Public Health England

UNIVERSITY OF EXETER

Accuracy and validity of HbA1c Point of Care Testing: A review of the scientific evidence and guidelines

Version: 3

Authors: Zhivko Zhelev, Jaime Peters, Morwenna Rogers, Rob Andrews, Timothy McDonald, Christopher Hyde

Exeter Test Group, University of Exeter Medical School

Date: 19 August 2020

Contents

List of Tables

List of Figures

List of abbreviations

CAP	College of American Pathologists
CLIA	Clinical Laboratory Approval Amendments
CV	Coefficient of variation
CVD	Cardiovascular disease
DCCT	Diabetes Control and Complications Trial
DM	Diabetes mellitus
DPP	Diabetes Prevention Programme
EQA	External quality assessment
ESCAP	Expert Scientific and Clinical Advisory Group
FDA	Food and Drug Administration
HbA1c	Haemoglobin A1c
IFCC	International Federation of Clinical Chemistry and Laboratory Medicine
IQA	Internal quality assessment
IS	International System of Units
NACB	National Academy of Clinical Biochemistry
NDH	Non-diabetic hyperglycaemia
NGSP	National Glycohemoglobin Standardization Program
PHE	Public Health England
POC	Point of care
POCT	Point of care testing
QC	Quality control
SRM	Secondary reference method
SD	Standard deviation
TE	Total error
TAE	Total allowable error

Glossary

HbA1c measurement and conversion	
IFCC or IS units	HbA1c is reported in mmol/mol
NGSP or DCCT units	HbA1c is reported in %
HbA1c conversion between IFCC and NGSP units	NGSP (%) = 0.0915 IFCC (mmol/mol) + 2.15
Absolute bias	NGSP=0.0915 IFCC
Imprecision in CV% at 50 mmol/ mol	NGSP = 0.685 IFCC
Imprecision in CV% in NGSP units	IFCC = 1.46 NGSP
Analytical validity	
Analytical validity study	A type of study that aims to determine how accurately an instrument can measure the analyte of interest. This may involve the evaluation of linearity (the ability of the assay to return values that are directly proportional to the concentration of the analyte in the sample), method comparison (an agreement with a reference method which may or may not be considered perfectly accurate), precision, interferences and other aspects. Usually, analytical validity studies are carried out in a laboratory using selected samples to cover the analytically relevant range of values and may not reflect the distribution of values observed in the respective patient population.
Bias	The difference between the value obtained by the instrument under evaluation and the reference value (which may nor may not be assumed to be the true value)
Mean bias	Mean bias is the mean difference between two methods in a series of paired measurements
Precision	Closeness of agreement between a series of measurements of a homogeneous sample, e.g. from replicate determinations by the same instrument. This may include within-run, within-laboratory and between-laboratory variation.
Coefficient of variation (CV%)	Coefficient of variation (as a measure of precision) is the ratio of the standard deviation to the mean in a series of measurements, expressed as %
Total error (TE)	Total error is the deviation from the reference value due to the combined effect of bias and imprecision
Total allowable error (TAE)	Total allowable error is an arbitrary threshold for total error defined by consensus and based on the clinical significance of the magnitude of the total error
Sigma-metrics	A method for benchmarking process performance according to the risk of failing a pre-determined standard, such as the total allowable error. For instance, the IFCC analytical performance criterion for HbA1c determination in routine laboratories is 2-sigma given a TAE of 10% (5 mmol/mol at 50 mmol/mol HbA1c); this means that >95% of results should be within this limit. 4-sigma is required for reference laboratories.
Diagnostic accuracy	
Diagnostic accuracy study	A type of study that aims to determine the error rate (accuracy) of one or more index tests in classifying patients as having or not

	having a specific target condition. The results from the index test are compared to a reference standard assumed to be 100% accurate and reported as true positive, false positive, false negative and true negative.
Index test	The test under evaluation.
Reference standard	The method used in a diagnostic accuracy study to determine if patients have the target condition. This may involve one or more tests and other procedures, such as adjudication, and is assumed to be 100% accurate. Therefore, all disagreements between the index test and the reference standard are interpreted as index test errors.
Single-gate diagnostic accuracy study	Diagnostic accuracy study design in which patients with unknown disease status, selected from the relevant population through random or consecutive sampling, undergoes the index test and the reference standard. For instance, the study may include consecutive patients that would normally undergo the NHS Health Check; each patient receives both the HbA1c POCT (index test) and laboratory testing (reference standard), and the results are included in a contingency table. This study design is superior to a two-gate design as the patient disease status is unknown and the spectrum of included patients is similar to the one observed in clinical practice.
Two-gate diagnostic accuracy study	Diagnostic accuracy study design in which patients with the target condition and those without the target condition are drawn from two different populations with pre-established disease status. For instance, the cases are selected from patients with DM who attend a diabetes clinic and controls are selected from patients without prior diagnosis of DM who go to their GP with unrelated complaints. Such studies are likely to overestimate the accuracy of the index test, since not the full spectrum of patients seen in clinical practice is included.
End-to-end study	Studies that investigate the impact of alternative testing strategies (e.g. POCT verses laboratory testing) on outcomes beyond accuracy, e.g. the proportion of patients enrolled in a prevention programme, move from one risk category to another (e.g. from NDH to DM) and die of a CVD event. Study designs may range from observational studies to RCTs.
Sensitivity	The proportion of patients with a positive index test result out of all patients who have the target condition
Specificity	The proportion of patients with a negative test results out of all patients who do not have the target condition
Positive predictive value (PPV)	The proportion of patients with the target condition out of all patients who have a positive result on the index test.
Negative predictive value (NPV)	The proportion of patients without the target condition out of all patients who have a negative result on the index test.
Clinical risk agreement	A type of total accuracy measure which quantifies the proportion of patients assigned to the same risk category by both the index test and the reference standard. As all total accuracy measures, it assumes that all types of error are valued equally which, in some circumstances, may not be the case. For instance, two POCT devices may have the same clinical risk agreement, but the

	impact would be different if patients with DM (according to the reference standard) are moved to the NDH category by the first device and to the normoglyceamic by the second, even if the proportion of misclassified patients is the same.
Cost-effectiveness	
Cost-minimisation study	A type of health economic evaluation in which alternative interventions are compared in terms of cost per intervention assuming equivalent clinical effectiveness.
Cost-effectiveness study	A type of health economic evaluation which compares the relative costs and outcomes (effects) of different interventions.

