study at greater risk of bias than other diagnostic accuracy studies (54, 55). Values for sensitivity of tuning fork (40%) appear to be inconsistent with other studies and also inconsistent with VibraTip results in same patients.

Table 3.6. *Pictogram summary of critical appraisal using QUADAS-2 (bias and applicability domains).*

Study	Risk of bias				Applicability concerns		
	Patient selection	Index test	Reference standard	Flow and timing	Patient selection	Index test	Reference standard
Levy (2010) (6)	☹	☹	☹	☺	?	☹	☹
Bowling *et al.* (2012) (1)	☹	☹	☹	☺	?	☺	?
Bracewell *et al.* (2012) (2)	☹	☹	☺	☺	?	☹	☺
Urbancic-Rovan *et al.* (2012) (9).	☹	☹	☹	?	?	☹	☹
Garbas *et al.* (2013) (3).	☹	☹	☹	☺	?	☹	☹
Nizar *et al.* (2014)(8)	☹	☹	☹	☺	☹	☹	☹

Key: ☺Low Risk ☹High Risk ? Unclear Risk

Figure 3.3. *Graphical presentation of study quality and reliability following critical appraisal with QUADAS-2.*



All the studies were considered to be at high risk of bias regarding patient selection, due to the potential for spectrum bias (56). This was due to a combination of poor reporting, poor methodological selection of patients, and selection of particularly high risk patients. If patients are selected that are at high risk of the target condition (i.e. have a high pre-test probability of having DPN), the predictive value of the index test may appear exaggerated (57). For one study, the population recruited had diagnosed DPN, and the target condition was 'at risk' feet, which was of limited applicability to the primary purpose of the scope (2).

All studies were considered at high risk of bias in the index test domain. This was partly due to poor quality reporting of how the index tests were used or poor methodological application of index tests (for instance index test order not randomised), but also due to design limitations, whereby it was likely that the operator had knowledge of the comparator or reference standard test results. In one study where this was not true because the diagnostic tests under study were randomised, the threshold for a positive result was not pre-specified (2).

The three studies that did not include a reference standard automatically scored for high risk of bias in this domain (3, 6, 9); however two of the remaining studies were also considered to be at risk of bias as, although they used an appropriate reference standard (neurothesiometer), it was likely the operator had prior knowledge of the index tests (1, 8). In one study, the order of tests was randomised and this was considered to be at low risk of bias in this domain (2). All the studies were considered to be at low risk in flow and timing domain due to the cross sectional nature of these studies.

The QUADAS-2 tool also assesses the external validity of diagnostic studies; that is, their applicability to the decision problem. Again, all the studies had at least one issue with external validity (see Figure 3.2). For patient selection, the generally poor reporting meant there was considerable uncertainty on whether the studies were directly relevant to the NHS, and this was further aggravated by the potential for patients to have been 'cherry picked' due to poor trial methodology (for instance non-consecutive recruitment). For the index test domain, it was unclear if the tests were performed as they are in standard clinical practice, in particular concerning the number of sites on the foot tested, and the number of insensate tests required for a positive diagnosis of DPN (see section 2.3.2). In one study it was found through receiver operator curve (ROC) analysis that 2/10 insensate sites gave the best diagnostic performance for VibraTip (2), however this technique is not reflected in the product literature (32). There were similar issues with the reference standard generalisability (i.e. is the correct target condition being

accurately diagnosed), with one study using an inappropriate target condition (2) and three studies not diagnosing a target condition at all (3, 6, 9).

A narrative of the study quality and risk of bias for each study is given below.

Levy 2010

The study by Levy (2010) was a cross sectional study that compared three diagnostic devices in patients with diabetes undergoing annual review in a hospital or podiatry clinic (6). The three devices included in the study were VibraTip, 10g monofilament, and 128 Hz tuning fork. However, as no reference standard was employed, this cannot be considered to be a diagnostic accuracy study and no diagnostic accuracy data could be drawn upon. The study was poorly described and because of this exhibited a high risk of bias in the domain of patient selection, due to the potential for spectrum bias. The index test domain was considered at high risk of bias because the threshold for positive or negative tests (i.e. proportion required for positive insensate test) was not described. There was no reference test; however the study was at low risk of bias in the flow and timing domain, due to the near simultaneous use of the index tests.

In terms of external validity, the study exhibited overall uncertain or high risk of bias, mainly due to poor reporting and uncertainty of the study conditions, population and conduct of the tests match clinical practice in the NHS.

Bowling *et al.* 2012

The study by Bowling *et al.* (2012) was a cross sectional diagnostic accuracy study that compared VibraTip or the Ipswich touch test with each of two reference standards; a neurothesiometer (threshold ≥25 V) or Neuropathy Disability Score (NDS), which is a composite outcome derived from pain sensation, vibration sensation, temperature sensation, and ankle reflex (threshold score ≥ 6) (1). The study was considered to be at a high risk of bias in the domain of patient selection because of the limited description of the patient characteristics and because recruitment did not appear to be randomised or consecutive. The study was considered to be at high risk of bias for the domains of index test and reference standard, as although thresholds were clearly described, the operator probably had prior knowledge of the index or reference test results (lack of randomisation). Additionally, the use of a neurothesiometer alone is not sufficient to detect the target condition (‘at risk’ feet). There was a low risk of bias in the flow and timing domain, as the studies were conducted at the same time. Clear instructions were provided to the eight healthcare professionals delivering the tests and to patients.

The study was not completely generalisable to the decision problem, primarily because the population studied had DPN of varying severity already diagnosed, and were therefore only representative of some of the population described in section ‘subgroups to be considered’ of the scope (‘people with diabetes having more frequent checks in a secondary care setting’). This meant both the patient selection and reference standard domains were at uncertain risk of bias. The index test domain for VibraTip was considered to be at low risk of bias because description of its use matched those stated in the product literature.

Bracewell *et al*. 2012

The study by Bracewell *et al.* (2012) was a cross sectional diagnostic accuracy study that compared four index tests with a neurothesiometer as a reference standard (2). It also attempted to establish the number of insensate sites that gave the best accuracy for each index test.

The index tests were VibraTip, Neurotip, 10g monofilament and 128 Hz tuning fork, and the population was patients with diabetes type 1 or 2 recruited from a secondary care setting. This study exhibited a high risk of bias for patient selection, as it was not clear how patients were recruited and the description of patient characteristics was poor. The study was also considered at high risk of bias for the domain of index test, because thresholds were derived from ROC analysis (using the same dataset). For the reference standard, the threshold voltage of the neurothesiometer was described as ≥25 V (the standard reference amplitude), and the study was considered to be at low risk in this domain. As all tests were performed in one episode, the study was at low risk of bias in flow and timing.

Regarding generalisability of the study, patient selection was poorly described, although as patients were from secondary care settings they were likely to be at greater risk for diabetic peripheral neuropathy (prevalence of neuropathy was measured at 41%). There was a high risk that the index test used was not generalisable as thresholds were calculated from ROC analysis and probably not representative of standard practice. Finally, the reference standard used, the neurothesiometer, was applicable to the decision question (low risk of bias).

Urbancic-Rovan *et al.* 2012

The study by Urbancic-Rovan *et al.* (2012) was described as a pilot study and was reported as a conference abstract which meant study details were lacking (9). This study compared five reference tests in 42 patients with diabetes: these reference tests were VibraTip, 128 Hz tuning fork, 10g monofilament,

Tip Therm (which detects impaired skin temperature sensation) (46) and Neuropad (described as a simple and cheap diagnostic tool for the evaluation of sweat gland function) (47). The patient population and setting were poorly described and at high risk of selection bias. Similarly, the use of index tests, (described as; 'as stated in the instructions in the product package') and randomisation procedures, was poorly reported meaning the index test domain was at high risk of bias. Although the target condition was correct, there was no reference test or gold standard used, so this domain was also at high risk of bias. This also limits the usefulness of this study. Although the tests were probably performed immediately consecutively, there was insufficient reporting to verify this.

Regarding the generalisability of the study, there was generally insufficient reporting to ascertain whether these study results were applicable to UK NHS practice or consistent with CG10. However, the lack of a reference standard meant that there was no diagnostic accuracy data with which to make comparisons with previously published studies (high risk of bias).

Garbas *et al.* 2013

The study by Garbas *et al.* (2013) was a follow-up study to the pilot performed by Urbancic-Rovan *et al.* (2012). It is reported as a conference poster; therefore details are also lacking (3). For this large study (n=496), only two index tests were investigated (VibraTip and 128 Hz tuning fork) and no reference standard was used. The study setting was not described and details of the patient selection procedure and patient characteristics were lacking, meaning the study was at high risk of selection bias. The description of the index tests, and in particular how they were used, and in which order, was poorly described and at high risk of bias. There was no reference test described meaning this domain was at high risk of bias. Although not clearly described, it was considered that it was likely the tests were performed during one episode, meaning the flow and timing domain was at low risk of bias.

In terms of generalisability to the UK and patient pathways described in CG10, the study reporting was generally inadequate to have confidence it was applicable to the decision problem. The lack of a reference standard meant that diagnostic accuracy data could not be obtained limiting the usefulness of this study.

Nizar *et al*. 2014

The study by Nizar *et al.* (2014) has recently been published in the Journal of Diabetes and Vascular Disease (8). This study was described as a cross sectional diagnostic study; however the patients (with type 1 or type 2

diabetes) had the presence of DPN diagnosed using a neurothesiometer, and were divided into equal groups of DPN positive and DPN negative cohorts (n=50 in each cohort) accordingly. This diagnostic case controlled design (not to be confused with epidemiological case control studies) has been shown to potentially exaggerate the diagnostic accuracy of the index test and may be subject to significant bias (54, 55). The patient population was selected and therefore at risk of selection bias by definition. This was illustrated by significant differences in the baseline characteristics between DPN positive and DPN negative cohorts in terms of age and type of diabetes (twice as many patients with DPN diagnosed had type 1 diabetes compared with those without the condition). Additionally, there was reporting of exclusion of 'malingering' patients in this study; the EAC was unclear what this meant.

The index tests investigated were VibraTip and a tuning fork, but the frequency of this was unstated. Both were applied to the medial malleoli where practical. Use of randomisation in order of application was not reported (high risk of bias). The study was considered to be at high risk of bias for the reference standard, because it is likely the operator was aware of the patients DPN status before the index and comparator tests were performed. The study was considered to be at low risk for the flow and timing domain.

Regarding generalisability, patient selection was adjudged to be compromised by the study design and the index test compromised by the lack of standard technique (high risk of bias). The amplitude threshold of the neurothesiometer was set at 20 V, which is not the threshold usually used when a neurothesiometer is being used as a reference standard for the detection of DPN, which is ≥25 V. Thus there was a high risk of bias in the reference test (target condition) domain.

Summary

A summary of the findings of the critical appraisal performed by the EAC is reported in Table 3.7.

Table 3.7. *Summary of the key quality concerns for each study, and the relevance and usefulness of each study for the decision problem at hand.*

Study	**Interval validity**	**External validity**	**Usefulness to decision problem**
Levy (2010) (6)	Generally poor, particularly for patient selection, index, and reference tests.	Lack of reference standard and poor reporting of methods and results.	Poor.
Bowling *et al.* (2012) (1)	Poor for patient selection, otherwise subject to potential bias in index and reference standard domains.	Patient selection and target condition of limited relevance to scope.	Medium. May match scope for 'subgroups to be considered'.
Bracewell *et al.* (2012) (2)	Poor for patient selection, subject to potential bias in index test domain as prespecified threshold not used.	Use of ROC analysis means caution should be applied when generalising.	High, most relevant and highest quality study on VibraTip.
Urbancic-Rovan *et al.* (2012) (9)	Poor for patient selection, index, and reference tests. Poorly reported.	Lack of reference standard and inadequate reporting of methods and results.	Poor.
Garbas *et al.* (2013) (3).	Poor for patient selection, index, and reference tests. Poorly reported.	Lack of reference standard and inadequate reporting of methods and results.	Poor.

Study	Interval validity	External validity	Usefulness to decision problem
Nizar *et al.* (2014) (8)	Poor for patient selection and description of index tests. Case control design susceptible to bias.	Patient selection and use of index tests high risk of bias and low applicability.	Poor. Inappropriate reference standard. Results reported for tuning fork not generalisable.

3.6 Results

3.6.1 Critique of sponsor's report of results

The sponsor summarised and presented the results for the 6 diagnostic accuracy studies in submission Table B9 (section 7.6). The sponsor did not report on the results of the other studies it considered were suitable for inclusion, namely the RCT (7) and the 2 technical papers (4, 5); the EAC's interpretation of the results from the technical papers are reported in Appendix 2. The EAC confirms the reported results from the diagnostic accuracy studies matched those reported in the original papers, although occasionally results were omitted. However, the sponsor did not attempt to appraise the author's conclusions or place the results in the context of the decision problem.

3.6.2 EAC's report of results

The primary results from the 6 diagnostic accuracy studies supplied by the sponsor are summarised in Table 3.8 and discussed in more detail individually below. Although all the studies provided comparative data, 3 studies did not employ a reference standard and therefore provided data of only limited applicability in the context of the decision problem (3, 6, 9). The comparative diagnostic accuracy data, including 95% confidence intervals, for the studies that included a reference standard are presented in Figure 3.4. The population recruited in 1 study had pre-existing DPN and the target condition was 'at risk feet', which limited the applicability of the study to the problem as set by the scope (1). Two studies, which had non-UK settings, were published in abstract or poster form only, and had a poor quality of reporting that made interpretation of comparative results difficult (3, 9). One study, the second largest reported (n=141), was considered to be of superior quality, have greater applicability, and better reporting of results than the other studies (2).

Table 3.8. *Summary of reported results of diagnostic studies on VibraTip.*

Study	**Index tests**	**Reference test(s)**	**'Agreement'**	**ROC**	**Diagnostic accuracy (of VibraTip)**
Levy (2010) (6)	VibraTip 10g monofilament 128 Hz tuning fork	None	*The same 55/100 patients felt both the tuning fork and the VibraTip and the remaining 45 felt neither (complete agreement). 53 subjects felt the monofilament, of whom 51 felt the VibraTip; 4 of the remaining 47 who did not feel the monofilament felt the VibraTip.	N/A	N/A
Bowling *et al.* (2012)	VibraTip (on hallux	Neurothesiometer	VibraTip vs Ipswich touch test: 100%	Not performed.	Sensitivity VibraTip: Neurothesiometer and NDS

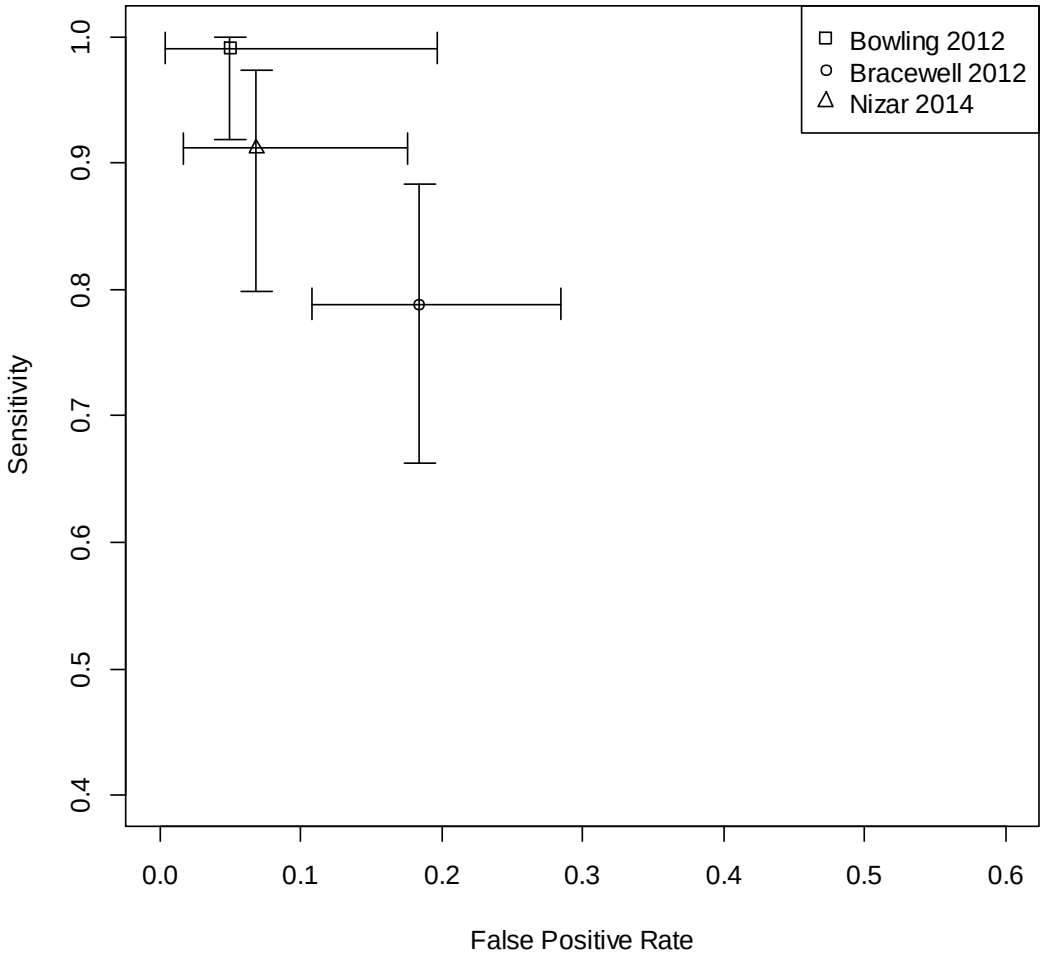
Study	Index tests	Reference test(s)	'Agreement'	ROC	Diagnostic accuracy (of VibraTip)
(1)	only) Ipswich touch test	(≥25 V threshold) NDS (threshold ≥6)	agreement (κ=1.0) VibraTip vs Neurothesiometer: κ=0.973, p<0.001 VibraTip vs NDS: κ=0.921, p<0.001		100%. Specificity VibraTip: Neurothesiometer 96.6%, NDS 90.3% PPV: Neurothesiometer 98.2%, NDS: 94.5% NPV: Neurothesiometer and NDS 100%. PLR: Neurothesiometer 29.0, NDS 10.3 NLR: Neurothesiometer and NDS 0
Bracewell *et al.* (2012) (2)	VibraTip NeuroTip 10g monofilament	Neurothesiometer (≥25 V threshold)	Intra-rater reliability: Cronbach's alpha coefficient 0.88 (n=18)	10g monofilament superior to tuning fork (p=0.0056) and Neurotip	Sensitivity: 0.79 Specificity: 0.82 PPV: 0.75

Study	Index tests	Reference test(s)	'Agreement'	ROC	Diagnostic accuracy (of VibraTip)
	128 Hz tuning fork			(p=0.0022), not superior to VibraTip (p=0.3214) (n=141)	NPV: 0.85 PLR: 4.39 NLR: 0.25
Urbancic-Rovan *et al.* (2012) (9).	VibraTip 10g monofilament 128 Hz tuning fork Tip Therm Neuropad	None	'Normal' results in 47.6% (VibraTip), 45.2% (tuning fork), 57.1% (Tip Therm), and 45.2% (Neuropad), no significant differences (p>0.05). Normal result 14.3% (10g monofilament). Significantly different (p<0.005).	N/A	N/A
Garbas *et al.* (2013) (3).	VibraTip	None	VibraTip (vibration sensation impaired): left foot 32.7%, right	N/A	N/A

Study	Index tests	Reference test(s)	'Agreement'	ROC	Diagnostic accuracy (of VibraTip)
	128 Hz Tuning fork		foot 31.3%. 128 Hz tuning fork: left foot 36.5%, right foot 36.3%. No significant difference in feet. Tuning fork more sensitive (left foot p=0.007, right foot p=0.012).		
Nizar *et al.* (2014)(8)	VibraTip Tuning fork	Neurothesiometer (20 V threshold)	'Difference' in sensitivity of 52% in favour of VibraTip (p<0.001)†. 'Difference' in specificity of 6% in favour of tuning fork (p<0.025) †.	Not performed.	Sensitivity: 0.92 (95% CI 0.808 to 0.978) Specificity: 0.94 (95% CI 0.834 to 0.987) PPV: 0.94 NPV: 0.92 PLR: 15.3

Study	Index tests	Reference test(s)	'Agreement'	ROC	Diagnostic accuracy (of VibraTip)
* Study reported agreement was measured using kappa-statistic and McNemar's test in Study design section; however, no statistical analysis presented in results. † Statistical analysis method not reported.					

Figure 3.4. *Sensitivity and false positive rate (1 – specificity) of Vibratip relative to the neurothesiometer reference standard; results from three studies* (3, 6, 9).



Levy 2010

The study described by Levy (2010) was not a true diagnostic accuracy study because no reference standard was used, so calculation of prevalence, sensitivity and specificity, and other associated diagnostic outcomes was not possible (6). The aim of the study was to measure the level of agreement between VibraTip and two other index tests, namely 128 Hz tuning fork and 10g monofilament; neither of these tests are gold standard tests (40). Instead, the author focused on the level of agreement between the tests. Contingency tables for these comparisons were inferred from the paper by the EAC as follows:

		Tuning fork		Monofilament	
		Felt	**Not felt**	**Felt**	**Not felt**
VibraTip	Felt	55	0	51	4
	Not Felt	0	45	2	43

(It was assumed that the labels of the first two bars in Figure 4 of the paper were transposed; otherwise the text and figure are not consistent). The number of patients perceiving pressure from sham 10g monofilament was not reported. Additionally, although statistical testing using kappa-statistic and McNemar's test were described in the methods, none were reported. For both comparisons, McNemar's test applied to these contingency tables by the EAC showed no significant differences between the techniques.

The data reported by Levy is of limited utility as it did not include comparison with a recognised reference standard and therefore analysis of diagnostic accuracy is not possible. Although some agreement data was reported, this is of limited value in the context of the decision problem. However, the superiority, equivalence, or otherwise of VibraTip over 10g monofilament could not be established with confidence from the reported data. In addition, the results that are presented are incomplete and lack a measure of significance.

Bowling *et al.* 2012

The study by Bowling *et al.* (2012) compared the diagnostic accuracy of two index tests (VibraTip and Ipswich touch test) with two reference standards (neurothesiometer [≥25 V] and NDS [score ≥ 6]) (1). All the patients recruited had pre-diagnosed DPN of varying severity and the target 'condition' was feet being 'at risk' of ulceration. The results, analysed using Cohen's kappa, showed that both index tests had perfect agreement with each other and had 'substantial agreement' with the reference standards.

The EAC inferred the contingency table for agreement between the Vibratip and neurothesiometer from the information provided in the paper, as follows:

		Neurothesiometer (25V)	
		Test +ve	**Test -ve**
VibraTip	**Test +ve**	54	1
	Test -ve	0	28

Where “Test +ve” means patients with “at risk” feet; i.e. those unable to feel the stimuli. From this table, the EAC calculated that relative to the neurothesiometer, the sensitivity of the Vibratip was 1.0 (95% confidence interval 0.934 to 1.000) and its sensitivity was 0.966 (0.822 to 0.999). The values for positive predictive value and negative predictive value quoted by the authors assumed the disease prevalence was the same as for their sample (i.e. 54/83 = 65%); the prevalence is likely to be lower in the population described in the scope and the quoted values for PPV and NPV may not be applicable (57).

In comparison with the NDS (not a reference standard), the EAC inferred the contingency table as follows:

		NDS (≥ 6)	
		Y	**N**
VibraTip	**Test +ve**	52	3
	Test -ve	0	28

The EAC applied McNemar’s test to this table and found there to be no significant differences between the two tests.

The results from the Bowling study showed that VibraTip is a useful test for predicting ‘at risk’ feet according to pre-specified diagnostic criteria using either a neurothesiometer (an unsuitable reference standard) or a clinical prediction rule. However, the population recruited was already diagnosed with DPN, and the target condition (‘at risk’ feet) was not that specified by the scope, which was people with diabetes. Thus the usefulness of this study in informing the decision problem was limited.

Bracewell *et al.* 2012

The study by Bracewell *et al.* (2012) compared four index tests (VibraTip, 10g monofilament, 128 Hz tuning fork, Neurotip) with a reference test (neurothesiometer, threshold ≥ 25 V) (2). This was the second largest of the studies on VibraTip, with a reported prevalence of diabetic peripheral neuropathy in this population of 41%. The first element of the study reported on the intra-rater reliability of VibraTip, and consisted of a researcher repeating the use of VibraTip on the same patient 2-3 weeks after the initial test (18 patients, unclear if these were subjects in main study). The successive readings resulted in an alpha coefficient (Cronbach’s alpha calculation) of 0.88, which the authors stated indicated that the device had ‘good intra-rater reliability’. No confidence interval was provided and it is not

clear why this test was selected to assess intra-rater repeatability of a test with a dichotomous outcome. It should also be noted that the patients recruited for this part of the study had particularly high risk of having diabetic peripheral neuropathy (with 72.2% having active or previous ulceration), giving rise to the potential for spectrum bias.

The main element of the study concerned the comparative diagnostic accuracy of the four index tests compared with the reference test (neurothesiometer). In total 141 patients were recruited for this part of the study, with 89% reported as having no history of foot ulcers (in contrast to the intra-rater reliability study). The authors performed ROC analysis to find the optimum number of insensate sites which gave the best diagnostic accuracy of peripheral diabetic neuropathy for each test. When comparing the area under the curve (AUC) of the index tests, the authors reported that the 10g monofilament was superior to the tuning fork ($p = 0.0056$) and the NeuroTip ($p = 0.0022$), but not VibraTip ($p = 0.3214$). Actual AUC values and comparisons of VibraTip with the other index tests were not reported and the form of hypothesis test was not described. From ROC analysis, the researchers determined that two or more insensate sites out of ten were optimal to diagnose diabetic peripheral neuropathy for VibraTip, 10g monofilament and Neurotip, whilst one or more sites was optimal for the tuning fork. It was assumed by the EAC these were the thresholds used for calculation of the diagnostic parameters.

For the comparison of Vibratip with the neurothesiometer, the EAC inferred that the contingency table was as follows:

		Neurothesiometer (25V)	
		Test +ve	**Test -ve**
VibraTip	**Test +ve**	46	15
	Test -ve	12	68

The EAC computed sensitivity of the Vibratip as 0.79 (95% CI 0.67 to 0.89) and the specificity as 0.82 (0.72 to 0.90). The values quoted by the authors for PPV and NPV assumed that the disease prevalence was the same as in their study population (58/141 = 41%). Different values would be found in different populations. The EAC derived contingency tables for the comparisons of tuning fork, monofilament and Neurotip, from which confidence intervals of sensitivity and specificity were derived. The EAC was not able to infer contingency tables for comparisons between the index techniques. The diagnostic results for VibraTip, the 10g monofilament, and the 128 Hz tuning fork are reported in Table 3.9 and presented graphically in Figure 3.5

Figure 3.5. *Sensitivity and false positive rate (1 – specificity) of VibraTip, 10g monofilament and the 128 Hz tuning fork (all relative to a neurothesiometer reference test).*

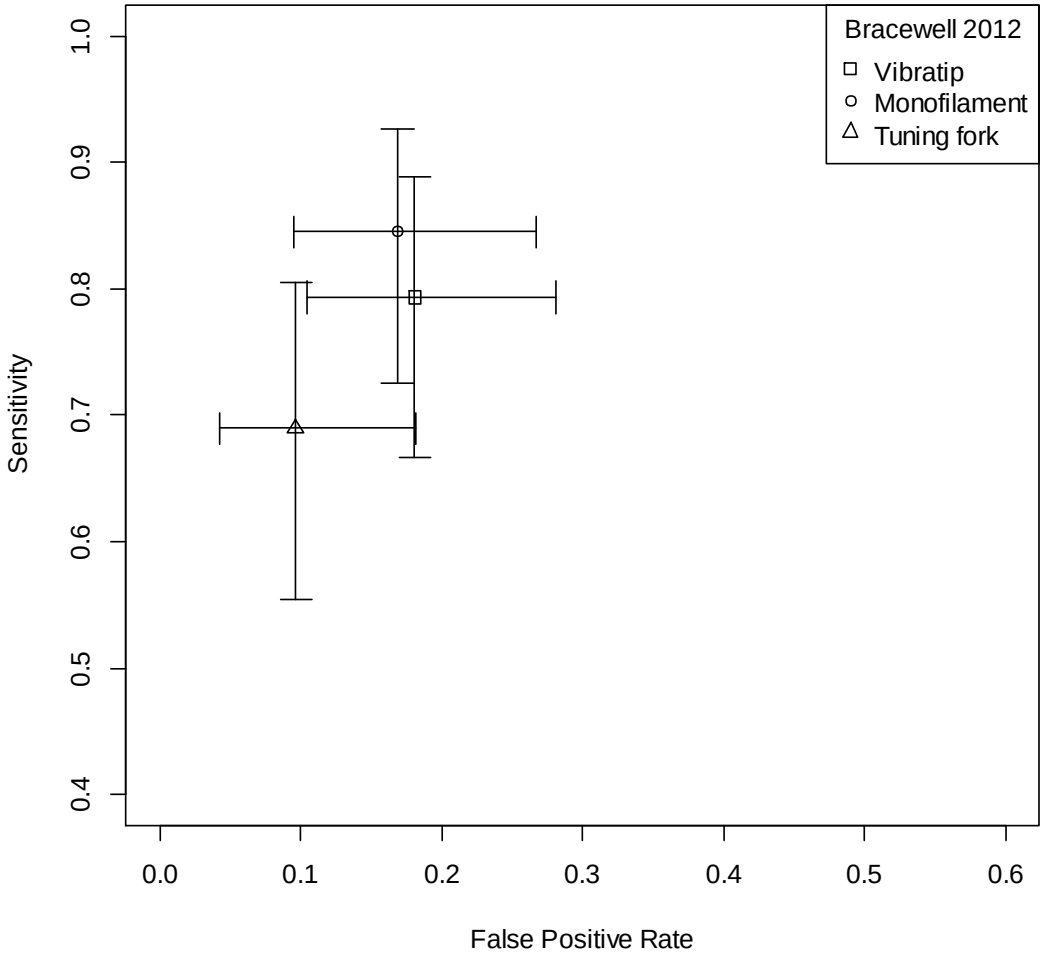


Table 3.9. *Diagnostic accuracy of VibraTip, 10g monofilament, and 128 Hz tuning fork. Data from Bracewell et al. (2012) (40), with 95% confidence intervals calculated by EAC*.

	Sensitivity (95% CI)	**Specificity (95% CI)**
VibraTip	0.79 (0.69 to 0.90)	0.82 (0.74 to 0.90)
10g monofilament	0.84 (0.75 to 0.94)	0.83 (0.75 to 0.91)
128 Hz tuning fork	0.69 (CI 0.57 to 0.81)	0.90 (0.84 to 0.97)

The results from Bracewell *et al.* provide evidence that VibraTip has similar diagnostic accuracy in detecting diabetic peripheral neuropathy as the 10g monofilament, 128 Hz tuning fork, and Neurotip. The study did not provide evidence that VibraTip was superior to any of the other index tests, and poor reporting of baseline characteristics and results may mean that the study might not be fully generalisable to all scenarios described in the scope. Although the intra-rater reliability [agreement] of VibraTip was measured and described as 'good', caution should be applied in the interpretation of this result as the sample was small, with patients at very high risk of neuropathy, and no comparison was made with other devices. Note, intra-rater reliability does not match the outcome described in the scope, which is *inter*-rater reliability (see Table 2.1).

Urbancic-Rovan *et al.* 2012

The study by Urbancic-Rovan *et al.* (2012) was a small pilot study (n=42) reported as a conference abstract (9). Similar to the study by Levy *et al.* (2010) (6), this was not a true diagnostic accuracy study because no reference standard was used, so calculation of diagnostic accuracy parameters was not possible. Instead, the study focused on the proportion of 'normal' index test results as the primary outcome. The index test devices studied were VibraTip, 128 Hz tuning fork, Tip Therm, Neuropad and 10g monofilament.

Whilst no diagnostic accuracy results were available because of the lack of reference standard, the results of this study suggest that the 10g monofilament had a much poorer sensitivity for the detection of insensate sites (positive in only 14.3% tests) compared with the other index tests, including VibraTip (positive in 47.6% of tests). However, it is unclear what the reason for this lack of diagnostic sensitivity might be because it is inconsistent with other studies, which have indicated the 10g monofilament has a much higher sensitivity (40).

Garbas *et al.*2013

The study by Garbas et al (2013) was an extension of the above pilot study by Urbancic-Rovan *et al.* (2012) and recruited a large number of participants (496 patients with diabetes). This study only investigated two index tests, VibraTip and 128 Hz tuning fork, and did not include a reference standard, so diagnostic accuracy parameters of the index tests could not be established. Instead results were expressed as proportion of feet which were sensory impaired, with the left and right feet of patients being independently assessed for both index tests (Table 3.10).

Table 3.10. *Summary of results presented by Garbas et al. (2013).*

	Left foot	**Right foot**	**Significant?**
	People (percentage)		
VibraTip (vibration sensation impaired)	162 (32.7)	155 (31.3)	Not significant†
128 Hz tuning fork (vibration sensation impaired)	181 (36.5)	180 (36.3)	Not significant†
Significant?	p=0.007*	p=0.012*	

* Statistical test used not stated.

† p value not stated.

The results from the Garbas study firstly indicated that there was no significant inter-foot variability regarding sensation testing for the two index tests. Secondly, the study claimed that the tuning fork was significantly more sensitive than VibraTip in detecting impairment of vibration sensation. There are insufficient data in the paper to infer the contingency table for the comparison between techniques, and therefore the main finding cannot be confirmed.

Nizar *et al.* 2014

The study by Nizar *et al.* (2014) was a diagnostic cross sectional study that used a case-control design which meant the researchers had prior knowledge of the patients' DPN status, and recruited them (n=100) accordingly to fix the prevalence of DPN at 50% (8). The index tests investigated were VibraTip and tuning fork on one site of the foot (medial malleoli or alternate pragmatic site if this was impractical). The authors did provide contingency tables for the comparisons of Vibratip and tuning fork with the neurothesiometer. The diagnostic accuracy values for VibraTip and tuning fork are reported in Table 3.11. Note that their estimates for PPV and NPV assume a disease

prevalence of 50%; it is not clear whether this was intentional (i.e. the likely prevalence in a population of interest), or simply a mistake (the sample prevalence was fixed at 50% for the purposes of the study).

Table 3.11. *Summary of results presented by Nizar et al. (2014).*

	VibraTip	**Tuning fork**
Sensitivity	92% (95% CI 80.8% to 97.8%)	40% (95% CI 26.4% to 54.8%)
Specificity	94% (95% CI 83.4% to 98.7%)	100% 95% CI (92.9% to 100.0%)
PPV*	94%	100%
NPV*	92%	63%
PLR	15.3	Infinity
NLR†	0.09	0.60

* Calculated assuming prevalence rate of 50%

† Negative likelihood ratio calculated by EAC for completeness.

The results from the Nizar study suggested that the tuning fork had a very poor sensitivity for the detection of DPN in the population investigated. In fact, more than half the subjects with peripheral neuropathy (30/50) were able to detect the tuning fork, compared with only 4/50 who were able to detect the VibraTip.

This result is unexpected and is not consistent with previously published results (58), including other comparative studies of tuning forks and VibraTip (2, 3, 6, 9). The specificity of the tuning fork test was 100% (i.e. all patients without DPN felt it) and a likely explanation for the very poor sensitivity to the tuning fork test in this study was that some subjects received an audible cue or similar hint that the tuning fork was vibrating, or that the fork had been

activated too strongly. The results suggest a sub-optimal protocol had been followed.

It is not clear why the investigators chose the reference threshold of the neurothesiometer as 20V rather than the standard ≥25 V, thus the ‘cases’ of DPN were not readily comparable with other studies. It has been previously demonstrated that the prediction of ulcer formation is very sensitive to changes in the vibration perception threshold (VPT) between the ranges of 16 to 24 V and >25 V (42). Also, as has been discussed in section 3.5.2, this study was subject to several methodological limitations and major potential sources of bias, including that the frequency of the tuning fork was not reported, and the anatomical site used for the devices (medial malleolus only) was probably not standard practice for DPN screening.

The EAC judges the sensitivity of the tuning fork as reported is outside the range of plausible values, and this and other difficulties with the study cast serious doubt on the validity and generalisability of the results in the context of the decision problem.

3.6.3. Summary of results

The EAC summarises the results as follows:

- Results from the three studies that did not employ a reference standard (3, 6, 9) were of limited use in informing the decision problem.

- Results from Bowling *et al.* (2012) (1) were probably of limited value in addressing the ‘core’ issues of the decision problem due to the population recruited and the target condition.

- Results from the remaining studies by Bracewell *et al.* (2012) (2) and Nizar *et al.* (2014) (8) were conflicting, with the latter study reporting implausibly low sensitivity for the tuning fork. This was probably due to poor study protocol during the application of the tuning fork.

- The studies were too heterogeneous to enable synthesis of data (see section 3.8).

- All the included studies were subject to potential bias, meaning there was considerable uncertainty in their validity and generalisability (see section 3.5).

Of the 2 studies (30, 28) directly relevant to the decision problem and employing a reference standard, the EAC judges Bracewell *et al.* (2012) (2) has the higher quality and is more applicable to the decision problem. This study closely matched the scope, gave comparative data on VibraTip, 10g monofilament, and the 128 Hz tuning fork, and utilized a recognised reference standard (neurothesiometer set at ≥25 V). Despite this, there were concerns over the potential for bias in this study, particularly regarding the use of ROC analysis to optimise the diagnostic technique, which might not reflect real-life practice.

The study by Nizar *et al.* (2014) (8) is seriously compromised by the use of a neurothesiometer set at 20 V rather than ≥25 V. The implications of this are significant:

- The reported target condition, DPN, was not diagnosed by a recognised reference standard. This is a significant concern because it is not known if the 50 patients classified as having DPN would have had DPN using conventional approaches.
- Tuning forks are routinely benchmarked against 25 V, not 20 V, so this is a potentially large source of bias.
- The study did not attempt to quality assure their results by repeating the trial using the standard ≥25 V threshold.

For these reasons, the EAC judges the data presented by the Nizar study are unreliable in the context of the decision problem, and therefore the study by Bracewell *et al.* should be preferred, with the provisos already discussed. This gives VibraTip a sensitivity of 0.79 (95% CI 0.69 to 0.90) and specificity or 0.82 (95% CI 0.74 to 0.90). The confidence intervals overlap those for sensitivity and specificity for a 10g monofilament and 128 Hz tuning fork.

3.7 Description of the adverse events reported by the sponsor

The sponsor reports that there were no adverse events expected or recorded in any evaluations and no adverse events have been reported from current usage.

The EAC agrees that the operation of the VibraTip device as a diagnostic aid, applied gently to the intact skin of a diabetic foot, has negligible risk of adverse events for the patient or operator and therefore no anticipated safety concerns. Nonetheless, the EAC independently searched the Medicines and Healthcare products Regulatory Agency (MHRA) website and US Food and Drug Administration (FDA) Manufacturer and User Facility Device Experience (MAUDE) database and found no instances of or medical device reports for VibraTip, respectively (both accessed 21/03/2014).

3.8 Description and critique of evidence synthesis and meta-analysis carried out by the sponsor

The sponsor did not attempt to synthesise data using meta-analyses, stating 'This would not be appropriate for a series of small observational studies'.

Data from the diagnostic accuracy studies were not synthesized by the EAC in the form of a meta-analysis. This was because there is considerable variation between the studies in terms of study methodology, populations recruited, reference standards used, comparators used, and outcomes reported. Hence the effect size obtained from combining the results from the different studies would be meaningless.

3.9 Additional work carried out by the External Assessment Centre in relation to clinical evidence

The EAC analysed two technical papers that were considered out of scope of the decision problem, but potentially relevant to the assessment overall. This analysis is reported in Appendix 2. These reports addressed the battery life of VibraTip (4), and the frequency and amplitude of VibraTip compared with a 128 Hz tuning fork (5).

Regarding the study by Horsfield and Levy (2013) (4), the EAC had some concerns over the lack of warning given by the VibraTip device when it is nearing the end of its battery life, and considered what might happen in clinical practice as the amplitude of the device decreased due to battery depletion. The implications of the limited battery life of VibraTip and its economic impact is discussed in section 4.5.

The study by Horsfield and Levy (2014, unpublished) (5) was also a laboratory bench test, which involved the design of a manikin toe. This study provided evidence that the frequency of vibration of VibraTip is closer to 200 Hz than the 128 Hz of the standard tuning fork. However, clinical experts contacted by the EAC were in agreement that it is amplitude, rather than frequency, which is most important in human sensation of vibration. The study also showed that VibraTip may create a more reproducible response than the tuning fork when used by non-expert clinicians. However, it is unclear what advantages, if any, this would confer to clinical practice.