Executive summary

Purpose of the review

To assess the accuracy and validity of HbA1c point of care testing (POCT) devices to inform decision making on their use in the context of the NHS Health Check programme.

Background

The NHS Health Check programme aims to improve the health and wellbeing of adults through earlier awareness, assessment, and management of the major risk factors and conditions driving premature death, disability and health inequalities in England. It is aimed at people between the ages of 40 and 74 years and is delivered outside the acute setting using both laboratory and POCT methods. The issue of the accuracy and validity of HbA1c POCT when used to screen patients at high risk of diabetes mellitus (DM) and non-diabetic hyperglyceamica (NDH) has recently been raised with the NHS Health Check Expert Scientific Clinical Advisory Panel (ESCAP). Poor accuracy of the test can lead to false reassurance in the case of false negative results or to over-diagnosis and over-treatment, if the result is false positive. Beyond certain level, both types of error could undermine the effectiveness and cost-effectiveness of the programme. To clarify this issue, the ESCAP has commissioned a review of the research evidence on the use of HbA1c POCT devices in the context of the NHS Health Check programme.

Review questions

1. What is the accuracy of HbA1c POCT devices when used in a non-laboratory setting compared to a venous blood sample analysed in a laboratory setting?
2. What delivery conditions are required to ensure that POCT devices used in non-laboratory settings are as accurate as possible?
3. What are the benefits of using POCT for HbA1c over lab-based testing?
4. What are the current national and international guidelines on the use of POCT for classifying/ diagnosing NDH and type 2 DM?

Methods

We searched Embase, MEDLINE and the Cochrane Library (CDSR and CENTRAL) for studies published in English between 1st January 2015 and 10th January 2020. The period before that was covered by including a systematic review of accuracy studies and a literature review of studies evaluating the advantages and disadvantages of HbA1c POCT devices. We also conducted hand-searches and contacted authors and manufacturers to identify additional titles and unpublished data; and searched relevant websites for guidelines.

Studies were included if they reported results relevant to any of the above questions. Outcomes included any of the following: analytical validity metrics including mean bias and precision, and sigma metrics which combine both; diagnostic accuracy outcomes, such as sensitivity and specificity; any outcomes on the advantages and disadvantages of HbA1c POCT when used to screen high risk adults for DM and NDH; relevant international guidelines and guidelines from the following countries: the UK, Ireland, USA, Canada, Australia, New Zealand, Norway, Sweden and the Netherlands. The last three countries were included at a later stage, after we had found publications reporting on the assessment of HbA1c POCT within External Quality Assessment (EQA) programmes.

The selection of relevant studies was done independently by two reviewers. One reviewer extracted the data which was then verified by a second reviewer. The methodological quality of the included studies was assessed using QUADAS-2 for accuracy studies, AMSTAR for systematic reviews, the

Cochrane Collaboration checklist for RCTs, AGREE II for guidelines and a checklist developed by our team and based on published tools for health economic models.

Due to differences in methodology, variable quality and highly heterogeneous results across studies, we decided not to conduct meta-analyses. Instead, we summarised the results in tables and presented them narratively. First, we report on the analytical validity of HbA1c POCT devices presenting the results from the systematic review conducted by Hirst 2017, the results from the newly identified studies and the analyses of EQA data. We considered the latter to provide better quality evidence (since they use standardised procedures and higher-level reference methods); to allow direct comparison between different devices; and to reflect the performance of the devices in clinical practice. Then we present the evidence pertaining to the diagnostic accuracy of these devices, which tells us what proportion of patients will be misclassified as false positives and false negatives. Finally, we present studies investigating the advantages and disadvantages of using HbA1c POCT to screen for NDH and DM and related national and international guidelines.

Results

The final number of studies included in the review was 64 (counting the two reviews we used as a starting point): 41 journal articles, one online publication, 13 conference abstracts and 9 guidelines. Of the 45 accuracy studies, 9 reported on analyses of EQA data and 8 reported on diagnostic accuracy. Twelve studies investigated the benefits and disadvantages of using HbA1c POCT devices (one was an RCT and 5 also reported on accuracy and were included above). Nine manufacturers replied to our data request; 3 of them provided additional data. We also included the 2019-2020 data for POCT devices from the College of American Pathologies (CAP) Survey.

DCA 2000/Vantage was evaluated in 39 studies included in the systematic review conducted by Hirst 2017 and in 25 new studies; the respective numbers for the other devices were: 12 and 21 for Afinion (including 3 papers reported on the new Dx version of the assay); 17 and 3 for A1c Now+; and smaller numbers for Cobas b101, B-analyst, Quo-Test, Quo-Lab, NycoCard, PixoTest, in2it, A1C EZ 2.0, DS5, RC20, I-Chroma, A1Care, and HbA1c 501. Analyses of EQA data were available only for DCA 2000/Vantage, Afinion and Quo-Test. Eighteen of the newly found accuracy studies evaluated two or more POCT HbA1c tests. Hirst 2017 reported pooled sensitivity and specificity estimates for DCA (n=24), Afinion (n=8), A1cNow (n=7), QuoTest (n=3) and NycoCard (n=4). The new studies reported on the diagnostic accuracy of DCA (n=3), Afinion (n=2), A1cNow+ (n=2), in2it (n=1) and A1c EZ 2.0 (n=1).

Although most analytical validity studies followed established protocols for evaluation of method comparison (e.g. CLSI EP-9) and precision (e.g. CLSI-EP5) they reported extremely heterogeneous results, most likely due to differences in the reference methods used. Most of the evaluations that used a higher level of reference method (e.g. the International Federation of Clinical Chemistry and Laboratory medicine (IFCC) secondary reference method (SRM)) reported that DCA and Afinion has the potential to meet the IFCC 2-sigma (or equivalent criteria) for analytical performance. This means that >95% of results will fall within the 10% total allowable error (TAE) equivalent to 5mmol/mol at 50 mmol/mol HbA1c. The same performance is expected from routine laboratory instruments. Data from EQA programmes suggest that these devices are able to meet the above (or similar) criteria when used in primary care and has similar analytical performance to routine laboratory instruments. Most likely, some of the other devices are also able to achieve this standard, but the evidence is less conclusive and more good quality data, especially from EQA programmes, is needed. However, optimal performance in a clinical setting could only be achieved if certain conditions, such as training, maintenance, internal and external quality control and collaboration with local laboratories are in place.

No good quality diagnostic accuracy studies were found. The pooled sensitivities and specificities reported in Hirst 2017 are based on analytical validity studies and do not provide unbiased estimates of the diagnostic accuracy of HbA1c POCT when used in primary care or community. The Norfolk Diabetes Prevention Study (NDPS) shows that a large proportion of patients will switch category (normoglyceamic, NDP and DM) when tested at baseline and 40 days later, even in perfect testing conditions: using a single high-performance laboratory instrument operated by a small number of highly trained staff! In this study 24% of initially normoglyceamic patients progressed to NDH; 21% of those in the NDH category regressed to normoglyceamic; 7% of those in the NDH category progressed to DM and 21% of the DM category regressed to NDH. Also, only 61% of patients had both HbA1c and fasting plasma glucose elevated at baseline! The authors recommended using both tests when screening for DM and NDH as the error rate of a single test was too high.

A number of expected advantages of using HbA1c POCT devices were suggested in the literature including immediately available results, ease of use, shorter analysis time, patient satisfaction and engagement, small footprint and portability, and being less demanding in terms of training and maintenance. However, most of the studies evaluating the advantages and disadvantages of HbA1c POCT were at high risk of bias and of limited applicability to the NHS Health Check programme. We found one RCT conducted in primary care in New Zealand which compared expected and realised benefits of using HbA1c POCT. No difference was found in terms of effectiveness (completed risk assessments and time) between POCT and laboratory testing; and a series of qualitative interviews highlighted the importance of contextual factors, proper integration of POCT in the practice workflow and dedicated resources. Also, we found no cost-effectiveness studies that compared HbA1c POCT and laboratory testing in the context of the NHS Health Check or a similar programme.

Four of the 9 guidelines included in the review (ADA 2009, IDF 2017, DAGDC 2009, NZMoH 2011) stated that HbA1c should only be measured in laboratories. Although the NHS Health Check Best Practice Guidance reported recommendations on using POCT from MHRA, these are not specifically for HbA1c. The NHS DPP Guidance recommended the use of HbA1c POCT only if there is evidence of cost-effective implementation and the MHRA quality framework recommendations are in place prior to the introduction of the POCT deviece. The NICE guidance (2017) stated that HbA1c testing, including POCT, should only be done by trained staff and follow expert consensus reports on use and conform to national quality specifications. The Canadian guidance (DCCPGEC 2018) only recommends HbA1c POCT when laboratory testing is not available for screening remote Indigenous populations. In these circumstances the testing must be part of a quality control programme where expertise and follow-up is available. The recent American guidance (ADA 2020) did recommend diagnostic use of HbA1c POCT but only “in settings licensed to perform moderate-to-high complexity tests.”

Conclusions

The identified research evidence suggests that DCA and Afinion has the potential to meet and, in some cases, exceed the criterion of 2-sigma required for HbA1c determination in routine laboratories, provided adequate training, maintenance, internal and external control programmes are in place. This means that >95% of the HbA1c results will be within the TAE of 10% (5 mmol/mol at HbA1c level of 50 mmol/mol). Other devices also seem to have the potential to meet this target, but further evidence is needed.

We found no good quality evidence on the diagnostic accuracy of HbA1c POCT when used to screen for DM and NDH and are unable to report reliable estimates of sensitivity and specificity. However, in the NDPS study a significant proportion of patients were re-classified when two consecutive

measurements 40 days apart were made, despite the fact that HbA1c was measured in nearly perfect conditions: using a single high-performance laboratory instrument by a small number of highly qualified staff. They also found considerable disagreement between HbA1c and fasting plasma glucose at baseline and suggested that both tests should be used to minimise the proportion of patients misdiagnosed when using a single screening test.

The research evidence pertaining to the advantages and disadvantages of using HbA1c POCT when screening for DM and NDH devices is also limited, at high risk of bias and with limited applicability to the NHS Health Check programme. An RCT conducted in New Zealand showed that careful planning, additional resources and integration of the POCT in the practice workflow is required if the expected advantages of using POCT are to be materialised. None of the included guidelines make specific recommendations for HbA1c POCT when used to screen for DM and NDH, but MHRA provide some guidance for the use of POCT devices that is applicable to the HbA1c POCT and the DPP Guidance recommend their use only if there is evidence that such testing strategy would be cost-effective.