3.10 Conclusions on the clinical evidence

VibraTip is a small medical device that is indicated for the detection of DPN as part of the risk stratification algorithm for foot ulceration in people with diabetes. Its place in the patient care pathway, as described by NICE CG10 (18) is clear and unambiguous, and that is as a direct replacement or

alternative for other diagnostic methods used to test for foot sensation during annual or more frequent foot reviews. In practice, VibraTip would replace either the 10g monofilament (typically in primary care) or the 128 Hz tuning fork (typically in secondary care).

3.10.1 Assessment of clinical evidence

The sponsor did not provide a replicable search strategy, but gave the EAC 8 studies that it considered are relevant to the scope. Five of these were published (1-3, 6, 9), 3 were unpublished at time of writing (5, 7, 8), and 1 was a technical study that the sponsor considered did not fit the scope (4). The EAC performed an independent literature search and identified 2 additional papers (10, 11), giving 11 papers specific to the VibraTip intervention.

After consideration of the decision problem using PICO (population, intervention, comparator, outcomes) analysis, the EAC excluded 5 papers as outside scope: 2 were technical papers (4, 5); 1 was an abstract of a full paper (11); 1 was a commentary (10); and 1 was an RCT but was excluded on the population domain (7). This left 6 studies for full consideration. These were all comparative cross sectional diagnostic accuracy studies that included VibraTip and other comparators including the 10g monofilament, 128 Hz tuning fork, Neurotip, Tip Therm, Neuropad, and Ipswich touch test. The studies were heterogeneous in terms of the population, the comparators, the target condition, and outcomes, which meant that data synthesis was not possible.

Of the 6 studies that were in scope, 3 did not provide a reference standard, and hence it was not possible to extract meaningful diagnostic accuracy information from them (3, 6, 9), reducing their usefulness in addressing the issues.

Of the 3 remaining studies, one study, which compared VibraTip with the Ipswich touch test, was conducted in people with pre-diagnosed DPN of varying severity (1). It used a neurothesiometer (threshold ≥25 V) or NDS (threshold ≥6 points) as the reference standard. This study was probably out of scope for patients undergoing annual screening for DPN at foot clinics, but may be relevant for those undergoing more frequent assessment following diagnosis of DPN (a subgroup specified in the scope).

The other 2 studies included patients with diabetes without a prior diagnosis of DPN. One of these had a diagnostic case control design, where 50 people were recruited with DPN, and 50 people without DPN (8). However, the reference standard used to make the diagnosis of DPN was a neurothesiometer set at a threshold of 20 V, rather than the standard ≥25 V,

meaning the ‘cases’ of DPN might not be accurate. Further, the authors found that the tuning fork (details not described) had a sensitivity of 40%, which is worse than random chance. The EAC therefore considered that the results provided from this study were not generalisable.

The remaining study, by Bracewell *et al.* (2012) (2), recruited 141 people with diabetes but with no documented prior diagnosis of DPN. All participants had their feet tested for sensation with VibraTip, NeuroTip, 10g monofilament, and128 Hz tuning fork, and an appropriate reference standard (neurothesiometer ≥25 V threshold) was used to determine the DPN status. Five anatomical sites were tested on each foot, with the number of insensate sites required for a diagnosis of DPN calculated by ROC analysis. This study was considered by the EAC to be of most relevance to informing the decision problem.

3.10.2 Key uncertainties in the clinical evidence

The EAC critically appraised the 6 studies that were in scope and considered their generalisability to the decision problem. All the studies identified had the potential for significant bias concerning internal validity and application to the decision problem. Regarding internal validity, most studies poorly reported recruitment strategy and baseline characteristics of the patients enrolled. Hence there was considerable risk of selection bias and spectrum bias. In most studies, because of the experimental design, a sole operator conducted all the diagnostic tests, introducing the possibility of cognitive bias when interpreting and recording results. This was especially true in studies where the operator already knew the results of the reference standard test.

None of the studies appraised used valid power calculations to estimate the number of participants required to provide a meaningful clinical result. Most of the studies were small in size and therefore probably underpowered. One of the studies was significantly larger than the others with 496 participants (3), but unfortunately did not employ a reference standard. The EAC’s preferred study (2) was second largest, but with 141 participants was still relatively small, which gives rise to significant uncertainty reflected by large confidence intervals (as calculated by the EAC, see Figure 3.5).

Regarding generalisability, one of the main sources of uncertainty resulted from the relatively poor quality of reporting. Another source of uncertainty resulted as a consequence of how the index and comparator tests were applied, for instance the number and location of anatomical sites they were used at, and the number of insensate sites required for a positive diagnosis of DPN. This varied considerably in the identified studies, and also varies considerably in clinical practice: this could have a significant impact on

results. For instance, in the study by Bracewell *et al.* (2), the number of insensate sites required for a positive result was calculated by ROC analysis; however pre-specified thresholds are generally preferred (51) to ensure that the results are generalisable. The target condition was problematic for most studies; in 3 studies because it was not diagnosed by a reference standard (3, 6, 9), in one 1 study because it was detection of 'at risk' feet, not DPN (1), and in 1 study because an inappropriate reference standard was used (8).

Another uncertainty is how the detection of DPN relates to clinical practice. The principal reason for testing for foot sensation and DPN is to stratify the risk of foot ulceration (18). However, since the detection of DPN is only one risk factor that contributes to the risk stratification algorithm specified by NICE (27), it is unclear how improvement in its diagnostic performance will affect the clinical care pathway.

Finally, the identified studies were all diagnostic accuracy studies, which have the aim of providing information on the diagnostic performance of the intervention and comparators. These studies are useful in providing information on the first outcome of the scope, namely 'Sensitivity and specificity in assessment of vibration perception and/or light touch'. However, the EAC has not been able to gather any additional information on the following outcomes listed in the scope for these studies, namely 'Sensitivity and specificity in assessment of grade of neuropathy'; 'Inter-rater agreement of assessment of grade of neuropathy'; 'Accuracy of risk assessment in ulcer formation'; 'Ulcer formation and amputation'; 'Time taken for sensory testing'; 'Quality of life (QoL)'; or 'Device-related adverse events'.

3.10.3 EAC's best estimate of diagnostic accuracy

The EAC considered that of the 6 studies within scope, only the study by Bracewell *et al.* (2012) (2) was suitable to inform the diagnostic accuracy aspect of the decision problem, although this was subject to limitations as described in section 3.10.2. The EAC used the published data to provide an estimate of uncertainty, presented as confidence intervals (see also Figure 3.4). Significance testing (p values) could not be calculated because of insufficient reporting of data (factorial tables comparing index tests would be required). The data for the sensitivity and specificity of VibraTip, 10g monofilament, and 128 Hz are summarised in Table 3.9, with VibraTip having a sensitivity of 0.79 (95% CI 0.69 to 0.90) and a specificity of 0.82 (95% CI 0.74 to 0.90). The EAC judges that the current limited evidence base suggests that VibraTip has equivalent diagnostic accuracy to the 10g monofilament and the 128 Hz tuning fork. However, larger diagnostic studies

would be required to rule out the possibility that there are no clinically important differences between the detection methods.

4 Economic evidence

4.1 Published economic evidence

4.1.1 Critique of the sponsor's search strategy

Insufficient information was provided to enable the EAC to reproduce or critically appraise the sponsor's search strategy. No description was given of any search for economic evidence on the technology. There was no indication that the databases given as a minimum requirement on the submission template (Medline, Embase, Medline (R) In-Process, EconLIT, NHS EED) had been searched.

Given the relative paucity of directly relevant studies identified for the clinical evidence submission, it is unlikely that additional economic studies relevant to the decision problem would have been identified. However, it would have been appropriate for the sponsor to report the methodology and results of a systematic search, even if the results were negative.

EAC's literature search for economic studies

The searches carried out by the EAC to identify clinical effectiveness evidence (reported in Section 3) were not restricted by study design. The searches were designed to retrieve all studies reporting on use of VibraTip to detect diabetic peripheral neuropathy in people with type 1 or 2 diabetes. As a result, these searches would have retrieved any studies of economic evidence on the technology.

In addition to these searches, the EAC designed and ran additional searches in specialist databases indexing economic research; EconLIT, Health Economic Evaluation Database (HEED) and the Cost-Effectiveness Analysis (CEA) Registry. The NHS Economic Evaluation Database (NHS EED) had already been searched by the EAC via the Cochrane Library as part of the searches for clinical effectiveness evidence. Full details of all the search resources and strategies used by the EAC search (included search date and the volume of results returned) are provided in Appendix 3.

4.1.2 Sponsor's study selection

No economic studies were identified or selected by the sponsor. The EAC did not identify any published economic studies concerning VibraTip.

4.1.3 Included and excluded studies

There were no studies found by the sponsor or the EAC. Thus inclusion and exclusion criteria are not relevant.

In submission section 8.1 ('Identification of studies'), the sponsor discussed the studies that were identified for the clinical evidence (1-3, 6, 8, 9), and summarised three key points ('critical conclusions') emerging from these studies. Namely:

- 'VibraTip is of comparable performance to the Neurothesiometer.
- VibraTip is comparable to the 10g Monofilament.
- VibraTip is superior to the Medical Calibrated 128 Hz Tuning Fork.'

The EAC considered that VibraTip had not been demonstrated as being comparable with the neurothesiometer and that there was no statistical evidence for VibraTip being superior (in terms of diagnostic accuracy) to the 128 Hz tuning fork. However, the EAC agreed that the limited clinical evidence from comparative studies between VibraTip and the 10g monofilament might indicate at best equivalent diagnostic accuracy of these devices, although there is considerable uncertainty across all the studies (see sections 3.10.2 and 3.10.3).

In section 8.1.2, the sponsor described the search strategy used to identify the cost of comparator technologies (10g monofilament, 128 Hz tuning fork, and neurothesiometer) and the criteria adopted to select the base case prices from a range of prices returned from google searches. The findings are critiqued by the EAC in section 4.2.5.

4.1.4 Overview of methodologies of all included economic studies

There were no included economic studies by the sponsor or the EAC.

In submission section 8.2.1, the sponsor populated Table C2 with information from studies informing the clinical evidence. The EAC concluded that none of the studies presented in this table were applicable to the economic evidence; all studies cited by the sponsor have been discussed in section 3 of this report.

4.1.5 Overview and critique of the sponsor's critical appraisal for each study

There were no studies identified by the sponsor or EAC to critically appraise. The sponsor populated the economic critical appraisal checklist with an

amalgamation of the previously discussed clinical evidence papers (submission, Table C3). This was not applicable to the economic evidence.

4.1.6 Does the sponsor's review of economic evidence draw conclusions from the data available?

There was no economic evidence presented by the sponsor, or identified by the EAC, to inform conclusions.

4.2 De novo cost analysis

The sponsor created its own *de novo* cost model which was appropriate given the lack of economic evidence available. The EAC critically appraised this model using the methodology of Drummond and Jefferson (1996) (59). The appraisal checklist is reported in Appendix 4. The structure of the model is now described.

4.2.1 PICO analysis

The population (patients), technology (diagnostic intervention, VibraTip), comparators, and outcomes used in the model are described in the following sections.

Patients

The sponsor's description of the population and setting was unclear.

The starting population in the model comprises people with diabetes, with a proportion of these in high risk groups, some with pre-diagnosed DPN. At later stages in the model, these high-risk patients are at higher risk of ulceration, developing ulceration, having ongoing ulceration, or undergoing amputation.

Technology

The intervention used in the model was VibraTip which was associated with a specific cost per-examination. In the base case, it was assumed that VibraTip would be adopted for 40% of the population with diabetes.

For all devices, patients considered to be at low risk of ulceration, were tested for DPN once per year; for patients at high risk of ulceration, testing was carried out 4 times per year.

Comparator(s)

The comparators described in the model were the 10g monofilament (used in the base case and the executable model provided to the EAC) and the 128 Hz tuning fork. Details of the neurothesiometer, including costs, were reported by

the sponsor, but this was correctly excluded in the model, as VibraTip was not expected to replace this specialised technology.

Outcomes

The outcomes were reported as total costs of the intervention arm compared with the 'current practice' arm. Of the outcomes described in the scope (see Table 2.1), only 'Ulcer formation and amputation' was represented in the model.

4.2.2 Model structure

Software

The executable model provided to the EAC was a Microsoft Excel spreadsheet. The spreadsheet had 7 worksheets which were as follows:

- 'Introduction', entitled 'Economic activity effects of NHS innovation'. This provided a brief description of the cell types used in the model (e.g. input cells, formula cells).
- 'Main menu'. This provided hyperlinks to the other worksheets in the model.
- 'Int' [internal]. This sheet provided the main structure and logic of the model (i.e. the decision tree).
- 'Input parameters'. This sheet was intended to provide the variable non-cost inputs required for the model. Some tables were clearly mislabelled, and costs of the devices were incorrectly reported. Not all the input variables were executable (i.e. changing them did not affect the model). To change some key variables in the model it was necessary to manually over-ride a cell used in the model itself (the 'Int' worksheet).
- 'Costs (Social Care)'. This worksheet was unpopulated and did not contribute any inputs into the model.
- "Costs (NHS)'. This worksheet provided cost inputs for the model and consisted of a table that was incorrectly labelled. Some of the costs reported in the table were used in the model, whilst others were not associated with an active cell. Costs included per-examination cost of the devices (VibraTip and comparators), cost of treatment of ulceration, and cost of amputation.

- ‘Results’. This worksheet calculated and reported the aggregate results of the base case model analysis on a population level, and the incremental difference in cost using VibraTip compared with the 10g monofilament. Results from the sensitivity analysis were also presented in this sheet. However, these results had been manually entered in the version of the Excel file provided, so it is not evident how they had been calculated.

Structure

The economic model provided (‘Int’ worksheet) was a decision tree with a three year time perspective. The model consisted of two arms from which the aggregate cost results were derived. These arms represented the flow of patients of the entire diabetic population of the UK through the relevant patient pathways for diabetic foot inspections.

Key issues arising from the structure are that the diagnostic accuracy of VibraTip and its comparators were not used in the base case of the model provided. Thus the results provided in the clinical evidence submission (principally concerning diagnostic accuracy) did not explicitly inform the economic evidence. In the base case, as diagnostic accuracy is not applied explicitly to each test, the model can be classed as a cost minimisation analysis. Such analyses are only appropriate if the intervention and comparators have demonstrated non-inferiority in terms of clinical effectiveness and safety. In consequence, any differences in costs should be solely due to the cost of the devices and related use, rather than the costs of preventative management of diabetic feet.

The first arm of the model was labelled as ‘current practice’, and is represented graphically (submission, section 9.1.4) and in tabular form (submission, section 9.1.5). The decision tree model was designed so that at each stage there were mainly dichotomous ‘choices’. In the ‘current practice’ arm, there was an initial stage where patients are classified as high risk or routine risk, followed by 3 years of similar decisions, resulting in 16 outcome states (2^4). Each state in the model was associated with three variables:

- The transition probability of entering that state;
- The number of patients in that state (calculated by the number of patients in parent state multiplied by transition probability);
- The cost associated with being in that state.

Patients had a risk of ulceration, repeat ulceration (second and third year only), and amputation (third year only). Not all clinical outcomes were

possible, for instance patients without ulcers in the second year could not go on to develop ulcers in the third year, and if diagnosed as routine risk (annual screening) in the first instance, it was not possible to then be diagnosed at high risk in later years. The EAC considered this was a limitation of the model in its representation of clinical reality, although it was unclear if this would bias results.

The second arm described by the model was the intervention arm. Usually, the structure of this arm matches that of the control arm (current practice). However, in this case the intervention arm categorised patients as receiving 'VibraTip' in the upper branch (40% of patients in base case), or receiving 10g monofilament in lower branch (60% of patients), instead of adopting a high or routine risk state. In the subsequent choices at one year, patients either did or did not develop ulcers in the VibraTip and comparator branches, the same as in the current practice arm. In the intervention arm, to compensate for omission of an initial choice of high or routine risk, the model averaged transition probabilities and device costs at each subsequent stage. Other than this, the intervention arm was subject to the same limitations as described in the current practice arm.

Functionality of model

Some input parameters were not used by the model and some were not identified in the correct parts of the workbook (i.e. the 'Input parameters' and 'NHS Costs' worksheets). In some places, table headings and descriptions were incorrect and did not relate to the decision problem. Input variables, which can be changed for sensitivity analysis, were not clearly separated from internal variables.

Only the base case results using the 10g monofilament as a comparator were presented as an executable result. Results from other scenarios, for instance using the 128 Hz tuning fork as the comparator or using sensitivity analyses, were reported in the accompanying narrative only.

The authors of model have added text descriptions and instructions to some cells, such as 'The difference between the Current Practice & Retained Practice Figures with NO VibraTip is £5 K or 0.03% and is the figure used for assessing cost benefit' [located in 'Input parameters' of worksheet]. The EAC does not know the significance of this comment.

4.2.3 Clinical parameters and variables

Clinical data

From the clinical evidence submission, the sponsor identified 8 studies that they considered were in scope and relevant to the decision problem. These studies, and a paper not originally included (4), were also cited (submission, section 8.2.1, Table C2). Six of these studies were diagnostic accuracy studies, 2 were technical papers, and 1 was a prospective RCT. Of these, the EAC deemed the 6 diagnostic accuracy studies were in the scope of the decision problem (1-3, 6, 8, 9) with 1 study being directly relevant to the scope and of a higher quality than the others (2).

However, as described in section 4.2.2, the economic model did not include any discussion of diagnostic or preventative intervention inputs. Hence there was no link between the clinical evidence and the assumptions adopted in the economic model.

Time horizon

The economic model submitted had a time horizon of 3 years. The sponsor did not justify this time horizon, other than stating '[The time 3 year horizon] Represents an overview of assessment and progression for Diabetics' (submission, Table C4).

In the decision tree model presented, dichotomous choices ('branches', years 1 and 2) and outcomes ('leaves', year 3) followed a strict temporal order, which generally resulted in patients deteriorating over the course of the model, and the proportion of the patients with serious complications increasing. However, the model was greatly simplified by being restricted in the main to only two decision options per state, and there were other structural limitations (e.g. no new ulcer formation possible in year 3).

One of the weaknesses of decision trees is that the number of states increases geometrically for each time step (the tree becomes too bushy) (60). For instance, if a 4 year horizon had been selected then there would have been 32 outcomes in each arm rather than 16 outcomes. However, a longer time period might have been more appropriate as diabetes and DPN are chronic conditions, with DPN frequently occurring years or decades after diabetes is diagnosed (61). For these reasons, a state transition model, such as a Markov model, may have been a preferred model choice if the sponsor had intended to report cost consequences or cost-effectiveness analyses, rather than cost minimisation.

Outcome measures

The economic model was a deterministic cost consequence model. Patients entering the model underwent a series of 'decisions', each of which was associated with predefined transition probabilities. Each of the consequences

(states) was associated with a monetary value. For the monitoring states, this value is based on the cost of using the device per patient. For the management of ulceration and amputation, the costs per event were much higher, but because the intervention and comparator devices had no impact on the outcomes from the clinical pathways in the base case, these values were identical for all tests. The sponsor restricted its reporting of results to the aggregate monetary expenditure only. Disaggregated monetary data, costs per patient, or the proportion of patients undergoing specific consequences such as foot ulceration, were not reported.

Discounting

In submission section 9.1.8 (Table C4), the sponsor reported that a discount rate of 3.5% had been 'Added in final figures'. However, there was no evidence from the model that annual discounting had been applied.

Other variables

Important parameters and transition variables that impacted on the patient pathways included the probability of developing DPN (which in the model was a surrogate for being at high risk of ulceration), risk of foot ulcers in low and high risk populations, risk of foot ulcers requiring retreatment, and risk of amputation in patients who have had a foot ulcer retreated. A discussion and critique of these values is given in section 4.2.4. Note these do not influence the base case results because each arm adopts the same transitional probabilities, and these costs were identical for all tests.

Input of sponsor's advisors

In submission section 9.2.5, the sponsor stated that they sought clinical advice and opinion from Professor Andrew Levy of the University of Bristol only. Professor Levy is the inventor of VibraTip. In section 9.3.4, the sponsor states that Dr David Huckle, the Chief Executive of Adams Business associates, independently developed the economic model used to inform the decision problem.

The sponsor did not elaborate on the nature of the advice received from Professor Levy or Dr Huckle, and how it was applied to the economic model or to inform the decision problem.