Future research studies should compare different screening strategies (e.g. HbA1c vs fasting plasma glucose using POCT vs laboratory testing) and aim to go beyond diagnostic accuracy, considering their clinical and cost-effectiveness.

Background

The NHS Health Check programme aims to improve the health and wellbeing of adults through earlier awareness, assessment, and management of the major risk factors and conditions driving premature death, disability and health inequalities in England. Over 15 million people between the ages of 40 and 74 years are eligible for an NHS Health Check once every five years [1].

During the risk assessment, standardised tests are used to measure the seven risk factors and calculate a person's risk of having a heart attack or stroke in the next 10 years. This includes undertaking an HbA1c or fasting blood glucose test with people identified as being at high risk of type 2 diabetes. The outcome of the assessment is then used to raise awareness of the risk factors, as well as inform a discussion on, and agreement of, the behavioural and medical approaches best suited to reducing the individual's risk of cardiovascular disease (CVD).

A key part of the NHS Health Check programme's governance structure is the Expert Scientific and Clinical Advisory Panel (ESCAP). ESCAP is an expert forum which provides advice on the content of and potential changes to the programme. It also acts in an advisory capacity to support successful roll-out, maintenance, evaluation and continued improvement based on emerging and best evidence.

The issue of the accuracy and validity of point-of-care testing (POCT) for HbA1c has recently been raised with the ESCAP. Poor accuracy may lead to false reassurance in the case of a false negative result or to over-diagnosis and over-treatment, in the case of a false positive result. A high rate of false positive and false negative results could undermining the effectiveness and cost-effectiveness of the programme. Given the expressed concerns, ESCAP has commissioned the current work to establish the accuracy and validity of HbA1c POCT devices when used in settings similar to those in which the NHS Health Check is delivered. The work will inform the Public Health England's (PHE) position on the use of HbA1c POCT devices in delivering the NHS Health Check programme.

Aim of the review

The review aimed to answer the following research questions:

1. What is the accuracy of HbA1c POCT devices when used in a non-laboratory setting compared to a venous blood sample analysed in a laboratory setting?
2. What delivery conditions are required to ensure that POCT devices used in non-laboratory settings are as accurate as possible?
3. What are the benefits of using POCT for HbA1c over laboratory-based testing?
4. What are the current national and international guidelines on the use of POCT for classifying/ diagnosing non-diabetic hyperglycaemia (NDH) and type 2 Diabetes Mellitus (DM)?

Methods

Prior to undertaking the review, we developed a review protocol which was discussed with and approved by the commissioning body. The protocol details the scope and methods of the review in keeping with the UK National Screening Committee's evidence review process.

Inclusion and exclusion criteria

We included systematic reviews and primary studies evaluating the analytical validity, diagnostic accuracy and the benefits and limitations of POCT devices used for determination of HbA1c. We

adopted the definition of HbA1c POCT device from the systematic review conducted by Hirst 2017 [2]: *"…any instrument designed to provide a rapid quantitative measurement of HbA1c using capillary blood at the point of care*" (p. 2). All HbA1c POCT devices were considered for inclusion provided they were commercially available before 30th March 2020. If necessary, we tried to contact authors to clarify the availability of the devices and excluded the study if no further information was available.

The device had to be used by a clinical personnel at the point of care in unselected adults who had not been previously diagnosed with DM or CVD. Given the limited number of diagnostic accuracy and analytical validity studies meeting the above criteria, we also included studies: 1) conducted in a laboratory setting; 2) in which venous blood was used and 3) studies in which blood from patients with DM or mixed groups (DM and non-DM patients) was used. In the absence of directly applicable research evidence, the results from such studies could provide an approximation of the performance of the device in a non-laboratory setting. We excluded studies conducted in children and in selected patients, such as patients with hemoglobinopathies, CVD, tuberculosis and other pre-existing conditions, as the performance of the evaluated devices might be different in such groups.

Accuracy studies were defined as studies in which the results from a POCT device (index test) were compared to those from a laboratory method (reference standard) and studies where the POCT device was evaluated as part of an External Quality Assessment (EQA) programme. Relevant outcomes included:

- Analytical accuracy (mean bias and precision, and related measures, such as total error and sigma metrics, explained below)
- Diagnostic accuracy (sensitivity and specificity, and related measures, such as likelihood ratios and predictive values)
- Factors that may affect the performance of POCT devices (e.g. background and training of the person performing the test)

We also included all studies (regardless of their design) which reported on the benefits and limitations of HbA1c POCT when used in settings similar to those in which the NHS Health Check programme is delivered. The outcomes included, but were not limited to, test-operators and users' experience and health-economic outcomes. Finally, we searched for and included guidelines on the use of POCT HbA1c for diagnosis of NDH and type 2 DM in a non-laboratory setting. Current international guidelines and national guidelines from any of the following countries were eligible for inclusion: the UK, Ireland, USA, Canada, Australia, New Zealand, Norway, Sweden and the Netherlands. The last three countries were added to the list at a later stage, after we had found relevant publications reporting on the analysis of EQA data.

Since this was a rapid review, we included only studies published in English since 1st January 2015. This date was chosen because in our scoping searches we identified two relevant reviews of reasonable quality that reported on the outcomes of interest and covered the period before 2015 [2, 3]. Conference abstracts were included if they reported the study in sufficient detail.

Search methods

An information specialist (MR) with extensive experience in systematic review searches adapted the search strategy used in Hirst 2017 [2] by adding the names of HbA1c POCT devices available from NHS Supply Chain (provided by the commissioning body and detailed in Error: Reference source not found). The strategy combines free text and subject headings for "glycated haemoglobin" and "point-of-care systems" (Appendix 1). We searched Embase, MEDLINE and the Cochrane Library

(CDSR and CENTRAL), limiting the searches to papers published in English since 1st January 2015. We also searched the TRIP Database, NICE Evidence, Google and Diabetes UK for relevant guidelines.

In addition, we hand-searched the reference lists of all included studies and other relevant publications and consulted experts for additional titles. We emailed the manufacturers of all devices evaluated in the primary studies and the devices listed in Table 1, and asked them to provide references to relevant publications and EQA data, if available.

Selection of studies

Two reviewers independently screened the titles and abstracts of all publications identified in the searches and the full texts of those selected in the first round. All disagreements were resolved through discussion and arbitration.

Data extraction

A data extraction form was developed and piloted using some of the studies included in Hirst 2017 [2]. One reviewer (ZZ) extracted all data which was then checked by a second reviewer (JP) and all discrepancies were resolved through discussion and arbitration.

Methodological quality assessment of the included studies

We used a tailored version of the QUADAS-2 checklist to evaluate the methodological quality of all diagnostic accuracy studies. QUADAS-2 covers four methodological quality domains: Patient selection, Index test, Reference standard and Flow and timing. The first three are assessed with respect of bias and applicability, while the fourth one, Flow and timing, is assessed only with respect of bias [4]. We tailored the tool (as recommended by its authors) by making the following changes.

In the Patient selection domain, we added a question checking whether the data was collected prospectively for the purpose of the study, since retrospective data collection increases the risk of error and bias. As both the index test and the reference standard are objective quantitative tests, we removed the questions related to blinding and, in the Index test domain, considered only applicability. No applicability concerns were noted if the POCT device was operated by non-laboratory staff in a primary care or community setting and capillary blood was used. In the Reference standard domain, the risk of bias was rated as 'low' if the laboratory instrument used as reference standard was explicitly defined as a secondary reference method (SRM) or the evaluation was done as part of an EQA survey. We also added in a question concerning funding by the manufacturer of the device, as this may increase the probability of favourable results.

QUADAS-2 is not entirely appropriate for an assessment of analytical validity studies, which aim to evaluate bias and precision at different levels of the analyte, rather than the accuracy of the test in classifying patients into different diagnostic categories. Therefore, for the analytical validity studies, we adapted the checklist by removing questions that seemed irrelevant, such as those related to patient selection and blinding, and adding questions concerning prospective/retrospective study design, the range of HbA1c values covered in the evaluation and whether the study received funding from the manufacturer.

In addition, we used AMSTAR to evaluate the methodological quality of systematic reviews [5]; the Cochrane Collaboration checklist for cluster RCTs [6]; AGREE II for guidelines [7] and a checklist developed by our team and based on Peñaloza 2015 [8] for health economic models.

Data synthesis

We decided not to conduct meta-analyses or to update the ones reported by Hirst 2017 [2]. The reasons were as follows:

- We found only a small number of diagnostic accuracy studies at high risk of bias and of limited applicability; most of them did not report the accuracy data in sufficient detail to allow pooling of results. The sensitivity and specificity estimates reported by Hirst 2017 [2] were based mainly on analytical validity studies which did not aim to evaluate the accuracy of the device at patient level; such studies may not reflect the actual performance of the device in clinical practice; the diagnostic accuracy data were extracted from correlation and Bland-Altman plots, which increases the risk of error. Therefore, updating the reported estimates would not have produced better quality evidence.
- The analytical validity studies included in Hirst 2017 [2] were of variable methodological quality and reported highly heterogeneous results leading to high level of statistical uncertainty in the pooled estimates. The new studies were equally heterogeneous and of variable quality; including them in the meta-analyses would not have improved the validity or precision of the pooled estimates of bias.
- Most of the studies investigating the benefits and limitations of HbA1c POCT devices were observational studies at high risk of bias and of limited generalisability and relevance to the NHS Health Check programme; they measured a range of different outcomes, which precluded pooling of results.

Instead of quantitative synthesis, we summarised the results in tables and presented them narratively, separately for each review question, and focusing on the evidence most relevant to the NHS Health Check programme. We reported the results from the accuracy studies separately for each device, first presenting the results from the meta-analysis reported by Hirst 2017 [2], then summarising the results from new studies that used a laboratory method as reference standard and, finally, summarising the results from studies analysing EQA data. The latter were deemed more likely to provide unbiased and generalizable evidence as they used standardised assessment procedures; included multiple devices and sites; included data for routine laboratory methods, thus allowing direct comparison; and captured the actual performance of the device as used in clinical practice.

Results

Results from the searches and selection of studies

Database searches were conducted on 10th January and guideline searches on 24th January 2020. Additional searches for guidelines published in Sweden, Norway and the Netherlands were carried out on 21st April 2020. Appendix 2 gives the number of hits per database while Figure 1 shows the initial number of titles from each source and the number of studies at each stage of the selection process.

The database searches resulted in 797 hits; additional 14 studies were suggested by PHE; 1 more study was found from searching reference lists and 26 guidelines were identified as potentially relevant. After removing all duplicates, 603 titles/abstracts were screened and 163 were retained for further assessment. Of those, 104 were journal articles assessed for eligibility at full text; 33 were conference abstracts for which additional information was sought; and 26 guidelines were assessed for relevance to the HbA1c POCT focus of the review.