4.2.4 Transition probabilities, resource identification, measurement and valuation

Description of sponsor's assumptions

The sponsor reported the assumptions in tabular form (submission, section 9.1.6, ‘Assumptions Table’). References for these assumptions are given in a Table at the end of this section (‘Sources of Information used for assumptions’). However, some specific values used were not easily traceable in the references cited. The EAC has summarised the assumptions used and its interpretation of them in Table 4.1.

Table 4.1 *Summary of sponsor's assumptions and EAC's interpretation.*

Sponsor's assumption	**Value stated**	**Reference provided by sponsor**	**EAC interpretation**
Diabetic Population Diagnosed	2.9 M	NICE figures as quoted in scope. CG10 (2004), CG 119 (2011), CG15 (2011).	Figure given in scope is 3.75 million, with 850,000 not diagnosed, equivalent to 2.9 million diagnosed.
Proportion of Patients Susceptible of developing DPN	60%	NICE figures as quoted in scope. CG10 (2004), CG 119 (2011), CG15 (2011).	This figure is given in the scope. However, it was not used in the model. Unreferenced in scope.
Prevalence of DPN for Diabetics	7.4%	NICE figures as quoted in scope. CG10 (2004), CG 119 (2011), CG15 (2011).	Scope states 'The current DPN prevalence rate of 7.4% for diabetes (diagnosed and undiagnosed)'. Unreferenced in scope.
Proportion of Diabetics that Develop Foot/Leg Ulcers	6%	Jeffcoate & Harding, Diabetic Foot Ulcers, Lancet, 2003 May 3362 (9368), 1545-51. Ramsay et al, Diabetes Care 22, 382-387, 1999. Boulton et al New England Journal of Medicine 2004, 352 48-55, 2004 from Manchester UK, reported 15% develop ulcers. Diabetes UK quote figure of 10%.	Three year cumulative incidence of DPN was 5.8% in retrospective cohort study of diabetic patients (62), not discussed in other studies cited (63, 64). In the sponsor's narrative, the sponsor estimated an annual incidence rate of ulcer formation of 2% in low risk groups and 5% in high risk groups, with an overall average of 5.1%. The EAC is unclear how these

Sponsor's assumption	Value stated	Reference provided by sponsor	EAC interpretation
			figures were arrived at or used in the model.
Cost of treatment for Foot Ulcers	£8500	Figures used with CG 10 and CG 119 also. Posnet et al The Burden of Chronic Wounds in the UK, 2008, Nursing Times 104:3 44-45 quoted £ 5,200	The EAC could not find the value of £8500 for cost of ulcer treatment in either CG10 or CG119 (full versions). Nor could the EAC find the reported cost of £5200 in the Nursing Times article (nor any other value for this parameter). The clinical presentation of foot ulcers is likely to be heterogeneous with wide-ranging costs and significant uncertainty.
Recurrence Rate of Ulcers after treatment	10%	Jeffcoate & Harding, Diabetic Foot Ulcers, Lancet, 2003 May 3362 (9368), 1545-51. Ramsay et al, Diabetes Care 22, 382-387, 1999. Boulton et al New England Journal of Medicine 2004, 352 48-55, 2004 from Manchester UK, reported 15% develop ulcers. Diabetes UK quote figure of 10%.	The EAC was unable to find the value of 10% of recurrence rate of ulcers in any of the cited papers (62-64). The figure of 10% could also not be found on the Diabetes UK website (www.diabetes.org.uk). This value was not used in the model.

Sponsor's assumption	Value stated	Reference provided by sponsor	EAC interpretation
Proportion of High Risk Diabetics resulting in Amputation	2%	Figures used with CG 10 and CG 119 also. Posnet et al The Burden of Chronic Wounds in the UK, 2008, Nursing Times 104:3 44-45 quoted £ 5,200	The EAC was unable to find the specific figure of 2% incidence of amputation in any of these references.
Cost of Amputation	£10,780	Figures quoted in NICE Costing Standard for In-Patient Management of Diabetic Foot, March 2011 with reference to NHS Tariff QZ11B (£ 6,759 or £ 9,405 without complications) & QZ11A (£ 13,413 or £ 13,552 with complications). Overall average of £ 10,780 used and compares with earlier figures for E115, E113 & E105 that ranged from £ 8.3 K to £ 14.5 K.	Figures from the NHS Tariffs, as cited in Costing statement: 'Diabetic foot problems: inpatient management of diabetic foot problems' by NICE (65), are correct. However, the sponsor has averaged the uncomplicated and complicated elective and non-elective tariffs without reference to the proportions of patients in each group.
Cost of VibraTip	£9.95	McCallan Medical Quoted figure.	Discussed in section 4.2.5.
Cost of 280 Hz Tuning Fork (Median used)	£28	Commercial figures from suppliers from web site searches (median used).	Discussed in section 4.2.5.
Cost of 10g Mono-Filament (Median	£15.20	Commercial figures from suppliers from web site	Discussed in section 4.2.5.

Sponsor's assumption	Value stated	Reference provided by sponsor	EAC interpretation
used)		searches (median used).	
Cost of Neurothesiometer	£1200	Company listed prices: *Horwell, Misty Medical, Science Lab Co.*	Discussed in section 4.2.5.
NHS Tariff for Foot Ulcer treatment for recoverable cost benefit	£700 +	NHS data. QZ 12Z for Foot Procedures and JC18Z to JC20Z range from £ 120 to £ 771 for the treatment with other costs additional.	These parameters were not used in model.
Improved diagnostic capacity versus current procedures	1-5%	From clinical studies potential from improved performance & ease of use with enhanced early diagnosis. Arbitrary 1% figure used for sensitivity analysis.	Discussed in sensitivity analysis (sections 4.2.6 and 4.3.2).
Cost of Death within NHS	£8000	Scottish Cancer Therapy Network Newsletter Autumn 2003	EAC was unable to retrieve this document. However, unclear if cost of cancer death is generalisable and death was not an outcomes of the model.
Consultation Costs – cost neutral in product comparison	£137	PSSU Unit costs of Health & Social Care.	Costs of NHS staff were not used in the model.

Sponsor's assumption	Value stated	Reference provided by sponsor	EAC interpretation
Reported pressure testing from 8,020 practices in QOF 2012/13.	85.6%	QOF April 2012-March 2013 at http://www.hscic.gov.uk/catalogue/PUB12262	EAC has found the number of tests for diabetic neuropathy in previous 15 months was 90.9% in 2013 (www.gpcontract.co.uk))

The EAC advises that in general, the sponsor's referencing of parameter values was weak. In some cases the EAC could not trace the specific parameters chosen by the sponsor. In addition, the rationale for selection of certain parameters was unclear and hence whether values may be subject to selection bias. This is a concern given the limited sensitivity analysis conducted (section 4.2.6).

Application of sponsor's assumptions to economic model

The EAC cross-referenced the assumptions listed by the sponsor in the submission with those used in the base case of the economic model (see Table 4.2).

Table 4.2 *Application of sponsor's assumptions to economic model*.

Sponsor's Assumptions (and value) from submission	**Application to model (current practice and intervention arms)**	**EAC comment**
Diabetic Population Diagnosed (2.9 million)	Starting population in both current practice and intervention arms.	Starting population (people with diabetes) appropriate.
Prevalence of DPN for Diabetics (7.4%)	Current practice: 7.4% used to define high and routine risks. Intervention: used to provide 'average' value of mixed cohort for transitional probability and cost values.	Structural differences between arms and subsequent need of averaging less transparent but should not bias results. Additionally, detection of DPN only part of the diagnostic workup to assess the risk of diabetic foot ulceration.
Proportion of Diabetics that Develop Foot/Leg Ulcers (6%)	Current practice: value of 5% used for high risk patients and 2% used for low risk. Intervention: values used but averaged (see previous row).	Value quoted in sponsor's submission not used in model. Structural differences between arms and subsequent need for averaging unsatisfactory but should not bias results.
Recurrence Rate of Ulcers after treatment (10%)	Current practice: transition probabilities stated as [1- 0.64] [cell F14 'Input parameters].	The EAC did not find reference to the value of '10%' in the spreadsheet.

Sponsor's Assumptions (and value) from submission	Application to model (current practice and intervention arms)	EAC comment
	In other cases, value linked to cell E14 ('Input parameters') with value of 0.82, but this has been manually over-ridden to 0.7 in the model. Intervention: as above.	The EAC could not determine where the modelled values of 0.64 (64%), 0.82 (82%), 0.3 (30%) originated from. These values were not explained in the submission.
Proportion of High Risk Diabetics resulting in Amputation (2%)	Current practice: amputation rate of 0.36 [cell R99 'Int', derived from cell D15 'Input parameters']. Intervention: only place in model where there are 3 branches: 0.0599 [cell R36 and R66 'Int', manually entered] 0.64 [cell R33 and R63] 0.3001 [cell R30 and R60]	Only patients who were initially at high risk and have ulceration for two consecutive years can undergo amputation (with no third year of ulceration possible) in standard care but possible in intervention arm. EAC did not find reference to the '2%' value in the model's spreadsheet. The EAC could not determine where the value of 0.0599 was derived. It was noted the state that preceded these were the only ones which had 3 outcomes: 'amputation' [0.0599], 'high risk testing [0.64], and 'Ulcer continued treatment' [0.3001]. The last transition probability was directly derived from the other two, suggesting it was placed to keep the model functional (e.g. so patient numbers were preserved).
Costs of intervention and comparators.	Per use cost employed. For intervention arm, averaged costs used (high risk patients have 4 tests per year).	Discussed in section 4.2.5.
'No event'	Current practice: cells R75, R91, and R105	In year 3 of the model, there are 5 'no event' cells which have zero

Sponsor's Assumptions (and value) from submission	Application to model (current practice and intervention arms)	EAC comment
	probability 0. Intervention: R12 and R42 probability 0.	probability attached to them. This means there is 100% probability of the inverse pathway occurring (patient retaining high or controlled risk status). The EAC could not determine the logic behind this aspect of the model.
Cost of treatment for Foot Ulcers (£8500)	Used as annual cost of treatment of foot ulcers in both arms.	Value independent of category of risk of patient. Total costs should be identical in each arm in in base case.
Cost of Amputation (£10,780)	Cost associated with amputation in year 3.	Total costs should be identical in each arm in base case.
Proportion of Patients Susceptible of developing DPN Cost of Death within NHS Consultation Costs		Reported in economic narrative but not used in economic model.

The EAC advises the derivation of values for many of the parameters was opaque. The EAC requested clarification from the sponsor seeking clarification on 4 of these parameters and structural issues (proportion of diabetics that develop foot/leg ulcers; ulcer recurrence after treatment; proportion of high risk diabetics resulting in amputation; 'no event cells'). The sponsor's response did not clarify the EAC's understanding of these issues.

4.2.5 Technology and comparators' costs

Unit costs of technology

In submission section 8.1.2, the sponsor estimated the costs of the intervention (VibraTip) and the comparators (10g monofilament, 128 Hz tuning fork, and [for completeness] neurothesiometer). The sponsor stated the costs of the comparators were 'from main manufacturer and distributor list prices for the products concerned' and that 'average list prices from these suppliers

[were] used'. The suppliers used were Williams Medical Supplies (www.wms.co.uk) and Canonbury Products (www.canonbury.com). The unit costs used in the sponsor's economic analysis did not include value added tax (VAT). The EAC has used the NHS Supply Chain (www.supplychain.nhs.uk), accessed on 14th March 2014, as the basis of its estimates for unit costs of the comparators. The comparative unit costs excluding VAT are listed in Table 4.3.

Table 4.3. *Sponsor's and EAC's estimates of unit costs.*

Diagnostic device	Sponsor estimate (excluding VAT)	EAC estimate (excluding VAT)	Comment
VibraTip	£9.95	£9.95	Price of £9.95* from manufacturer's website: www.mccallanonline.com
10g monofilament	£15.20	£16.00	Bailey retractable monofilament with Duraban
128 Hz tuning fork	£28.00	£31.28	'Reusable Gardiner Brown' 128 Hz tuning fork.
Neurothesiometer	£1200.00	£1412.03	Not used in economic analysis
* A price of £6.50 (plus postage and packing is cited on the device website (www.vibratip.com) but this does not seem to be available (no payment mechanism on website).			

VibraTip is a small device which could be misplaced or lost. To counter this, a peripheral component, VibraClip, consisting of a Lanyard attachment, is available for £9.95 (excluding VAT) from Williams Medical Supplies (www.wms.co.uk). This cost has not been included in any analysis by the sponsor or subsequently by the EAC.

Per-examination costs

The relevant costs of the intervention and comparators used in the sponsor's *de novo* economic analysis was the *per-examination cost* (device cost associated per one clinical examination), rather than the unit purchase price (submission, section 9.3.6).

The sponsor listed 6 assumptions made during its calculations of per-use costs. Table 4.4 summarises these and provides the EAC's critique of them.

Table 4.4 *Sponsor's assumptions when calculating per-examination costs and EAC 's critique.*

Sponsor's assumptions	**EAC's critique**
The comparator products require greater skill in use and additional time to perform the assessments than VibraTip. These comments are supported by all of the 9 references noted for the Economic Assessment above.	The sponsor makes reference to papers provided in the clinical evidence submission and listed in Table C9. The EAC did not find quantitative evidence for the need for greater skill or greater time to perform the comparator tests in any of these papers.
The time to perform the necessary procedures with VibraTip can be faster than the comparator products (6) but this is time that might not be valuably occupied and is not considered as a recoverable cost in the economic calculations.	Time savings from VibraTip over its comparators has not been quantitatively demonstrated. However, the EAC agrees that any time savings made would not necessarily free time resources because clinic appointment slots are usually pre-determined and testing of foot sensation forms only one part of the overall assessment.
In the same way the ease of use, general utility and limited training required for VibraTip compared to the comparator products is not quantified for cost savings.	The EAC agrees with this statement.
The product usage costs are based on cost/test related to the useful life of each comparator device.	Agreed and expanded in this report (see section 4.5).
As each device has a finite useful life, replacement of comparator products with VibraTip is assessed in the Economic Model from the level of "Adoption" of VibraTip against these other products.	Level of adoption is varied in the sponsor's sensitivity analyses (section 4.2.6).
Use of the high cost gold standard Neurothesiometer is not considered in the main economic Model as VibraTip would not be used as a replacement product although the comparisons undertaken and referenced above [Table C2] showed that VibraTip gave	The EAC agrees that VibraTip would not replace the neurothesiometer, but does not agree that the clinical evidence shows that VibraTip has equivalent diagnostic accuracy to the neurothesiometer (none of the clinical studies were designed to

comparable clinical performance	demonstrate this).

From these assumptions, the sponsor calculated the *per-examination* cost of VibraTip and its comparators, using a combination of useful life of the device and clinic throughput. These contributory factors were not applied identically for each device, as described in the following sections.

10g monofilament

The sponsor made the following assumptions when calculating the per-examination costs of the 10g monofilament:

- The monofilament can be used 10 times before it requires resting for a 24 hour period. As it is possible more than 10 patients would be seen in clinic per day, 2 monofilaments per clinic would be required to provide availability.
- The maximum useful life of the monofilament before it expires is 1 year.
- Three clinics are held per week.

Therefore, the cost per-examination of the monofilament are:

- £30.40 [2 x unit cost] / (10 [uses per day] x 3 [clinics per week] x 50 [working weeks per year]) x 2 (units bought) = £30.40 / 3000 = £0.0101. £0.01 (1 pence) was the value used in the economic model.

The EAC considered this was probably a reasonable estimate of the per-examination cost of using the monofilament, assuming clinics were operating at the maximum capacity described. If this assumption did not hold (i.e. there were less than 10 patients examined per clinic), then the per-examination cost of the 10g monofilament would exceed £0.01. Alternatively, if there are 4 clinics a week the cost would fall to £0.0076, with the cost falling as number of clinics rises.

It is recommended that the 10g monofilament is rested for 24 hours following 10 patient examinations in CG10 (18) based on a technical study by Booth et al. (2000) (24). The sponsor cited that the expiration date of 1 year was from guidance from the US Food and Drug Administration (FDA), which does not apply to the UK. However, this seems a reasonable estimate. The only 10g monofilament listed on the NHS Supply Chain is the Bailey retractable, and the instructions for use for this device state that the useful life is 6 months for heavy (daily) use or 24 months for more occasional use. Thus the useful life

of the monofilament is dependent on the throughput of the clinic. A more recent technical study (12) has indicated that the useful life of some branded monofilaments may be limited to 70 to 90 patients, although it is noticeable from this study that the Bailey's monofilament compared favourably with others in terms of initial calibration and subsequent deterioration. This uncertainty is addressed further in section 4.5.

The sponsor estimated that on average the 10g monofilament would be used in 3 diabetic clinics per week. The EAC could not verify this estimate but it is likely that the usage pattern of the monofilament is dependent on the location and setting. The sponsor should have measured the impact of such uncertainty through sensitivity analysis.

128 Hz tuning fork

The sponsor determined the cost per-examination of the tuning fork with the following assumptions:

- Appointment slot time 15 minutes (time to conduct all clinical examination and history).
- Two tuning forks required per clinic, 3 clinics held per week.
- Total usage per year calculated as 2 [tuning forks] x 4 [uses per hour] x 6 [hours of clinic] x 3 [clinics per week] x 50 [working weeks] = 7200.

Therefore the cost per-examination of the tuning fork was calculated as £56.00 [cost of tuning fork x 2] / 7200 = £0.0077 (approximately £0.008, or 0.8 pence). This was the value used in the economic model.

An inconsistency presented in the analysis was that the estimated clinic throughput for the tuning fork (24 patients per clinic) was different to that described for the 10g monofilament, resulting in cost denominators (number of patients per clinic per year) of 3000 and 7200 respectively for the devices. If the same assumptions had been used for the tuning fork as for the 10g monofilament, then the cost of the 128 Hz tuning fork would have been higher (£0.019 per-examination).

However, the EAC considered that the main flaw in the cost analysis presented was that the tuning fork was given a useful life of 1 year before replacement. The EAC is unaware of any evidence to suggest the tuning fork deteriorates over a year's use. It has been claimed, for instance by Levy (2010) (6), that tuning forks tend to be stolen, but this risk has not been verified and is probably unquantifiable. Feedback from clinical experts was almost unanimous that tuning forks have an unlimited useful life and are only

replaced when they go missing. Increasing the useful life associated with the 128 Hz tuning fork would clearly reduce its per-examination costs.

The sponsor also claimed in this section that the tuning fork had poorer clinical performance than its comparators, owing to the possibility of patient cueing and greater variability in use. However, this is not consistent with the clinical evidence or the assumptions adopted in the economic model.

VibraTip

For VibraTip, the sponsor did not attempt to estimate clinic throughput or usage over the course of 1 year, but instead extrapolated evidence on the battery life to the clinical setting such that:

- Number of single tests = 5000.
- Cost of VibraTip = £9.95.
- Therefore cost per test with VibraTip = £9.95 / 5000 = £0.002 (0.2 pence).

The EAC considered that this analysis was not a fair comparison with the other tests, and that the true per-examination cost of VibraTip might be as high as £0.0995 (9.95 pence). The reasons for this are discussed in section 4.5.

Miscellaneous costs

The sponsor noted other possible resource savings from adoption of VibraTip rather than other devices:

- Increased primary care management, diverting management away from hospital clinics.
- The potential for increased testing, with improved patient care resulting in fewer incidences of ulcers and their treatment costs.
- Savings from reduced amputations and reduced need for hospitalisation.
- Possible reduced social care resources.
- Minimal training of staff required.

There is no evidence to support these claims, particularly relative to the 10g monofilament.

4.2.6 Sensitivity analysis

The main purpose of sensitivity analysis in economic modeling is to provide a level of confidence that the results are valid (within a range), and allow for an analysis of uncertainty. The simplest method to achieve this is one or two-way deterministic sensitivity analysis (66).

In submission sections 9.4.1 and 9.4.2, the sponsor discussed what it considered were the two major sources of uncertainty. The first was the projected uptake (adoption) of VibraTip compared to its comparators (10g monofilament or 128 Hz tuning fork), whilst the second concerned adherence to NICE clinical guidelines (CG10) and the possible implications of improving compliance, particularly the element of screening for DPN.

Adoption of VibraTip

The base case intervention arm assumed that 40% of tests for DPN (in both low and high risk groups) would be performed using VibraTip ('intervention'), with the remaining 60% using 10g monofilament or 128 Hz tuning fork. In the current practice arm, all patients received either 10g monofilament or 128 Hz tuning fork.

The sponsor conducted sensitivity analyses with proportional adoption rates of VibraTip of 0%, 20%, 40% and 100%.

This sensitivity analyses did not test the underlying uncertainties in the model. In the base case, the patient pathways modelled were not affected by the relative use of VibraTip or its comparators, with only the per-examination cost of the device itself influencing the results. Thus, if VibraTip is cost-saving in the base case any changes to its adoption uptake should be directly proportional to the savings made. Altering the proportion of adoption affects the magnitude of savings, but not the direction.

Patient pathways

As a second sensitivity analysis, the sponsor assumed VibraTip use was associated with a 1% relative risk reduction in ulcer formation compared with comparator devices. Thus in patients who were at high risk, the probability of having an ulcer decreased from 5% to 4.95%, and for patients at low risk, the probability decreased from 2% to 1.98% (submission, Table C10.2). Given the high cost of treating ulcers and resultant amputation in the model, this has a material impact on costs (see section 4.3.2). The sponsor also combined this assumption with the adoption uptake, to provide crude two-way deterministic analysis.

The EAC considered that this analysis did not measure parameter uncertainty, but instead could be considered to be a type of scenario analysis. The assumption of a 1% reduction in ulcer rates was unsubstantiated but seemed to be associated with the potential from improved performance and ease of use with enhanced early diagnosis of DPN, and subsequent avoidance of foot ulcers. As noted in Section 3.10, the available limited clinical evidence indicates that VibraTip has equivalent or non-inferior diagnostic accuracy to the 10g monofilament and the 128 Hz tuning fork.

Potential sensitivity analysis

The sponsor failed to address the uncertainties associated with cost per-examination of devices; relative diagnostic accuracies; the risks and uncertainties associated with the probability of ulcer formation and amputation; and subsequent management costs.

4.3 Results of de novo cost analysis

4.3.1 Base-case analysis results

The sponsor reported the base case results by three methods. In the first (and most appropriate) method (submission, 9.5.1), the aggregate costs associated with the current practice and intervention arms were summed and presented. These costs represented the budgetary impact of using the devices in the entire modelled diabetic population (2.9 million) over the three year period of the model, rather than the costs per patient. In the second method (submission, 9.5.2) the sponsor used the spreadsheet to calculate the 'theoretical product test costs'. Later in the same section, the sponsor calculated what is termed 'real world product test costs'.

Sponsor's reported base case results

The sponsor reported the results of the base case of the model (submission, Table C2). The cumulative costs associated with each branch of the decision tree were summed for the current practice and intervention arms of the model. In the base case, adoption of VibraTip was set to 40%. The results were:

- VibraTip: £1,467.86 M
- Diabetic Monofilament : £1,467.91 M
- 128 Hz Tuning Fork : £1,467.90 M

The sponsor did not report the incremental differences in costs between VibraTip and its comparators. The EAC calculated these savings as £0.05 M

(or £50,000) compared with 10g monofilament and £0.04 M (or £40,000) compared with the tuning fork.

From these results, the EAC calculated that the sponsor's claimed savings were:

- Use of Vibratip saves £50,000 compared with the use of monofilament for monitoring 2.9 million diabetics (equivalent to saving 1.7 pence per patient over 3 years)
- Use of Vibratip saves £40,000 compared with the use of tuning fork for monitoring 2.9 million diabetics (equivalent to saving 1.4 pence per patient over 3 years).

Comparison with 10g monofilament

In the spreadsheet supplied to the EAC, the 10g monofilament was used as the only comparator. The results for the base case analysis for the 10g monofilament, as reported in the spreadsheet in 'Results' worksheet (rather than the word document submission) were as follows:

- Vibratip: £1,467,900,713
- 10g monofilament: £1,467,930,096
- Cost savings from VibraTip compared with monofilament: £29,384

These values are similar, but not identical to, the results presented in the written submission.

The EAC investigated the reason for the discrepancy between these values. Given the assumptions, the proportion of people who remain on current practice (in this case 10g monofilament) should be *exactly* the same as the current practice arm, when scaled up. By adjusting cell F23 of 'Int' to 0 (redirecting all patients to us 10g monofilament rather than VibraTip) the results changed to:

- Current practice arm (100% use 10g monofilament): £1,467,930,096
- Retained practice arm (100% use 10g monofilament): £1,467,935,122

Hence the results were similar, but not exactly the same, with the retained practice arm being £5025 more than the current practice arm. Proportionately, this difference is 0.000342%.

The EAC noted that in the ‘Input parameters’ worksheet a note stated ‘The difference between the Current Practice & Retained Practice Figures with NO VibraTip is £ 5 K or 0.03% and is the figure used for assessing cost benefit’. Thus it appears the sponsor was aware of the slight discrepancy.

Overall the EAC notes that the results presented by the sponsor in the submission were not the same as those of the executable model and that these may have a small error. Of greater concern, the results as presented might not inform the decision on which device minimises costs to the NHS because of the inclusion of an incorrect unit cost for VibraTip (see section 4.5).

Comparison with tuning fork

The sponsor did not provide the EAC with an executable spreadsheet which included the 128 Hz tuning fork as the comparator. Since the EAC could not fully replicate or interpret the sponsor’s base case results for the 10g monofilament, it did not attempt to replicate the sponsor’s results for the 128 Hz tuning fork.

Theoretical product test costs

For this analysis, the sponsor assumed an uptake of 100% of all devices and calculated the cost by multiplying the number of uses throughout the three year period by the average cost per-examination of each device. These results are:

- Diabetic Monofilament at £134,684.
- 128 Hz Tuning Fork at £107,747.
- VibraTip at £21,537.

The sponsor then calculated the average tests required per year, the number of devices per year, and the total product cost per year. This calculation gave the same results as above but per year rather than over 3 years (i.e. 10g monofilament: £45,500; tuning fork £33,000 and VibraTip £7100). These cost differences arose because the sponsor assumed fewer VibraTip devices are necessary (718) than 10g monofilaments (2993) or tuning forks (1181).

The EAC has been able to replicate these results, including those for the tuning fork (by adjusting the cost inputs). However, the EAC advises the underlying assumptions may be flawed, particularly in respect of the number of tests per VibraTip in its useful life (section 4.5) Hence the results should be interpreted with caution

‘Real world’ product test costs

In the theoretical analysis described above, only 718 VibraTip devices per year were required to supply the NHS in England and Wales, which is not plausible. The sponsor also undertook a ‘real world’ analysis for VibraTip and the comparators (10g monofilament and 128 Hz tuning fork). The sponsor provided a range of assumptions based on the diabetic population tested annually for DPN (67); the average number of tests performed on each diabetic patient per year (1.55, derived from economic model); the average number of tests per general practice registered with QoF (approximately 8,000 practices (67), average 480 tests performed), and assumed two devices would be required per practice.

Using only the number of GP practices (8000), number of devices per practice, and the procurement cost of the devices only (see section 4.2.5), the sponsor made the following cost estimates:

- Monofilament = £243,000
- Tuning Fork = £448,000, or £224,000 if only 1 device per practice
- Total NHS costs for 8000 practices: VibraTip = £159,000.

In this analysis, VibraTip is reported as less expensive based on the unit cost of the device. At £9.95 excluding VAT it is less expensive to purchase than its comparators (10g monofilament at £15.20 and 128 Hz tuning fork at £28.00). However, this ignores the factors informing the per-examination usage (section 4.2.5). Additionally, this analysis assumed all the testing would take place in primary care settings, regardless of risk of foot ulceration. The EAC concluded the ‘real life’ costing analysis did not usefully inform the decision problem.

4.3.2 Sensitivity analysis results

The sponsor reported the results of the sensitivity analysis (described in submission, section 4.2.6) in Table C10.1 (single-way sensitivity analysis, submission, section 9.5.6) and Table C10.2 (multi-way scenario analysis, submission, section 9.5.7).

Single-way sensitivity analysis

In the first sensitivity analysis, the sponsor varied the uptake of VibraTip from the baseline of 40% to 20% and 100% (rationale for these values not stated). The savings were proportionate to the uptake of VibraTip. Compared with the 10g monofilament, the sponsor reported that the annual savings made by VibraTip would be £17,200, £34,400 and £86,000 at 20%, 40% and 100%

adoption levels respectively. The EAC attempted to replicate this analysis by altering the value of the cell E10 in ‘Input parameters’ (actively linked to cell F23 of ‘Int’) and retrieved the corresponding values of £12,179, £29,384, and £80,997. Thus the value of £5000 has been added to the sensitivity analysis, which corresponds to the sponsor’s comment on the ‘Input parameters’ worksheet as described previously (section 4.3.1), but the EAC does not understand the reason for this adjustment.

The sponsor also reported the results for the sensitivity analysis compared with the tuning fork as £12,900, £25,800, and £64,500 for levels of adoption at 20%, 40% and 100% respectively. As the EAC was not provided with the workings for this analysis it could not replicate it.

Multi-way scenario analysis

For the multivariate sensitivity analysis the sponsor varied the uptake of VibraTip and assumed adopting VibraTip would result in a 1% reduction in the rate of formation of diabetic foot ulcers in people who had not had an ulcer in the previous year. The sponsor reported that, compared with 10g monofilament, this assumption would lead to savings of £3,210,000 (20% adoption), £6,420,000 (40% adoption), and £16,060,000 (100% adoption). Compared with the 128 Hz tuning fork, the respective savings were £3,180,000, £6,350,000 and £15,880,000.

The EAC was able to replicate these results following the instructions for manual changes provided in the spreadsheet (‘Input parameters’). However, the EAC were unable to confirm the validity of these transition variables. As discussed in section 4.2.6, there is no evidence to support the assumed risk reduction and thus the result of the sensitivity analysis is not informative.

4.3.3 Subgroup analysis

The subgroups were described in the scope as ‘People with diabetes having routine annual checks in a primary care setting; people with diabetes having more frequent checks in a secondary care setting’. As these subgroups were integral to the model, subgroup analysis in this context was not appropriate or possible.

4.3.4 Model validation

The sponsor stated ‘The clinical evidence and assumptions were all set out above. The internal checks for the structure of the model, particularly for values between the Current and Retained/VibraTip structures, have been left in the attached model’. The EAC saw no evidence that full quality assurance validation had been performed.

4.4 Interpretation of economic evidence

Consistency with published economic literature

In submission section 9.8.1 (consistency with published literature), the sponsor simply stated ‘Results were as expected from the product device price differences’. This statement reflected the key importance of unit and per-examination costs in the sponsor’s model, and the lack of published economic evidence available to draw comparisons with.

Relevance to NHS settings

In submission section 9.8.2 (how well the economic evidence matches the setting in the scope) the sponsor stated the data covers ‘all applied sectors carrying out Diabetic clinics’. The EAC considered this was true except for the ‘real world product test costs’ which assumed all tests are in GP practices (described in section 4.3.1).

Strengths and weaknesses of analysis

In submission section 9.8.3 (strengths and weakness of the analysis), the sponsor correctly identified ‘The key assumption relates to the life time of each device as this impacts the cost per test and the ultimate analytical results’. The EAC’s judgement of the strengths and weaknesses of the submission is now described.

Strengths of analysis

The EAC considered the main strengths of the analysis were that it matched the scope well in some categories. The population and settings was people with diabetes undergoing foot examination in either primary or secondary care settings. The intervention (VibraTip) and comparators (10g monofilament and 128 Hz tuning fork) also correctly matched the scope. However, only one outcome described in the scope (‘Ulcer formation and amputation’) was described in the economic model.

Methodologically, the use of a decision tree was an appropriate (but perhaps not the optimal) choice to simulate the short-term outcomes of diabetic foot inspection (68). However, as the base case model did not introduce different diagnostic accuracies and hence changes in the transition probabilities of clinical pathways due to the effect of the technology, it became a *de facto* cost minimisation model.

Weaknesses of analysis

The EAC considered the main weakness of the analysis was the failure to discuss how the findings from the clinical evidence in section 3 informed the assumption of clinical equivalence implicitly assumed in the economic model. The rationale for clinical equivalence and the demonstration of equivalent clinical outcomes to support the argument should have been made explicitly.

This meant that the model, in its base case, was a *de facto* cost minimisation model. This may have been appropriate considering the available clinical evidence base (section 3.10), but this being the case some of the model's functionality was redundant other than to enable sensitivity (scenario) analyses. Other weaknesses of the *de novo* model included:

- The sponsor's calculations of per-examination costs for VibraTip and its comparators (sections 4.2.5 and 4.5).
- Poor referencing of assumptions used in the model; inconsistencies between those in the model and those in the narrative submission; and no explanation on why values were chosen or the process to identify possible values (except in respect of cost of devices).
- The use of two slightly different structures to model the risk of an ulcer in the population.
- Replication of the sponsor's results required adjustment of internal model parameters in addition to the declared model inputs.
- Poor sensitivity analyses which failed to address the uncertainty associated with key parameters including the cost per test, useful life of devices and different sensitivity and specificities.
- The sensitivity analyses that were performed did not inform the sensitivity of the base case results to these uncertainties:
    - Results from the sensitivity analyses assuming different uptakes of VibraTip were linear and not an informative measure of uncertainty in the model.
    - The second sensitivity analysis was based on an unproven assumption of improved ulcer detection rate for VibraTip relative to the other devices, which reported large savings, but was based on an assumption without evidence.

4.5 Additional work undertaken by the External Assessment Centre in relation to economic evidence

As the sponsor's base case economic model was a *de facto* cost minimisation model (whereby diagnostic accuracy was assumed to be equal in the intervention and comparator arms), the costs of the diagnostic testing *per se* are crucial to the analysis. The EAC considered that the unit costs estimated by the sponsor were largely accurate. However, the EAC considered that the sponsor's calculation of the *per-examination* costs, particularly that of VibraTip, were flawed, and therefore provided additional estimates for these values.

For its calculation of the per-examination costs for the 10g monofilament and the 128 Hz tuning fork, the sponsors used estimates based on simulated clinic use (see section 4.2.5) and expected useful life. Although there was some inconsistency in these values, and there was only limited evidence to support some of the assumptions, these unit cost estimates are plausible.

However, for the per-examination costs associated with VibraTip, the sponsor assumed the device could deliver 5000 tests within its average battery life. This approach is inconsistent with that taken for the comparator devices, which introduces additional uncertainty.

The estimate of 5000 diagnostic tests was derived from a conservative estimate of the 6000 actuations that were reported in the technical study by Horsfield and Levy (2013) (4) (described in Appendix 2), and assumes that only one site on one foot is tested. Whilst there is considerable uncertainty concerning the optimal clinical practice when testing for DPN, it is anticipated that as a minimum both feet are tested, but also often multiple sites on each foot.

The submission cites the Horsfield and Levy study (4), which reported that the amplitude of VibraTip did not decay significantly over the course of the first 1000 activations in one of their test protocols, leading the authors to conclude that 'VibraTip would provide a very consistent source of vibration to test at least 100 patients, if not considerably more'.

The EAC has been informed by the clinical advisors that variations in the stimulation amplitude may have important clinical consequences. In this scenario, where VibraTip was simulated as being used 5 times on each foot (similar to the usual use of the 10g monofilament), and the device discarded after amplitude significantly declined at 1,000 uses, the cost per patient examination would be £0.0995 (nearly 10 pence) (£9.95/(1000/10)).

VibraTip use in this way was reported by Bracewell *et al.* (2012) (2), which is the principal paper informing the EAC's conclusion on the clinical evidence (section 3.10.3).

The EAC considered that for the 128 Hz tuning fork, it was not possible to provide a meaningful per-examination cost in this way, because it has an unlimited useful life.

It is known that the useful life of the 10g monofilament is limited, and this issue has been addressed in the literature. A technical study performed by Lavery *et al.* (2012) on several brands of monofilaments showed considerable variation in their durability after repeat testing (12). However, the Bailey's 10g monofilament (used widely in the NHS) fared comparatively well, with the results indicating the device produced a buckling force within the limits of acceptability (between 9 and 11 grams) after 1800 to 2400 tests. If the device is used 5 times on both feet, this suggests the Bailey's monofilament would have a useful life of approximately 200 patients before requiring replacement.

The EAC has recalculated the per-examination costs of VibraTip (using battery life) and the 10g monofilament (using degradation of buckling force) in a range of scenarios in Table 4.5 and the sponsor's unit cost estimates.

Table 4.5 *Cost per-examination estimates for VibraTip and the 10g monofilament*.

	Number of activations in useful life*	**Number of sites per-examination (protocol)**	**Cost per-examination (pence)**
VibraTip	5000	1 (hallux on 1 foot, sponsor's submission).	0.20
	5000	2 (hallux on both feet)	0.40
	5000	4 (hallux and malleolus on both feet, clinical experts)	0.80
	5000	10 (5 sites, both feet (2, 6))	1.99
	1000	1 (hallux on 1 foot, sponsor's submission).	1.00
	1000	2 (hallux on both feet)	1.99
	1000	4 (hallux and malleolus on both feet, clinical experts)	3.98
	1000	10 (5 sites, both feet (2, 6))	9.95
Bailey's 10g monofilament.	2000	4 (hallux and malleolus on both feet, non-clinical protocol for direct cost comparison only)	3.04
	2000	10 sites (5 sites, both feet, manufacturer's instructions, (12))	7.60
	800	10 sites, 'worst case' scenario' reported by Lavery *et al.* (2012) (12)	19.00

* 5000 activations is representative of number of activations before battery fails, 1000 is representative of number of activations at consistent amplitude.

Note: sponsors per-examination estimate for 10g monofilament 1 pence, for 128 Hz tuning fork 0.8 pence.

Thus the EAC estimates, that depending on the importance of maintaining a constant amplitude for VibraTip, the number of sites per foot and number of feet tested, the per-examination cost of VibraTip is between 0.2 pence (estimate used in sponsor's economic model) and 9.95 pence (usage estimate consistent with heavy use stated by Horsfield and Levy (4), and that adopted in the diagnostic accuracy study by Bracewell et al. (2)). This compares with the sponsor's estimates of 1.00 pence for the 10g monofilament and 0.80 pence for the tuning fork, and revised estimate by the EAC of between 3.04 pence and 19.00 pence (sponsor's estimate of worst case scenario). The EAC's most plausible estimate, using data reported in the technical study by Lavery *et al.* (2011), where 2000 uses of the monofilament at five sites on each foot was practicable, gave a monetary value of around 7.60 pence (12). A potentially important issue with both VibraTip (through battery discharge) and the 10g monofilament (through reduced plasticity) is that the operator may be unaware the devices are losing functionality. In both cases clinical use beyond their effective useful life would result in reduced sensory force being applied to the patient. In this scenario, it is likely that a person with true DPN would be less likely to feel VibraTip or the 10g monofilament, and be correctly diagnosed (true positive). However, a person without DPN would also be less likely to detect the devices, and therefore would risk being incorrectly diagnosed with DPN (false positive). This could potentially lead to inappropriate referral, over treatment, and increased costs, but would be unlikely to have a detrimental effect on the patient's health.

4.6 Conclusions on the economic evidence

The sponsor did not provide any published economic evidence on VibraTip and no economic studies were identified by the EAC. The sponsor provided the EAC with a *de novo* economic model structured as a decision tree that simulated the patient pathways of diabetic patients over three years. Outcomes in the model included patients developing high risk of ulceration, developing ulcers, and requiring amputation. Each state in the model was associated with a transition probability, a number of patients (as a proportion of the diabetic population) and a cost. The costs were associated with device usage, management of ulcers or amputation.

The EAC critically appraised the model and the accompanying narrative in the sponsor's submission. The model, in its base case, had no diagnostic input and thus implicitly assumes diagnostic equivalence.

Given the assumed diagnostic equivalence, each arm in the model had the same clinical outcomes and clinical costs; only the per-examination device costs contributed to the difference in aggregate costs (making it a *de facto* cost minimisation model). The EAC also noted that the two arms of the model,

current practice and intervention (which incorporated ‘retained current practice’) were structurally different. In general, the input parameters for transition probabilities and costs were not judged to be robust or transparently derived from the references provided, making it difficult for the EAC to replicate the sponsor’s results.

The EAC disagreed with the methodology used by the sponsor to estimate device *per-examination* costs, and considered that, depending on its useful battery life and the number of foot sites per patient used, VibraTip would cost between 0.2 pence and 9.95 pence per person examined. This compared with the sponsor’s estimate of 1.00 pence and 0.80 pence for the 10g monofilament and 128 Hz tuning fork respectively, and the EAC’s revised estimate of 7.60 pence for the 10g monofilament.