The information provided by the eight companies that replied to our data request is detailed in Table 2 and could be summarised as follows:

- The in2it device has been discontinued and currently BioRad do not offer any HbA1c POCT devices
- RC20 (Sekisui Medical) is not available in Europe; RC-W, a model similar to RC20, is scheduled to become available on the European market after October 2020; no additional data or publications related to RC20 were provided
- No additional data or publications were provided for the NycoCard (which, according to the manufacturer, is intended only for monitoring of diabetes), Pixotest A1c, A1Care and the Eurolyser CUBE
- Only EQA certification data were provided for Allegro and Finecare
- A list of publications was provided for Afinion (Abbott) of which one new study, not captured by our searches, was identified. The study was published both in English and Spanish, and compared Afinion AS100, DCA Vantage and the In2it POCT systems [9]

In addition, we downloaded the results for 2019 and 2020 from the College of American Pathologists (CAP) Survey, which is freely available to access online: http://www.ngsp.org/CAPdata.asp. The results for DCA Vantage, Afinion AS100 and Afinion 2 (the only POCT devices enrolled in the programme) were used in the analysis of analytical performance of those devices.

Characteristics of the included studies

The final number of studies included in the review was 64 of which 41 were journal articles, 1 was an online publication, 13 were conference abstracts and 9 were guidelines. One of the included studies was a systematic review and meta-analysis of accuracy studies [2]; one was a literature review looking more broadly at the use of HbA1c POCT in non-laboratory setting [3]; 45 were accuracy studies (including 9 papers reporting on analyses of EQA data), 9 were studies investigating the benefits and disadvantages of HbA1c POCT devices and 1 was an observational study reporting on the POCT in the contest of the UK NHS Diabetes Prevention Programme (DPP). The studies were conducted in the following countries (not counting the reviews): 18 studies were conducted in the USA, 5 in the UK, 4 in Spain, 2 each in Canada, China, New Zealand, Norway, Taiwan and the Netherlands, and 1 each in Australia, Belgium, Ecuador, Iran, Japan, Mexico, South Africa, Sweden, Switzerland, the United Arab Emirates and Peru. In addition, one study was conducted in the European Reference Laboratory for Glycohemoglobin by authors from the Netherlands and the UK [10]; one paper reported on the EurA1c trial conducted in 17 European countries [11]; and two papers reported on the analysis of data from 3 related studies conducted in the USA and, possibly, Norway [12, 13].

Eighteen of the accuracy studies evaluated two or more POCT HbA1c devices. DCA 2000/Vantage was evaluated in 25 studies; Afinion in 21 studies (3 of the papers reported on the new Dx version of the assay); Cobas b101 was evaluated in 4 studies; A1c Now+ and Quo-Test in 3; B-analyst, NycoCard, PixoTest and in2it in 2; and A1C EZ 2.0, DS5, RC20, I-Chroma, Quo-Lab, The A1Care, and the HbA1c 501 were evaluated in one study each. Only DCA 2000/Vantage, Afinion and Quo-Test were included in an analysis of EQA data published in a peer-reviewed journal (11, 7 and 1 paper, respectively). In addition, we received EQA certification data for Allegro and Finacare provided by the manufacturers (Table 2) and downloaded EQA data for DCA Vantage, Afinion AS100 and Afinion 2 from the CAP Survey website.

The only relevant systematic review we found was conducted by Hirst and colleagues who searched MEDLINE, Embase and the Web of Science up until June 2015 [2]. Studies were included in the review if they evaluated the accuracy of a POCT HbA1c device against a laboratory-based method

and reported the mean difference between the POCT and laboratory HbA1c. Sixty one studies were included in the meta-analyses; the following devices were evaluated: DCA 2000/Vantage (n = 39), A1cNow (n = 17), Afinion (n = 12), Quo-Test (n = 7), Nycocard (n = 6), B-analyst (n = 5), Cobas b101 (n = 5), Innovastar (n = 5), Quo-Lab (n = 3), HemoCue (n = 2), Clover (n = 2), SDA1c Care (n = 1) and A1cgear (n = 1).

We did not find any relevant systematic reviews investigating the benefits and limitations of the HbA1c POCT devices. Instead, we included a literature review published in 2016 which, in addition to test accuracy studies, reviewed studies looking at the benefits, limitations, acceptability and cost-effectiveness of such devices [3]. The review was conducted by the Diagnostic Evidence Co-operative Oxford and aimed to answer the following question: *"In a primary care setting, what is the utility of HbA1c point-of-care testing (POCT) devices in the detection/diagnosis* of *diabetes mellitus (DM), compared to standard laboratory methods for HbA1c analysis?"* (p.1)

According to the report "*a strategic literature search was performed on MEDLINE, Embase, Scopus, CINAHL, Cochrane library, TRIP and Web of Science*" (p. 6). The dates and other parameters of the search as well as the methods for selecting studies, extracting data and assessing the methodological quality of included studies were not reported. Some of the authors of this report were also authors on the Hirst 2017 paper [2]. The report listed 24 HbA1c POCT devices and their key characteristics as claimed by the manufacturers. Research evidence on the accuracy and/or benefits and limitations of the following devices was included: A1cNow+, Cobas b101, A1c Gear, DCA 2000(+)/Vantage, InnovaStar, B-Analyst, SD A1c Care, Afinion AS100, NycoCard, Quo-Test, Quo-Lab and Clover. Most of the accuracy studies included in Hirst 2017 [2] and in the literature review [3] were the same, so we extracted only results pertaining to the benefits and limitations of the evaluated HbA1c POCT devices.

Finally, of the 9 included guidelines, 2 were international guidelines, 3 were from the UK and 1 each from Australia, Canada, New Zealand and the USA. They were published between 2009 and 2020.

Methodological quality of the included studies

The results from the methodological quality assessment of the included diagnostic accuracy and analytical validity studies are summarised in Table 3 and Table 4Table 4, respectively. All diagnostic accuracy studies were considered at high risk of bias in at least one domain. Only one study reported consecutive sampling and 4 out the 8 studies reported prospective design. None of the studies reported using an SRM as reference standard and one study reported receiving funding from the manufacturer. In terms of applicability, only 2 studies included participants similar to those in the NHS Health Check programme (one of which was conducted in the UK as part of a CVD screening programme) and only for one study we could ascertain that the POCT device was used in a community or primary care setting by non-laboratory staff (Table 3).