In the base case analysis, the sponsor claimed 40% uptake of VibraTip would result in savings, over 3 years, of approximately £50,000 compared with 10g monofilament, and £40,000 compared with the 128 Hz tuning fork. The EAC could not exactly replicate these figures with the model provided. However, as these values were solely related to the per-examination costs, any changes to these values would directly impact on the overall aggregate costs. In scenarios where VibraTip is more expensive than current practice then using that device will increase NHS costs. The sponsor provided two sensitivity analyses. In the first one-way deterministic analysis it was shown that increasing the level of VibraTip adoption would result in proportionate cost savings (or expenditure if VibraTip was more expensive). In the second analysis, the sponsor assumed that the use of VibraTip (at an adoption rate of 40%) would reduce the incidence of ulcer formation by 1%, and calculated that this resulted in savings of £6,420,000 compared with 10g monofilament and £6,350,000 compared with the 128 Hz tuning fork. However, there is no evidence to support this assumption.

In summary, the EAC considered that the results of *de novo* economic model did not usefully inform the decision problem. The model was built to accommodate different diagnostic accuracies with associated different clinical outcomes and costs. However, many of the parameters and much of the functionality of the model was redundant because no difference was assumed in the diagnostic accuracies of the device. If the clinical evidence had shown differences in accuracy then it is important these could be modelled.

The main issues were:

- There was no discussion of the assumptions on diagnostic accuracy and their consistency or otherwise with the results of the clinical effectiveness section.
- The unit cost adopted for VibraTip and its comparators.

- An absence of sensitivity analyses on all unit costs.

If a cost minimisation analysis was adopted, including the *per-examination* device costs only, VibraTip might be more expensive than either comparator under heavy usage (4), with a plausible per-examination cost of between 3.98 pence (for 4 sites per person) and 9.95 pence (for 10 sites per person). This compares with a per-examination cost of 7.60 pence for the 10g monofilament assuming 10 sites per person. Given this, the EAC considers that the economic case for VibraTip is not demonstrated.

5 Conclusions

Currently NICE guidelines for the assessment of diabetic feet recommend that the 10g monofilament or a device testing for vibratory sensation (in practice the 128 Hz tuning fork) should be used to test for the presence of DPN, as part of a clinical algorithm to assess for the risk of foot ulceration (18). The current evidence base for these specific devices is limited in terms of quantity and quality (38), and the guidelines have been developed pragmatically to reflect this; these are under review and due to be updated in 2015 (14). VibraTip is a novel device which is used to test for vibratory perception and is intended as a direct replacement or alternative for the 10g monofilament or 128 Hz tuning fork. The sponsor, McCallan Medical, has presented the clinical evidence and economic case to support the adoption of VibraTip for the detection of DPN.

In its clinical evidence submission, the sponsor provided 6 diagnostic accuracy studies which the EAC considered fitted the scope of the decision problem (1-3, 6, 8, 9), and a further 2 technical papers, which, whilst not in scope, provided useful information concerning the performance and endurance of the device (4, 5). Of the diagnostic accuracy studies, 3 were of limited usefulness because they did not have a reference standard (3, 6, 9), 1 because it used a modified reference standard (8), and 1 because it recruited a population who had pre-existing DPN and its target condition was 'at risk' feet (1). This left 1 study, by Bracewell *et al.* (2012) that the EAC considered was of reasonable quality and most closely matched the decision problem scope (2).

The study by Bracewell *et al.* recruited patients with diabetes and investigated the diagnostic accuracy of VibraTip compared with the 10g monofilament, the 128 Hz tuning fork, and NeuroTip, and used an appropriate reference standard (neurothesiometer with a threshold of 25 V). The results of this study indicated that VibraTip had a sensitivity of 0.79 (95% CI 0.69 to 0.90). The sensitivity of the 10g monofilament was 0.84 (95% CI 0.75 to 0.94) and the 128 Hz tuning fork was 0.69 (95% CI 0.57 to 0.81). The specificity of VibraTip was 0.82 (95% CI 0.74 to 0.90), the 10g monofilament was 0.83 (95% CI 0.75 to 0.91), and the 128 Hz tuning fork was 0.90 (0.84 to 0.97). These results had overlapping confidence intervals (see Figure 3.5), indicating that the tests might be diagnostically equivalent. However, there was no indication that VibraTip was diagnostically superior to either device, and it was noted that the 10g monofilament in fact gave better results in terms of both sensitivity and specificity, although statistical superiority cannot be claimed.

The Bracewell study had two important limitations which the EAC considers were potential sources of uncertainty. Firstly, the diagnostic accuracy

parameters for each device appeared to have been optimised using ROC analysis, with 2/10 sites testing insensate by VibraTip indicating the presence of DPN (2/10 for the 10g monofilament and 1/10 for the 128 Hz tuning fork). It is unclear how generalisable this is to usual clinical practice. Secondly, the study was relatively small, with 141 participants. This gives rise to the possibility of a type 2 error, where there was a difference in the diagnostic accuracy of the devices, but the study was not adequately powered to detect it. This could be significant, because even small differences in diagnostic accuracy might have important consequences concerning the post-diagnostic pathway for at risk diabetic feet.

The sponsor and the EAC did not identify any existing economic studies on VibraTip. The sponsor provided the EAC with a *de novo* economic model, which was based on a decision tree that mapped patient pathways over a 3 year time horizon. The model had two arms, a current practice arm (which in the base case was 10g monofilament) and an intervention arm, where patients were tested with VibraTip (40% base case) or retained practice (60% 10g monofilament base case). Patients with diabetes (population 2.9 million) were modelled with transition stages and outcomes including being classed as high risk of DPN (requiring more frequent testing), developing ulceration and continued ulcers, and requiring amputation. The sponsor performed deterministic sensitivity testing where the adoption uptake of VibraTip was assumed to be 20% or 100%, and this was combined with a second sensitivity analysis, whereby an assumption was made that the introduction of VibraTip was associated with a 1% reduction in ulcer formation.

The sponsor reported that in the base case, over 3 years, the introduction of VibraTip would lead to overall cost savings to the NHS of approximately £50,000 compared with the 10g monofilament and £40,000 compared with the 128 Hz tuning fork. The differences arise from the different per-examination device costs; all clinical outcomes and associated costs were identical.

For sensitivity analysis, the sponsor reported that increased adoption of VibraTip would lead to proportionate increases in savings, as would be expected. If it was assumed that the use of VibraTip would lead to a 1% reduction in ulcer formation, then VibraTip would lead to savings of £6,430,000 compared with the 10g monofilament and £6,350,000 compared with the 128 Hz tuning fork, at the level of 40% adoption and over the three years.

The EAC considered that in the base case, the model was a *de facto* cost minimisation study because the sponsor had implicitly assumed the clinical evidence showed non–inferiority across the devices. The EAC agrees with

this interpretation of the clinical evidence but notes the uncertainty around all the clinical studies.

Hence much of the functionality in the model was not required. The EAC has identified various issues with the model's structure and poor sensitivity analysis (see section 4.2.6) but these are not rehearsed because they do not inform the results required to answer the decision problem.

The EAC performed some further work on estimating the per-examination costs of the devices by using the technical papers by Horsfield and Levy (2012) (4) and Lavery *et al.* (2011) (12). The principal limitation in the useful life of a VibraTip is its characteristics as its battery discharges. The EAC calculated that the per-examination cost of VibraTip ranged between 0.2 pence and 10.0 pence, with a value of between 4.0 pence and 10.0 pence being most plausible. This compared with an estimated per-examination cost of the 10g monofilament (in which the plasticity of the filament degrades) of 3.04 pence to 7.60 pence. No equivalent measures were available for the 128 Hz tuning fork which has an indefinite useful life. Thus, the EAC concludes that the economic case for VibraTip has not been made.

The EAC noted that VibraTip has some potential advantages over the other technologies, particularly the 128 Hz tuning fork (6). These include its size and portability, ease of use, and reduced auditory and temperature cueing. None of these claims has quantitative evidence to support them, and there is therefore no evidence of demonstrable benefits to patients or the NHS. In comparison, the advantages of the 10g monofilament and the 128 Hz tuning fork mainly concern healthcare professionals' familiarity with the devices, particularly the 10g monofilament, which is widely used in primary care, has a consistent amplitude (over 1000 uses), and a lesser requirement for training.

It has been demonstrated previously that early detection of feet at risk of diabetic ulceration is both beneficial to patients and cost-effective to healthcare services (69, 70). Thus, the potential benefits of VibraTip will be dependent on its diagnostic accuracy compared with the present technologies employed. However, the EAC considers at present the available clinical and economic evidence for VibraTip is equivocal.

6 Implications for research

The 10g monofilament and the 128 Hz tuning fork are established diagnostic devices for the detection of DPN and are currently recommended by national guidelines (18). Comparative research of VibraTip and its comparator technologies is currently lacking, therefore the EAC would recommend that further primary research is carried out.

As the benefit from early detection of feet at risk of diabetic ulceration is not disputed, the EAC would recommend that research should focus on the diagnostic accuracy of VibraTip compared with the 10g monofilament and the 128 Hz tuning fork. Such a study should ideally have the following characteristics:

- Cross-sectional diagnostic design.
- Sequential or random recruitment from a population of people with diabetes, but with no previous history of DPN or ulceration.
- Adequately powered with predefined outcomes, estimate of significant clinical effect, and appropriate power calculations.
- Predefined thresholds for detection of DPN in line with standard practice.
- Use of neurothesiometer with a threshold set at 25 V as the reference standard. Preferably, the operator of the tests should be blind to these results.

Another source of uncertainty the EAC encountered during the evaluation of VibraTip was the lack of standardisation of technique when applying VibraTip or its comparators, relating to which are the most appropriate sites on the foot, how many sites should be tested, and how many insensate sites are required for a positive test of DPN. This could be achieved through suitably designed diagnostic study using ROC analysis. Ideally, standardised practice could be implemented.

References

1. Bowling FL, Abbott CA, Harris WE, Atanasov S, Malik RA, Boulton AJ. A pocket-sized disposable device for testing the integrity of sensation in the outpatient setting. Diabetic Medicine. 2012;29(12):1550-2.
2. Bracewell N, Game F, Jeffcoate W, Scammell BE. Clinical evaluation of a new device in the assessment of peripheral sensory neuropathy in diabetes. Diabetic Medicine. 2012;29(12):1553-5.
3. Garbas G, Urbancic-Rovan V, Slak M, Hohnjec M. Can we replace 128-Hz tuning fork testing by Vibratip®. Diabetic Foot Study Group; 2013; Sitges, Spain.
4. Horsfield I, Levy A. VibraTip durability in clinical practice: How long does it last? British Journal of Diabetes and Vascular Disease. 2013;13(2):91-5.
5. Horsfield I, Levy A. Clinical consistency of vibration sense testing: development of a manikin toe with embedded accelerometer for testing and training. In Press. 2014.
6. Levy A. Preliminary data on VibraTip, a new source of standardised vibration for bedside assessment of peripheral neuropathy. British Journal of Diabetes and Vascular Disease. 2010;10(6):284-6.
7. Levy A, Greenwood R. Self care and misplaced confidence. In Press. 2014.
8. Nizar H, Munro N, Nightingale P, Feher MD. Diagnostic accuracy of the VibraTip in detection of diabetic peripheral neuropathy. British Journal of Diabetes and Vascular Disease. 2014;14(1):26-9.
9. Urbancic-Rovan V, M. S, Hihnjec M, editors. Screening for neuropathy - do additional testing methods make sense? Diabetic Foot Study Group (DFSG) 28-30 September; 2012; Berlin-Potsdam.
10. Baker N. An alternative to a 10-g monofilament or tuning fork? Two new, simple, easy-to-use screening tests for determining foot ulcer risk in people with diabetes. Diabetic Medicine. 2012;29(12):1477-9.
11. Bracewell NJ, Scammell BE, Jeffcoate JW, Game FL. Evaluating a new device in the assessment of peripheral sensory neuropathy in diabetes: the Vibratip® study. Diabetologia. 2011;54.
12. Lavery LA, Lavery DE, Lavery DC, Lafontaine J, Bharara M, Najafi B. Accuracy and durability of Semmes-Weinstein monofilaments: what is the useful service life? Diabetes Research & Clinical Practice. 2012;97(3):399-404.
13. DH. Quality and Outcomes Framework guidance for GMS contract 2013/14. In: DH, editor. London2013.
14. National Collaborating Centre for Chronic Conditions. Type 2 diabetes: National clinical guideline for management in primary and secondary care (update). London: Royal College of Physicians 2008.
15. Mooradian AD, Chehade JM. Diabetes mellitus in older adults. American Journal of Therapeutics. 2012;19(2):145-59.
16. NICE. Diabetic foot problems: final scope. London 2013.
17. Callaghan BC, Cheng HT, Stables CL, Smith AL, Feldman EL. Diabetic neuropathy: clinical manifestations and current treatments. Lancet Neurology. 2012;11(6):521-34.
18. NICE. Type 2 diabetes foot problems: Prevention and management of foot problems. London: National Institute for Health and Care Excellence; 2004.

19. DCCT. The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. The Diabetes Control and Complications Trial Research Group. New England Journal of Medicine. 1993;329(14):977-86.
20. Boucek P. Advanced Diabetic Neuropathy: A Point of no Return? The Review of Diabetic Studies. 2006;3(3):143-50.
21. Vinik AI, Strotmeyer ES, Nakave AA, Patel CV. Diabetic neuropathy in older adults. Clinics in Geriatric Medicine. 2008;24(3):407-35, v.
22. Vinik AI, Suwanwalaikorn S, Stansberry KB, Holland MT, McNitt PM, Colen LE. Quantitative measurement of cutaneous perception in diabetic neuropathy. Muscle & Nerve. 1995;18(6):574-84.
23. Gilman S. Joint position sense and vibration sense: anatomical organisation and assessment. Journal of Neurology, Neurosurgery & Psychiatry. 2002;73(5):473-7.
24. Booth J, Young MJ. Differences in the performance of commercially available 10-g monofilaments. Diabetes Care. 2000;23(7):984-8.
25. NICE. Type 1 diabetes: diagnosis and management of type 1 diabetes in children, young people and adults. London 2004.
26. SIGN. The management of diabetes: a national guideline. Edinburgh: Scottish Intercollegiate Guidelines Network; 2010.
27. NICE. Type 2 diabetes: Prevention and management of foot problems. Quick reference guide. London 2004.
28. School of Health and realted Research (ScHARR). Clinical Guidelines for Type 2 Diabetes: Prevention and management of foot problems (revised version). London: NICE; 2004.
29. Pham H, Armstrong DG, Harvey C, Harkless LB, Giurini JM, Veves A. Screening techniques to identify people at high risk for diabetic foot ulceration: a prospective multicenter trial. Diabetes Care. 2000;23(5):606-11.
30. Jude E. A Study to Evaluate the Performance of VibraTip® by Different Clinical Users (NCT01878682). (Not yet recruiting) 2014.
31. Partanen J, Niskanen L, Lehtinen J, Mervaala E, Siitonen O, Uusitupa M. Natural history of peripheral neuropathy in patients with non-insulin-dependent diabetes mellitus. New England Journal of Medicine. 1995;333(2):89-94.
32. McCallan Medical Ltd. VibraTip Press Pack2014 [cited 2014 10th March]; www.vibratip.com. Available from: http://www.vibratip.com/pdf/vibratip_press_pack_v2010.pdf.
33. FRAME. Diabetes Foot Screening. 2011 [cited 2014 26th March]; Available from: http://www.diabetesframe.org/
34. IGWGDF. International Working Group for Diabetic Foot : Diabetic neuropathy. 2012 [cited 2014 March 26th]; Available from: http://iwgdf.org/consensus/diabetic-neuropathy/
35. ADA. American Diabetes Association: Clinical Practice Recommendations. 2014 [cited 2014 26th March]; Available from: http://www.diabetes.org/research-and-practice/we-support-your-doctor/clinical-practice-recommendations.html?loc=rp-slabnav.

36. Edmonds ME, Blundell MP, Morris ME. Improved survival of the diabetic foot: The role of a specialised foot clinic. Quarterly Journal of Medicine. 1986;60(232):763-71.
37. Gefen A. Technologies for detecting loss of multiple sensory modalities in diabetic foot neuropathy. Diabetes Management. 2012;2(2):149-55.
38. Buckley L, Docter S, Morona J, Reddin E, Mnatzaganina G, Junega V, et al. Prevention, identification and management of diabetic foot complications. School of population health and clinical practice. Adelaide 2011.
39. Dros J, Wewerinke A, Bindels PJ, van HCH. Accuracy of monofilament testing to diagnose peripheral neuropathy: a systematic review. Annals of Family Medicine. 2009;7(6):555-8.
40. Kanji JN, Anglin RE, Hunt DL, Panju A. Does this patient with diabetes have large-fiber peripheral neuropathy? JAMA. 2010;303(15):1526-32.
41. Feng Y, Schlosser FJ, Sumpio BE. The Semmes Weinstein monofilament examination is a significant predictor of the risk of foot ulceration and amputation in patients with diabetes mellitus. Journal of Vascular Surgery. 2011;53(1):220-6.e1-5.
42. Young MJ, Breddy JL, Veves A, Boulton AJ. The prediction of diabetic neuropathic foot ulceration using vibration perception thresholds. A prospective study. Diabetes Care. 1994;17(6):557-60.
43. Mathias V. Good vibrations will help diabetics lead a better life. Western Daily Press. 2010.
44. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. BMJ (Online). 2009;339(7716):332-6.
45. Rayman G, Vas PR, Baker N, Taylor CG, Jr., Gooday C, Alder AI, et al. The Ipswich Touch Test: a simple and novel method to identify inpatients with diabetes at risk of foot ulceration. Diabetes Care. 2011;34(7):1517-8.
46. Viswanathan V, Snehalatha C, Seena R, Ramachandran A. Early recognition of diabetic neuropathy: evaluation of a simple outpatient procedure using thermal perception. Postgraduate Medical Journal. 2002;78(923):541-2.
47. Yoshioka K, Okada H. Useful application of the neuropad test for assessment of diabetic polyneuropathy. Internal Medicine. 2012;51(23):3241-5.
48. Weintrob N, Amitay I, Lilos P, Shalitin S, Lazar L, Josefsberg Z. Bedside neuropathy disability score compared to quantitative sensory testing for measurement of diabetic neuropathy in children, adolescents, and young adults with type 1 diabetes. Journal of Diabetes & its Complications. 2007;21(1):13-9.
49. Bandolier. Bandolier bias guide. 2001 [cited 2001 4th April]; Available from: http://www.medicine.ox.ac.uk/bandolier/Extraforbando/Bias.pdf.
50. Schueler S, Schuetz GM, Dewey M. The revised QUADAS-2 tool. Annals of Internal Medicine. 2012;156(4):323; author reply -4.
51. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Annals of Internal Medicine. 2011;155(8):529-36.
52. NICE. The Guidelines Manual. London: s2009.
53. NICE. Diagnostics Assessment Programme manual. In: Excellence NIfHaC, editor. London 2011.

54. Rutjes AW, Reitsma JB, Di M, Smidt N, van JCJ, Bossuyt PM. Evidence of bias and variation in diagnostic accuracy studies. CMAJ Canadian Medical Association Journal. 2006;174(4):469-76.
55. Lijmer JG, Mol BW, Heisterkamp S, Bonsel GJ, Prins MH, van JHJ, et al. Empirical evidence of design-related bias in studies of diagnostic tests. JAMA. 1999;282(11):1061-6.
56. Kelly S, Berry E, Roderick P, Harris KM, Cullingworth J, Gathercole L, et al. The identification of bias in studies of the diagnostic performance of imaging modalities. British Journal of Radiology. 1997;70 6(838):1028-35.
57. Altman DG, Bland JM. Diagnostic tests 2: Predictive values. BMJ. 1994;309(6947).
58. Kastenbauer T, Sauseng S, Brath H, Abrahamian H, Irsigler K. The value of the Rydel-Seiffer tuning fork as a predictor of diabetic polyneuropathy compared with a neurothesiometer. Diabetic Medicine. 2004;21(6):563-7.
59. Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. BMJ. 1996;313(7052):275-83.
60. Petrou S, Gray A. Economic evaluation using decision analytical modelling: design, conduct, analysis, and reporting. BMJ. 2011;342.
61. Coppini DV, Wellmer A, Weng C, Young PJ, Anand P, Sonksen PH. The natural history of diabetic peripheral neuropathy determined by a 12 year prospective study using vibration perception thresholds. Journal of Clinical Neuroscience. 2001;8(6):520-4.
62. Ramsey SD, Newton K, Blough D, McCulloch DK, Sandhu N, Reiber GE, et al. Incidence, outcomes, and cost of foot ulcers in patients with diabetes. Diabetes Care. 1999;22(3):382-7.
63. Boulton AJ, Kirsner RS, Vileikyte L. Clinical practice. Neuropathic diabetic foot ulcers. New England Journal of Medicine. 2004;351(1):48-55.
64. Jeffcoate WJ, Harding KG. Diabetic foot ulcers. Lancet. 2003;361(9368):1545-51.
65. NICE. Costing statement: ‘Diabetic foot problems: inpatient management of diabetic foot problems’. London 2011.
66. Taylor M. What is sensitivity analysis? 2009 [cited 2014 10th April]; Available from: http://www.medicine.ox.ac.uk/bandolier/painres/download/whatis/What_is_sens_analy.pdf.
67. Health & Social Care Information Centre (HSCIC). Quality and Outcomes Framework. Health & Social Care Information Centre; 2014 [cited 2014 15th April]; Available from: http://qof.hscic.gov.uk/index.asp.
68. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. Methods for the economic evaluation of healthcare programmes: Oxford University Press; 2005.
69. Li QY, Tang J, He EH, Li YM, Zhou Y, Zhang X, et al. Clinical utility of three-dimensional contrast-enhanced ultrasound in the differentiation between noninvasive and invasive neoplasms of urinary bladder. European Journal of Radiology. 2012;81(11):2936-42.

70. Ortegon MM, Redekop WK, Niessen LW. Cost-effectiveness of prevention and treatment of the diabetic foot: a Markov analysis. Diabetes Care. 2004;27(4):901-7.

Appendix 1. Supplementary information of literature searches performed by EAC.

The databases and information sources searched are reported in Table A1.

Table A1 *Databases and information sources searched by EAC.*

Database / information source	Interface / URL
MEDLINE and MEDLINE In-Process	OvidSP
Embase	OvidSP
Cochrane Database of Systematic Reviews (CDSR)	Cochrane Library/Wiley Interscience
Cochrane Central Register of Controlled Trials (CENTRAL)	Cochrane Library/Wiley Interscience
Database of Abstracts of Reviews of Effectiveness (DARE)	Cochrane Library/Wiley Interscience
Health Technology Assessment Database (HTA)	Cochrane Library/Wiley Interscience
NHS Economic Evaluation Database (NHS EED)	Cochrane Library/Wiley Interscience
ClinicalTrials.gov	http://www.clinicaltrials.gov
WHO International Clinical Trials Registry Platform (ICTRP)	http://apps.who.int/trialsearch/
metaRegister of Controlled Trials (mRCT)	http://www.controlled-trials.com/mrct/
Diabetes UK conference	See search details below
Society of Chiropodists & Podiatrists Annual Conference	See search details below
Diabetic Foot Study Group (DFSG) Proceedings	See search details below
Malvern Diabetic Foot Conference	See search details below
Diabetic Foot Global Conference	See search details below
International Diabetic Foot Conference	See search details below
American Diabetes Association	See search details below
International Diabetes Federation	See search details below
BioMedSearch.com	http://www.biomedsearch.com/
VibraTip website	http://www.vibratip.com/
Association of British Clinical Diabetologists website	http://www.diabetologists-abcd.org.uk/home.htm
Association of British Neurologists website	http://www.theabn.org/
British Peripheral Nerve Society website	http://www.bpns.org.uk/
British Society for Clinical	http://www.bscn.org.uk/

Database / information source	Interface / URL
Neurophysiology website	
Diabetes UK website	https://www.diabetes.org.uk/
Primary Care Diabetes Society website	http://www.pcdsociety.org/
Society for Endocrinology website	https://www.endocrinology.org/index.aspx
Society of Chiropodists & Podiatrists website	http://www.scpod.org/
The Neurological Alliance website	http://www.neural.org.uk/
Google	https://www.google.co.uk/

The searches identified 1034 records (Table A2). Following deduplication 735 records were assessed for relevance.

Table A2. *Records identified by EAC*.

Resource	Records identified
MEDLINE and MEDLINE In-Process	288
Embase	632
Cochrane Database of Systematic Reviews (CDSR)	3
Cochrane Central Register of Controlled Trials (CENTRAL)	37
Database of Abstracts of Reviews of Effectiveness (DARE)	2
Health Technology Assessment Database (HTA)	1
NHS Economic Evaluation Database (NHS EED)	0
ClinicalTrials.gov	33
WHO International Clinical Trials Registry Platform (ICTRP)	23
metaRegister of Controlled Trials (mRCT)	0
Diabetes UK conference	4
Society of Chiropodists & Podiatrists Annual Conference	3
Diabetic Foot Study Group (DFSG) Proceedings	4
Malvern Diabetic Foot Conference	0
Diabetic Foot Global Conference	0
International Diabetic Foot Conference	1
American Diabetes Association	0
International Diabetes Federation	0
BioMedSearch.com	0
VibraTip website	1
Association of British Clinical Diabetologists website	0
Association of British Neurologists website	0
British Peripheral Nerve Society website	0
British Society for Clinical Neurophysiology website	0
Diabetes UK website	0
Primary Care Diabetes Society website	0
Society for Endocrinology website	0

Resource	Records identified
Society of Chiropodists & Podiatrists website	0
The Neurological Alliance website	0
Google	2
TOTAL	**1034**
TOTAL after deduplication	**735**

Search strategy details

1: Source: MEDLINE(R) In-Process & Other Non-Indexed Citations and MEDLINE(R) 1946 to Present

Interface / URL: OvidSP
Search date: 10/03/14
Retrieved records: 288
Search strategy:

1 exp Diabetes Mellitus/ 307249
2 diabet$.ti,ab,kf,jw. 427229
3 (dpn or iddm or dm1 or mody or dm2 or niddm or iidm or dm).ti,ab,kf. 42824
4 or/1-3 488471
5 Vibration/ 19511
6 (vibrat$ or vibrotactil$ or vibro-tactil$).ti,ab,kf. 47725
7 or/5-6 53175
8 4 and 7 983
9 (vibratip$ or vibra tip$).ti,ab,kf. 2
10 8 or 9 983
11 exp animals/ not humans/ 3898900
12 10 not 11 968
13 limit 12 to (english language and yr="2007 -Current") 289
14 remove duplicates from 13 288

2: Source: Embase 1974 to 2014 March 06

Interface / URL: OvidSP
Search date: 07/03/14
Retrieved records: 632
Search strategy:

1 exp diabetes mellitus/ 607284
2 diabet$.ti,ab,kw,jx. 614048

3 (dpn or iddm or dm1 or mody or dm2 or niddm or iidm or dm).ti,ab,kw. 61111
4 or/1-3 756575
5 vibration/ 23774
6 vibration sense/ 2122
7 (vibrat$ or vibrotactil$ or vibro-tactil$).ti,ab,kw. 43484
8 or/5-7 49805
9 4 and 8 1485
10 (vibratip$ or vibra tip$).ti,ab,kw,dm,dv. 7
11 9 or 10 1486
12 (animal/ or animal experiment/ or animal model/ or nonhuman/) not exp human/ 5133879
13 11 not 12 1470
14 limit 13 to (english language and yr="2007 -Current") 653
15 remove duplicates from 14 632

3: Source: Cochrane Database of Systematic Reviews (CDSR) - Issue 3 of 12, March 2014

Interface / URL: Cochrane Library/Wiley Interscience
Search date: 10/03/14
Retrieved records: 3
Search strategy:

#1 MeSH descriptor: [Diabetes Mellitus] explode all trees 15413
#2 diabet*:ti,ab,kw 29439
#3 (dpn or iddm or dm1 or mody or dm2 or niddm or iidm or dm):ti,ab,kw 4593
#4 #1 or #2 or #3 31761
#5 MeSH descriptor: [Vibration] this term only 598
#6 (vibrat* or vibrotactil* or vibro-tactil*):ti,ab,kw 1209
#7 #5 or #6 1209
#8 #4 and #7 136
#9 (vibratip* or vibra next tip*) 2
#10 #8 or #9 from 2007 to 2014 34
#11 #10 in Cochrane Reviews (Reviews and Protocols) 3

4: Source: Database of Abstracts of Reviews of Effectiveness (DARE) - Issue 1 of 4, Jan 2014

Interface / URL: Cochrane Library/Wiley Interscience
Search date: 10/03/14
Retrieved records: 2

Search strategy:

#1 MeSH descriptor: [Diabetes Mellitus] explode all trees 15413
#2 diabet* 34309
#3 (dpn or iddm or dm1 or mody or dm2 or niddm or iidm or dm) 13422
#4 #1 or #2 or #3 44696
#5 MeSH descriptor: [Vibration] this term only 598
#6 (vibrat* or vibrotactil* or vibro-tactil*) 1401
#7 #5 or #6 1401
#8 #4 and #7 221
#9 (vibratip* or vibra next tip*) 2
#10 #8 or #9 from 2007 to 2014 103
#11 #10 in Other Reviews 2

5: Source: Health Technology Assessment Database (HTA) - Issue 1 of 4, Jan 2014

Interface / URL: Cochrane Library/Wiley Interscience
Search date: 10/03/14
Retrieved records: 1
Search strategy:

#1 MeSH descriptor: [Diabetes Mellitus] explode all trees 15413
#2 diabet* 34309
#3 (dpn or iddm or dm1 or mody or dm2 or niddm or iidm or dm) 13422
#4 #1 or #2 or #3 44696
#5 MeSH descriptor: [Vibration] this term only 598
#6 (vibrat* or vibrotactil* or vibro-tactil*) 1401
#7 #5 or #6 1401
#8 #4 and #7 221
#9 (vibratip* or vibra next tip*) 2
#10 #8 or #9 from 2007 to 2014 103
#11 #10 in Technology Assessments 1

6: Source: NHS Economic Evaluation Database (NHS EED) - Issue 1 of 4, Jan 2014

Interface / URL: Cochrane Library/Wiley Interscience
Search date: 10/03/14
Retrieved records: 0
Search strategy:

#1 MeSH descriptor: [Diabetes Mellitus] explode all trees 15413
#2 diabet* 34309
#3 (dpn or iddm or dm1 or mody or dm2 or niddm or iidm or dm) 13422
#4 #1 or #2 or #3 44696
#5 MeSH descriptor: [Vibration] this term only 598
#6 (vibrat* or vibrotactil* or vibro-tactil*) 1401
#7 #5 or #6 1401
#8 #4 and #7 221
#9 (vibratip* or vibra next tip*) 2
#10 #8 or #9 from 2007 to 2014 103
#11 #10 in Economic Evaluations 0

7: Source: Cochrane Central Register of Controlled Trials (CENTRAL) - Issue 1 of 12, Jan 2014

Interface / URL: Cochrane Library/Wiley Interscience
Search date: 10/03/14
Retrieved records: 37
Search strategy:

#1 MeSH descriptor: [Diabetes Mellitus] explode all trees 15413
#2 diabet* 34309
#3 (dpn or iddm or dm1 or mody or dm2 or niddm or iidm or dm) 13422
#4 #1 or #2 or #3 44696
#5 MeSH descriptor: [Vibration] this term only 598
#6 (vibrat* or vibrotactil* or vibro-tactil*) 1401
#7 #5 or #6 1401
#8 #4 and #7 221
#9 (vibratip* or vibra next tip*) 2
#10 #8 or #9 from 2007 to 2014 103
#11 #10 in Trials 37

8: Source: ClinicalTrials.gov

Interface / URL: http://www.clinicaltrials.gov/ct2/home
Search date: 10/03/14
Retrieved records: 33
Search strategy:

Advanced search interface used (http://clinicaltrials.gov/ct2/search/advanced). Following searches carried out - search terms entered in the 'Search terms' box:

1. vibratip OR vibratips OR "vibra tip" OR "vibra tips" = 1 result

2. (diabetes OR diabetic OR diabetics OR dpn OR iddm OR dm1 OR mody OR dm2 OR niddm OR iidm OR dm) AND (vibration OR vibrations OR vibrate OR vibrates OR vibrated OR vibrating OR vibratory OR vibrotactile OR "vibro-tactile") = 32

3. (diabetes OR diabetic OR diabetics OR dpn OR iddm OR dm1 OR mody OR dm2 OR niddm OR iidm OR dm) AND (vibratometer OR vibratometers OR vibratometry) = 0 results

9: Source: International Clinical Trials Registry Platform (ICTRP)
Interface / URL: http://apps.who.int/trialsearch/
Search date: 10/03/14
Retrieved records: 23
Search strategy:

Basic search interface at the above url used. Following searches carried out:

1. vibratip* OR vibra tip* = 0 results

2. diabet* AND vibrat* OR diabet* AND vibrotactil* OR diabet* AND vibro-tactil* = 10 (10 records for 10 trials)

3. dpn AND vibrat* OR dpn AND vibrotactil* OR dpn AND vibro-tactil* = 0 results

4. iddm AND vibrat* OR iddm AND vibrotactil* OR iddm AND vibro-tactil* = 4

5. dm1 AND vibrat* OR dm1 AND vibrotactil* OR dm1 AND vibro-tactil* = 0 results

6. mody AND vibrat* OR mody AND vibrotactil* OR mody AND vibro-tactil* = 4

7. dm2 AND vibrat* OR dm2 AND vibrotactil* OR dm2 AND vibro-tactil* = 0 results

8. niddm AND vibrat* OR niddm AND vibrotactil* OR niddm AND vibro-tactil* = 4

9. iidm AND vibrat* OR iidm AND vibrotactil* OR iidm AND vibro-tactil* = 0 results

10. dm AND vibrat* OR dm AND vibrotactil* OR dm AND vibro-tactil* = 1 result

10: Source: metaRegister of Controlled Trials (mRCT)
Interface / URL: http://www.controlled-trials.com/mrct/
Search date: 10/03/14
Retrieved records: 0
Search strategy:

All registers searched apart from the NIH ClinicalTrials.gov Register – searched above. Following searches carried out:

1. vibratip OR vibratips OR "vibra tip" OR "vibra tips" = 0 results

2. vibration OR vibrations OR vibrate OR vibrates OR vibrated OR vibrating OR vibratory OR vibrotactile OR vibro-tactile OR vibratometer OR vibratometers OR vibratometry = 0 results downloaded (20 results retrieved by search and assessed for relevance online – no relevant identified)

Conference abstract searches

11: Source: Diabetes UK conference
Search date: 11/03/14
Retrieved records: 4
Search strategy:

2014: Diabetes UK Professional Conference 5-7 March 2014 ACC Liverpool

Diabetic Medicine. March 2014 Volume 31, Issue s1, Page 1-195. Special Issue: Abstacts of the Diabetes UK Professional Conference 2014, Arena and Convention Centre, Liverpool, UK, 5–7 March 2014. http://onlinelibrary.wiley.com/doi/10.1111/dme.2014.31.issue-3/issuetoc

Ctrl-f search for the term 'vibr' was carried out across the pdf documents listed below. Abstracts containing the terms were assessed for relevance / checked against existing records in EndNote for duplication.

http://onlinelibrary.wiley.com/doi/10.1111/dme.12377_1/pdf
http://onlinelibrary.wiley.com/doi/10.1111/dme.12377_2/pdf
http://onlinelibrary.wiley.com/doi/10.1111/dme.12377_3/pdf
http://onlinelibrary.wiley.com/doi/10.1111/dme.12377_4/pdf
http://onlinelibrary.wiley.com/doi/10.1111/dme.12377_5/pdf
http://onlinelibrary.wiley.com/doi/10.1111/dme.12377_6/pdf
http://onlinelibrary.wiley.com/doi/10.1111/dme.12377_7/pdf
http://onlinelibrary.wiley.com/doi/10.1111/dme.12377_8/pdf
http://onlinelibrary.wiley.com/doi/10.1111/dme.12377_9/pdf
http://onlinelibrary.wiley.com/doi/10.1111/dme.12377_10/pdf
http://onlinelibrary.wiley.com/doi/10.1111/dme.12377_11/pdf
http://onlinelibrary.wiley.com/doi/10.1111/dme.12377_12/pdf
http://onlinelibrary.wiley.com/doi/10.1111/dme.12377_13/pdf
http://onlinelibrary.wiley.com/doi/10.1111/dme.12377_14/pdf
http://onlinelibrary.wiley.com/doi/10.1111/dme.12378_1/pdf - 1 result
http://onlinelibrary.wiley.com/doi/10.1111/dme.12378_2/pdf - 3 results
http://onlinelibrary.wiley.com/doi/10.1111/dme.12408/pdf

2013 abstracts are indexed in Embase – searched via the Embase search detailed above

2012 abstracts are indexed in Embase – searched via the Embase search detailed above

2011 abstracts are indexed in Embase – searched via the Embase search detailed above

12: Source: Society of Chiropodists & Podiatrists Annual Conference

Search date: 13/03/14 – 14/03/14; 17/03/14
Retrieved records: 3
Search strategy:

Note: 2014 meeting has not been held at the time of searching (to be held in November).

2013: Society of Chiropodists & Podiatrists 2013 Annual Conference and Exhibition, Liverpool

Titles of presentations in the SCP Delegate Brochure and Programme Nov 2013 (http://www.scpod.org/easysiteweb/getresource.axd?assetid=37987&type=0&servicetype=1)were scanned for potentially relevant abstracts using a Ctrl F search on each of the following terms 'diabet', 'vibr' and 'neuropath'. The pdf for those identified (provided by SCP Communications Officer) was viewed and a Ctrl F search for 'vibr' carried out across the document.

0 results selected.

Communications Officer subsequently provided a copy of the 2013 conference handbook. Ctrl-F search on the term 'vibr' carried out across the Conference Handbook (provided by SCP Communications Officer). Identified results assessed for relevance.

0 results selected.

2012: Society of Chiropodists & Podiatrists Annual Conference 2012 - Glasgow, Scotland at the SECC on 11 - 13 October.

Ctrl-F search on the term 'vibr' carried out across the Conference Handbook (provided by SCP Communications Officer). Identified results were assessed for relevance.

0 results selected.

2011: Society of Chiropodists & Podiatrists Annual Conference 2011 – Harrogate

Titles of presentations in the SCP Delegate Brochure and Programme 2011 (provided by SCP Communications Officer) were scanned for potentially relevant abstracts using a Ctrl F search on each of the following terms 'diabet', 'vibr' and 'neuropath'. No abstracts are available.

Possibly relevant titles selected: 3

Communications Officer subsequently provided a copy of the conference handbook. Ctrl-F search on the term 'vibr' carried out across the Conference

Handbook (provided by SCP Communications Officer). Identified results were assessed for relevance.

0 results selected.

13: Source: Diabetic Foot Study Group (DFSG) meetings
Search date: 13/03/14
Retrieved records: 4
Search strategy:

Note: 2014 meeting has not been held at the time of searching (to be held in September 2014). No meeting was held in 2011.

2013: http://dfsg.org/previous-meetings-and-abstracts/abstract-2013.html

Titles scanned for potential relevance. Full abstract was viewed for titles containing the terms, which were judged potentially relevant. These were assessed for relevance / checked against existing records in EndNote for duplication. = 2 results

2012: http://new.dfsg.org/previous-meetings-and-abstracts/abstract-2012.html

Titles scanned for potential relevance. Full abstract was viewed for titles containing the terms, which were judged potentially relevant. These were assessed for relevance / checked against existing records in EndNote for duplication. = 2 results.

14: Source: Malvern Diabetic Foot Conference
Search date: 13/03/14
Retrieved records: 0
Search strategy:

Note: 2014 conference (the 15th) has not been held at the time of searching (to be held in May 2014). No conference was held in 2013 or 2011.

2012: 14th Malvern Diabetic Foot Conference - 9th - 11th May

No conference abstracts found online. Contacted organisers who sent pdf of programme with the abstracts. A ctrl-f search for the term ‘vibr’ was carried out across the pdf documents.

0 relevant records identified.

15: Source: Diabetic Foot Global Conference
Search date: Not searched
Retrieved records: n/a
Search strategy: n/a

Note: 2014 conference (the 15th) has not been held at the time of searching (to be held March 22nd). No abstracts found online for 2013, 2012, 2011. Communicated by e-mail with member of the conference organising company who confirmed that abstracts from past conferences are not made available.

16: Source: International Diabetic Foot Conference
Search date: 13/03/14
Retrieved records: 1
Search strategy:

Note: No 2014 conference held as yet. No abstracts found online for 2012, 2011. E-mailed conference organisers twice. No reply as of 01/04/14.

2013: 4th International Diabetic Foot Conference (IDFC), 14-15 November, 2013

Only source of abstracts that can be found online is the Conference Brochure, available as an ISSUU document. Unable to find any functionality for searching terms across document. Body text in document is difficult to read – only able to assess the titles. Titles assessed.