Prospective data collection was reported or could be assumed from the paper only for 13 of the 35 analytical validity studies. Twenty nine of them reported an assessment of the POCT device over a wide analytical range defined as approximately 31 to 97 mmol/mol (5% to 11%) HbA1c. Only 9 studies reported that the reference standard was an SRM; in 11 it was clear that the time between the index test and the reference standard was <24 hours and 9 studies reported funding by the manufacturer. In terms of applicability of the index test, only 5 studies reported that the POCT device was used in a community or primary care setting by non-laboratory staff (Table 4).

The systematic review conducted by Hirst 2017 [2] met most of the relevant criteria of the AMSTAR 2 checklist and was judged to be of good methodological quality. AMSTAR 2 is intended mainly for systematic reviews of RCTs and some of the quality items are not relevant for systematic reviews of

diagnostic accuracy studies. The results from the meta-analyses were considered to be broadly applicable to our review question despite the inclusion of studies that would not have met our inclusion criteria, such as studies in children. The authors conducted a series of sub-group analyses to investigate the impact of such study characteristics as well as the methodological quality of the included studies. Unfortunately, the authors did not report the QUADAS checklist in sufficient detail, which makes it difficult to interpret the results from the sensitivity analyses in which studies at high risk of bias were excluded.

We identified one RCT, the EPOCH trial conducted in New Zealand [14], which was considered to be at low risk of bias when assessed against the Cochrane Collaboration checklist for cluster RCTs [6].

The included guidelines met most of the AGREE II criteria except for the following three:

- The views and preferences of the target population (patients, public, etc.) have been sought.
- A procedure for updating the guideline is provided
- The guideline presents monitoring and/or auditing criteria

Exceptions were the American Diabetes Association guideline and the Recommendations for the use of POCT in the NHS DPP which, as far as we can tell from the publications, did not meet many of the criteria (see Table 5).

Accuracy of POCT HbA1c devices

Below we summarise the results from the included accuracy studies, separately for each device. We first report the results from the meta-analysis in Hirst 2017 [2], then the results from new studies comparing the device to a laboratory instrument and, finally, we present the results from studies analysing EQA data. The results are presented as reported, either in National Glycohemoglobin Standardization Program (NGSP) units as %HbA1c or in the International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) International System (IS) units as mmol/mol.

DCA 2000 and DCA Vantage

Hirst 2017

Thirty nine studies were included in the meta-analysis of DCA devices. The mean bias ranged from -0.96% (Standard Deviation (SD) 0.66) to 0.28% (SD 0.63); pooling the studies produced mean bias of -0.25% (95%CI -0.33 to -0.18) with a wide prediction interval from -0.76% to 0.25%. This means that the results from 95% of any future studies with similar characteristics are expected to fall within these limits.

Excluding the studies where the review authors had to make estimates from graphs due to missing data led to slight improvement in pooled mean bias, -0.17% (95%CI -0.26 to -0.08), while excluding studies at high risk of methodological bias led to even greater negative mean bias, -0.36% (95%CI -0.48 to -0.24). On the other hand, including only studies which used NGSP or IFCC reference laboratories (n=6) led to a pooled mean bias of 0% (95%CI -0.16 to 0.15). No prediction regions were reported for the sensitivity analyses estimates, but the level of statistical uncertainty is likely to be similar to that in the main analysis.

The SD of the mean bias also varied widely, ranging from 0.20% to 0.66% across studies. The pooled SD of the mean bias was 0.38 (95%CI 0.34 to 0.42), also with a very wide prediction interval ranging from 0.16% to 0.60%.

Additional results from the meta-regression and subgroup analysis indicated that:

- The year of publication had no significant impact on the mean bias estimate, but there was a trend towards less negative bias over time (coefficient 0.014% HbA1c per year, p = 0.081)
- SD of the mean bias significantly decreased over time (–0.01% HbA1c per year, 95% CI –0.015 to –0.006, p < 0.0001) suggesting that the variability in measurements contributing to the mean bias is decreasing
- Studies published between in 2014 and 2015 (the last two years of the review's search period) also had a non-significant mean bias of –0.08% (95% CI –0.21 to 0.06).

The authors found no difference in the mean bias between:

- A POCT and laboratory setting
- Laboratory (n=14) or clinical (n=13) operators
- Venous (n=17) or capillary (n=11) blood samples
- Children (n=6) and adults (n=30)

Seventeen studies reported imprecision based on replicate analyses of a single sample at low (<6.0%, 42 mmol/mol), medium (6.0-8.0%, 42-64 mmol/mol) or high (>8.0%, 64 mmol/mol) HbA1c (calculated from raw values and presented in % units). The coefficient of variation (CV) at each level was 2.5%, 2.0% and 2.8%, respectively, with a total CV of 2.5%. Median between-laboratory CV from the 2016 survey of the College of American Pathologists (CAP) proficiency testing programme was 2.9% which was within the CAP threshold of <3% and, for the comparator laboratory instruments included in the same survey, ranged from 1.6% to 2.6% (table 4 in the paper).

Pooling study-level sensitivities and specificities (n= 24) estimated from correlation and Bland-Altman plots at 6.5% (48 mmol/mol) produced summary estimates of 95.4% (95%CI 91.5-97.5) and 99.1T (95%CI 96.3-99.8), respectively, which remained robust when the analysis was repeated including only studies published between 2006 and 2015 [2]. No subgroup analysis was conducted to investigate the analytical performance or diagnostic accuracy of the assay at the cutoff of 6.0% (42 mmol/mol) HbA1c used to differentiate between NDH and normoglycaemia.