1 possible abstract of interest identified

17: Source: American Diabetes Association
Search date: 17/03/14
Retrieved records: 0
Search strategy:

Note: 2014 conference not held as yet (to be held in June 2014)

2013: Used advanced search function at http://www.abstractsonline.com/Plan/AdvancedSearch.aspx to conduct the searches below. Identified results assessed for relevance.

Session type: ALL
Session Title Keyword: vibr*
Results = 0

Session type: ALL
Abstract Title: vibr*
Results = 2 assessed, 0 selected.

Session type: ALL
Keyword: vibr*
Results = 2 assessed, 0 selected.

E-poster portal also search at: http://ada.apprisor.org/login.cfm using the following:

Poster topic: all
Vibr in second search box
Date limit: 2013
Results = 9 assessed, 0 selected

2012: abstracts indexed in Embase – search via the Embase search above.

2011: abstracts indexed in Embase – search via the Embase search above.

18: Source: International Diabetes Federation
Search date: 17/03/14
Retrieved records: 0
Search strategy:

Note: 2014 conference not held.

2013: searched using abstract search function at: http://conference2.idf.org/MEL2013/World%20Diabetes%20Congress%202013/data/HtmlApp/main.html#open-sessions. Searched on the following term:

vibr = 12 results assessed (0 selected)

2012: no conference held

2011: scientific programme searched via: http://conference2.idf.org/dubai2011/CM.NET.WebUI/CM.NET.WEBUI.SCPR/

SCPRsearchpage.aspx?conferenceid=05000000-0000-0000-0000-000000000001

Searched following in keyword field:

vibr = 0 results

neuropathy = 12 sets of results. Went into each set and carried out a ctrl-f search on 'neuropath'. Assessed each identified result for relevance.

0 results selected.

19: Source: BioMedSearch.com
Search date: 17/03/14
Interface / URL: http://www.biomedsearch.com/
Retrieved records: 0
Search strategy:

Advanced search used at: http://www.biomedsearch.com/search.html. Selected collections = Science Docs / Articles.

Search 1: vibratip* OR "vibra tip*" = 0 results

Search 2: (diabet*[TIAB] OR dpn[TIAB] OR iddm[TIAB] OR dm1[TIAB] OR mody[TIAB] OR dm2[TIAB] OR niddm[TIAB] OR iidm[TIAB] OR dm[TIAB]) AND vibr*[TIAB] = 13 results assessed for relevance – all excluded as irrelevant or duplicates.

Website searches

20: Source: Vibratip website
Interface / URL: http://www.vibratip.com/
Search date: 13/03/14
Retrieved records: 1
Search strategy:

The website was browsed for references on studies of Vibratip. 1 record added.

21: Source: Association of British Clinical Diabetologists website
Interface / URL: http://www.diabetologists-abcd.org.uk/home.htm

Search date: 13/03/14
Retrieved records: 0
Search strategy:

Used the 'Search documents' function on homepage to search individually on the following terms:

vibratip vibratips vibra-tip vibra-tips vibra

22: Source: Association of British Neurologists website
Interface / URL: http://www.theabn.org/
Search date: 13/03/14
Retrieved records: 0
Search strategy:

Used 'Search site' function on homepage to search individually on the following terms:

vibratip vibratips vibra-tip vibra-tips vibra

23: Source: British Peripheral Nerve Society website
Interface / URL: http://www.bpns.org.uk/
Search date: 13/03/14
Retrieved records: 0
Search strategy:

Used search function on the homepage to search on the following terms:

vibratip OR vibratips OR vibra-tip OR vibra-tips OR vibra

24: Source: British Society for Clinical Neurophysiology website
Interface / URL: http://www.bscn.org.uk/
Search date: 13/03/14
Retrieved records: 0
Search strategy:
No search function found on the site. Searched using Google with the following strings:

site:.bscn.org.uk/ vibratip
site:.bscn.org.uk/ vibratips
site:.bscn.org.uk/ "vibra-tip"

site:.bscn.org.uk/ “vibra-tips”
site:.bscn.org.uk/ vibra

25: Source: Diabetes UK website
Interface / URL: https://www.diabetes.org.uk/
Search date: 13/03/14
Retrieved records: 0
Search strategy:

Search function on homepage used to search individually on the following terms:

vibratip vibratips vibra-tip vibra-tips vibra

26: Source: Primary Care Diabetes Society website
Interface / URL: http://www.pcdsociety.org/
Search date: 14/03/14
Retrieved records: 0
Search strategy:

Used the ‘Search DOTN’ (Diabetes On The Net) search function to search individually on the following terms:

vibratip vibratips vibra-tip vibra-tips vibra

27: Source: Society for Endocrinology website
Interface / URL: https://www.endocrinology.org/index.aspx
Search date: 14/03/14
Retrieved records: 0
Search strategy:

Used search function on homepage to search individually on the following terms:

vibratip vibratips vibra-tip vibra-tips vibra

28: Source: Society of Chiropodists & Podiatrists website
Interface / URL: http://www.scpod.org/#
Search date: 14/03/14
Retrieved records: 0
Search strategy:

Used search function on homepage to search individually on the following terms:

vibratip vibratips vibra-tip vibra-tips vibra

29: Source: The Neurological Alliance website
Interface / URL: http://www.neural.org.uk/
Search date: 14/03/14
Retrieved records: 0
Search strategy:

No search function found on the site. Searched using Google with the following strings:

site:.neural.org.uk/ vibratip
site:.neural.org.uk/ vibratips
site:.neural.org.uk/ “vibra-tip”
site:.neural.org.uk/ “vibra-tips”
site:.neural.org.uk/ vibra

30: Source: Google
Interface / URL: https://www.google.co.uk/
Search date: 14/03/14
Retrieved records: 2
Search strategy:

vibratip OR vibratips filetype:pdf

Search returned ‘About 97 results’. Results scanned. Decision made by searcher based on information provided on results pages as to which results to click on and view. 2 results indicating potential study on Vibratip (not already sourced) selected.

Appendix 2. Technical studies submitted by sponsor

Two papers within the clinical evidence submission from the sponsor were technical studies (one unpublished). Whilst the EAC judged that these did not directly relate to the decision problem, they contained useful information regarding some practical issues surrounding the device, such as battery life and operator technique that are relevant to the interpretation of the clinical evidence. It was agreed with the MTEP team that the EAC should proceed to describe and appraise these studies in the narrative within the assessment report.

A2.1 Overview of methodologies and results of submitted technical studies

Horsfield and Levy 2013

The diagnostic repeatability of VibraTip in detecting DPN is dependent on the consistency of the stimulus presented to the patient.

Horsfield and Levy (2013) (4) considered whether the nature of the stimulus varies as the battery discharges through repetitive use. They measured the stimulus during three different activation pattern methods. The first was designed to quantify VibraTip battery performance under continual discharge. The second was designed to assess the change in frequency and amplitude of vibrations as the battery becomes progressively exhausted with successive activations (500 ms at 2s intervals). The third was designed to mimic the realistic usage of the device within clinical practice. In the third method, which was most relevant to the clinical studies described earlier, VibraTip was subjected to 10 half second activations with two second intervals, followed by 10 minute rests. This cycle was then repeated 350 times over a 10 hour working day for five days, allowing the device to rest overnight. Results demonstrated that over the first 1000 activations, there was no significant reduction in either frequency or amplitude of the vibrations. Amplitude fell to approximately 74% and then to 65% of the initial amplitude after 2500 and 3500 activations, respectively. The frequency of vibration, however, only declined to 94% of initial frequency after 3500 activations. The paper concludes that with routine clinical activity, *i.e.* rests between patients and overnight, VibraTip could provide consistent performance for many months and could test at least 100 patients with no significant reduction in frequency or amplitude of vibrations. With successive applications, the amplitude reduced more rapidly, with no initial plateau phase, until it fell to about 45% of initial amplitude after about 2000 applications. The authors concluded that the performance of the device depends on its pattern of use.

Horsfield and Levy 2014 (unpublished)

Both inter- and intra-operator consistency are factors that should be considered when evaluating the sensitivity of both the VibraTip and 128 Hz tuning fork. The intensity of the vibration stimulus presented to the patient and ultimately the sensitivity of the tuning fork is dependent on the striking force applied by the operator to impart the vibration and also the time elapsed since it was struck. VibraTip is claimed to deliver a fixed vibration stimulus with a frequency slightly above 128 Hz; however, the pressure applied by the operator when presenting the VibraTip to the diabetic foot may confound quantification of vibration sense and may be responsible for inter- and intra-operator inconsistency.

A set of experiments were performed by Horsfield and Levy (5) in order to determine the clinical consistency and reliability of vibration stimulus application of both VibraTip and a 128 Hz tuning fork. A manikin toe structure was modelled to mimic human bone and tissue structures and incorporated an accelerometer. This allowed the waveform and intensity of vibration applied to the surface of the toe to be recorded. Junior and senior doctors (n=21) familiar with using tuning forks but with no prior VibraTip training selected either VibraTip or the tuning fork and applied it five times in succession to the surface of the toe, before switching to the alternative device. This process was repeated, allowing 10 applications to be recorded for each doctor using each device. An oscilloscope recorded the waveform from the accelerometer for 3 seconds from the point of application of the stimulus to the manikin toe. The graphical presentation of results demonstrate that the amplitudes recorded from the application of VibraTip were generally greater than those elicited by the tuning fork. Unexpectedly, the mean frequencies measured from both devices were variable. The authors conclude that although VibraTip is as consistent a source of vibration as a standard tuning fork, reliability of the delivered vibration intensity correlates with the manner in which the devices are used. It is therefore suggested that the manikin toe could serve as a useful training aid to improve consistency of presentation of vibration stimulus to the diabetic foot and therefore improve vibration perception testing.

A2.2 Overview and critique of the sponsor's critical appraisal of technical studies

Horsfield and Levy 2013

The sponsor stated that this laboratory assessment of the durability of VibraTip was intended to find out how long the device lasts when used in a variety of different ways. The sponsor described the outcome as that the

device would last for many months of routine use without significantly altering its behaviour.

The sponsor did not provide a critical discussion regarding the limitations of the study.

Horsfield and Levy 2014 (unpublished)

The sponsor summarised the details of this unpublished study in submission Table B4. The population was defined as 21 medics - from foundation level to consultant. The intervention was described as testing consistency of vibration sense testing and the comparators were VibraTip versus 128 Hz tuning fork.

The sponsor did not provide a critical discussion regarding the limitations of this study and its generalisability to ‘real life’ patients.

A2.3 EAC’s critical appraisal of technical studies

The EAC offers a narrative appraisal of the strengths and weaknesses of the two technical studies, considering quality of reporting, potential bias, protocol design and any potential for errors.

Horsfield and Levy 2013

The study design has a number of strengths, principally in assessing three different activation patterns of the VibraTip device and measuring how these influence the expiration rate of the battery and the amplitude and frequency of the stimulus. The experimental set up and methods are well described and the findings are clearly presented.

For the activation pattern designed to represent the device in clinical use, it is unclear why a half second discharge was chosen, when the Instructions for Use submitted by the sponsor recommend a 1second activation when testing the diabetic foot. Doubling the duration of each activation would be expected to drain the battery more rapidly. It is therefore uncertain whether the conclusion that the device could be used on at least 100 patients before disposal would be realised in clinical practice.

Horsfield and Levy 2014 (unpublished)

The strengths of this study lie in the considered design of the human tissue-mimicking materials of the manikin toe in attempting to replicate the clinical assessment of the hallux of the diabetic foot.

The 21 users studied were hospital-based junior and senior doctors from the medical admissions unit and short stay ward of a single acute hospital.

Although each was described as being familiar with a tuning fork, the graphical results show that five of the 21 users (24%) applied negligible stimulus to the manikin toe. This suggests that no tuning fork vibration signal was actually applied by these five, whereas all 21 users applied a VibraTip vibration frequency of at least 200 Hz, despite not having used it before. A second user group could have been studied, including clinical users more familiar with performing routine checks for DPN, such as podiatrists. This would have provided an indication of consistency between different types of users.

Complex waveform propagation through the manikin toe is suggested by the authors as a possible reason for measured variation in frequency for the tuning fork, but this could also account for variations in the measured amplitude from both devices. Apart from applied pressure, other reasons for measurement variations include different device positioning and, for tuning forks, differences in user technique. A single marked testing position on the manikin toe would have reduced user variation due to positioning, because the accelerometer was located at one position.

The Horsfield and Levy 2013 study had previously demonstrated that the amplitude of vibration generated by the VibraTip declines over time with repeated activations. It was not reported in the 2014 study whether the same VibraTip was used by all 21 doctors and whether prior usage (and therefore remaining battery life) may have had any bearing on the resultant amplitudes of vibration measured.

A2.4 EAC's conclusions on the technical studies

Horsfield and Levy 2013

This study demonstrates that there perhaps ought to be a concern in clinical practice that the VibraTip will reach a point at which it is no longer delivering a clinically useful vibration stimulus on presentation to the diabetic foot. On questioning by the EAC, the sponsor stated that the clinical user would physically sense the loss of vibration, with the VibraTip motor 'spluttering' and thereby signalling to the user to dispose of the device and replace with new.

In addition to uncertainty about battery life, the study also demonstrates that the amplitude of the stimulus reduces significantly during the working life of the device, and that the rate of reduction depends on its duty cycle. No evidence was presented by the sponsor to demonstrate the effect of observed reductions in the stimulus amplitude (of approximately 50%) on diagnostic accuracy. The EAC concludes that such poor repeatability of the stimulus (from device to device, or within the same device at different times) is a

significant limitation to the clinical repeatability of the test, until proved otherwise. The battery state of the device was not reported in the clinical studies.

Horsfield and Levy 2014

This study demonstrates that the VibraTip delivers a consistent frequency stimulus of around 200Hz, somewhat in excess of the ‘near 128 Hz’ that is stated in product literature, as relates to the frequency of a standard tuning fork. In contrast, a neurothesiometer operates at 50Hz. The EAC therefore questioned the Expert Advisers about the relative importance of frequency and amplitude of vibration stimulus and whether there would be any clinical difference in testing for DPN with a vibration at 128 Hz (tuning fork) or a frequency of ~200Hz (Vibratip), or 50 Hz (neurothesiometer). There was not a consensus view on this aspect; therefore the EAC concludes that it is uncertain whether the operating frequency of the Vibratip elicits the same physiological response as other forms of vibration perception testing.

Appendix 3: EAC searches for economic studies

The EAC searched the following additional databases for economic evidence (Table A3.1)

Table A3.1. *Economic evidence databases.*

Database / information source	Interface / URL
Econlit	OvidSP
Health Economic Evaluation Database (HEED)	EBSCOHOST
Cost-Effectiveness Analysis (CEA) Registry	https://research.tufts-nemc.org/cear4/Home.aspx

The searches identified 0 records (Table A3.2).

Table A3.2. *Search results for economic studies.*

Resource	Records identified
Econlit	0
Health Economic Evaluation Database (HEED)	0
Cost-Effectiveness Analysis (CEA) Registry	0
TOTAL	**0**
TOTAL after deduplication	**0**

Search strategy details

1: Source: Econlit 1886 to March 2014
Interface / URL: OvidSP
Search date: 09/04/14
Retrieved records: 0
Search strategy:

1 diabet$.af. 340
2 (dpn or iddm or dm1 or mody or dm2 or niddm or iidm or dm).af. 435
3 1 or 2 770
4 (vibrat$ or vibrotactil$ or vibro-tactil$).af. 35
5 3 and 4 0
6 (vibratip$ or vibra tip$).af. 0
7 5 or 6 0

2: Source: Health Economic Evaluation Database (HEED)
Interface / URL: EBSCOHOST
Search date: 09/04/14
Retrieved records: 0

Search strategy:

S8 S7 Limiters - Published Date: 20070101-20141231 0
S7 S5 OR S6 1
S6 TX(vibratip* OR "vibra tip*") 0
S5 S3 AND S4 1
S4 TX(vibrat* OR vibrotactil* OR "vibro-tactil*") 3
S3 S1 OR S2 2,289
S2 TX(dpn OR iddm OR dm1 OR mody OR dm2 OR niddm OR iidm OR dm) 186
S1 TX diabet* 2,241

3: Source: Cost-Effectiveness Analysis (CEA) Registry
Interface / URL: https://research.tufts-nemc.org/cear4/Home.aspx
Search date: 09/04/14
Retrieved records: 0
Search strategy:

Basic search used at: https://research.tufts-nemc.org/cear4/SearchingtheCEARegistry/SearchtheCEARegistry.aspx ('search for articles' selected).

Following searches were run. Returned results were viewed online to check for diabetes context.

vibration = 0
vibrations = 0
vibrating = 0
vibrate = 0
vibrates = 0
vibrated = 0
vibratory = 0
vibrotactile = 0 (1 result returned – non diabetes context)
vibro-tactile = 0
vibro tactile = 0
vibratip = 0
vibratips = 0
vibra-tip = 0
vibra-tips = 0
vibra tip = 0
vibra tips = 0

Appendix 4: Quality assessment of sponsor's *de novo* economic model

Study name: Sponsor's *de novo* economic model		
Study design	*De novo* model	
Study question	**Response (yes/no/not clear/N/A)**	**Comments**
1. Was the research question stated?	Yes	
2. Was the economic importance of the research question stated?	Yes	
3. Was/were the viewpoint(s) of the analysis clearly stated and justified?	Not clear	Perspective not explicitly stated although implicitly from viewpoint of NHS. Social care costs in model not populated.
4. Was a rationale reported for the choice of the alternative programmes or interventions compared?	Yes	
5. Were the alternatives being compared clearly described?	No	In the model there are two arms, current practice and practice including VibraTip. These appear to be structurally different.
6. Was the form of economic evaluation stated?	No	Not explicitly stated. This was a cost consequence model incorporating a decision tree.
7. Was the choice of form of economic evaluation justified in relation to the questions addressed?	Not clear	May have been more appropriate to use a state transition model.
8. Was/were the source(s) of effectiveness estimates used stated?	No	There is no relationship between results reported in clinical evidence submission and the economic evidence submission. Sensitivity analysis based on ungrounded assumption.
9. Were details of the design and results of the effectiveness study given (if based on a single study)?	N/A	
10. Were details of the methods of synthesis or meta-analysis of estimates given (if based on an overview of a number of effectiveness studies)?	N/A	
11. Were the primary outcome measure(s) for the economic evaluation clearly stated?	Yes	Ulcer formation and amputation.
12. Were the methods used to value health states and other benefits stated?	N/A	
13. Were the details of the subjects from whom valuations were obtained given?	N/A	
14. Were productivity changes (if included) reported separately?	N/A	
15. Was the relevance of	N/A	

productivity changes to the study question discussed?		
16. Were quantities of resources reported separately from their unit cost?	Yes	Cost of test or treatment per patient and number of patients.
17. Were the methods for the estimation of quantities and unit costs described?	Yes	But source of costs questionable.
18. Were currency and price data recorded?	Yes	Costs in pounds sterling only.
19. Were details of price adjustments for inflation or currency conversion given?	No	No costs have been inflated and no currency conversions were carried out.
20. Were details of any model used given?	No	
21. Was there a justification for the choice of model used and the key parameters on which it was based?	No	
22. Was the time horizon of cost and benefits stated?	Yes	Time horizon of three years used, but not justified.
23. Was the discount rate stated?	Not clear.	Sponsor states '3.5% Added in final figures'. Discounting should be done for every year and no indication the discount was actually performed in the final year.
24. Was the choice of rate justified?	No	But standard rate used by NICE (52).
25. Was an explanation given if cost or benefits were not discounted?	No	
26. Were the details of statistical test(s) and confidence intervals given for stochastic data?	N/A	
27. Was the approach to sensitivity analysis described?	Yes	But sensitivity data not used to quantify uncertainty.
28. Was the choice of variables for sensitivity analysis justified?	No	
29. Were the ranges over which the parameters were varied stated?	Yes	
30. Were relevant alternatives compared? (That is, were appropriate comparisons made when conducting the incremental analysis?)	Yes	
31. Was an incremental analysis reported?	Yes	
32. Were major outcomes presented in a disaggregated as well as aggregated form?	No	
33. Was the answer to the study question given?	Yes	
34. Did conclusions follow from the data reported?	Yes	But data was flawed.
35. Were conclusions accompanied by the appropriate caveats?	No	

36. Were generalisability issues addressed?	Not clear	Since no diagnostic clinical studies used in economic models
Adapted from Drummond MF, Jefferson TO (1996) Guidelines for authors and peer reviewers of economic submissions to the BMJ (59). Cited in Centre for Reviews and Dissemination (2008) Systematic reviews. CRD's guidance for undertaking reviews in healthcare. York: Centre for Reviews and Dissemination		