New studies comparing DCA 2000/Vantage to a laboratory instrument

Sixteen new studies evaluated DCA 2000/Vantage using a laboratory method as reference standard (Table 7 Diagnostic accuracy studies). DCA Vantage was used in all but one study [15], which did not specify the model. Fourteen studies reported on mean bias and 4 on imprecision.

Mean bias ranged from -0.28% [9] to 0.32% [16] (excluding Diaz-Garzon 2017 [17] which reported mean bias of -0.76% and -1.69% at 5.5% and 11.1% HbA1c levels, respectively). The range was consistent with the 95% prediction interval reported by Hirst 2017 (from -0.76% to 0.25%).

Four of the studies reported CV in NGSP units: 2.8% and 2.3% at low level and 2.5% and 2.7% at high level HbA1c [17, 18] and total CV of 1.7% [19] and 3.7% [9]. The fifth study reported CV in IFCC unites: 0.88% and 1.79% at low and high levels of HbA1c, respectively [20].

One retrospective study reported 88.6% sensitivity and 96.3% specificity (without CIs) for the diagnosis of DM at the 6.5% (48 mmol/mol) cutoff [15]; and one study reported sensitivity of 96.0% (95%CI 95.0% to 98.0%) and specificity of 89.0% (95%CI 86.0 to 92.0) at the 7.0% (53 mmol/mol) HbA1c for glycemic control in patients with type 2 DM [21]. One more study available only as conference abstract reported clinical risk agreement based on reclassification of patients diagnosed as normoglyceamic, NDH and DM. DCA Vantage achieved 100.0% agreement outperforming the A1CNow+ (94.0%) and demonstrating similar performance as the laboratory methods (98.0%,

p=0.17) (Table 7). All three studies were at high risk of bias and had limited applicability to our review questions; hence, their results should be treated with caution (Table 3).

The authors of two studies concluded that DCA Vantage was as accurate as laboratory methods [19, 20] and in one paper the assay was recommended for screening after adjusting for bias [22]. The conclusions from the rest of the studies were that clinicians should be cautious when using POCT and/or should only use them for monitoring.

Studies analysing data from an EQA programme

Eleven studies reported on the performance of DCA 2000/Vantage in an EQA programme. Data from the following countries were analysed: Australia, Canada, the Netherlands, Norway, Sweden and multiple countries participating in the EurA1c trial. The period covered in the analyses ranged from a single survey [23] to 15 years [24] and included from one [22, 25] to >4000 [26] POCT sites. The setting in which POCT was used varied, but was predominantly primary care or community (Norway and Sweden) and indigenous populations living in remote areas with limited access to central laboratory services (Australia and Canada). In addition, we included data from the CAP survey in 2020 (first five samples) and 2019 (all 3 x 5 samples). The EQA pass criteria varied across programmes, but were the same or similar to those for routine laboratory methods (Table 9).

Three European studies, the EurA1c trial 2018 [11], Lenters-Westra 2017 [23] and Delatour 2019 [27], reported that DTA Vantage attained the international guidance target of >2 sigma at total allowable error (TAE) of 10.0%. In the EurA1c trial, 11 out of 12 POCT laboratories achieved >2 sigma at 10.0% TAE and 5 out of 12 achieved >4 sigma at 10.0% TAE and >2 sigma at the more stringent criterion of 6.0% TAE. In the Equalis EQA programme (Sweden) the performance of the assay showed considerable variation over time, with the percentage of POCT sites that met the criterion of TAE of 0.33% (3.6 mmol/mol) at 6.7% HbA1c (50 mmol/mol) fluctuating between 80.0% and 93.0% (but consistently below the programme's target of 95.0%) [26].

Two studies reported data from the Noklus EQA proficiency testing programme in Norway.
Orvim Solvik 2018 showed that DCA devices had similar performance when used in community pharmacies (following training and technical support) and GP practices. Data from three EQA surveys conducted between October 2016 to April 2017 and including seven pharmacies and more than 400 GP practices were analysed. Between 56.0% and 100.0% of the pharmacies achieved "very good" performance for trueness (mean value of duplicate measurements ±2.0% of the target interval) and between 71.0% and 100.0% for precision (≤0.2% difference between duplicate measurements); no "poor" results were reported (trueness: ≥5.4% difference; precision: ≥0.5%). Of the GP practices, between 75.0% and 87.0% achieved "very good" performance for trueness and 84.0% to 94.0% for precision; between 0.2% and 1.2% and between 0.5% and 2.0% had "poor" results for trueness and precision, respectively [28].

Stavellin 2019 analysed data from 7 Noklus EQA surveys conducted between January 2017 and October 2018. Although the focus of the study was the performance of Afinion in primary care, results for DCA and Quo-Test in the same setting were also reported. The pass criteria in the study were as follows: mean bias <0.3% HbA1c (equivalent to the CAP Survey criteria); within-laboratory CV <2.0% and between-laboratory CV <3.5% (as recommended by the National Academy of Clinical Biochemistry for diagnosis of DM). All (100.0%) of the 459 primary care DCA users met the above criteria for each of the 14 samples. The authors concluded that the performance of DCA was similar to that of laboratory instruments and was suitable for diagnosing DM [29].
Shephard 2017 analysed 15 years of QAAMS data from Australia showing that the performance of DCA 2000/Vantage meet the programme criteria and matches the quality achieved by the Australian laboratories. The number of devices included in the programme increased from 45 in 1999 to 200 in

2016. On average, 89.5% (SD 5.5; range 77.0–96.0%) of the results from 2002 to 2016 (using mainly DCA 2000) and 94.0% (SD 1.3; range 92.0–96.0%) from 2009 to 2016 (when DCA Vantage was introduced) met the programme's criteria. Median imprecision across device operators averaged 2.8% (SD 0.5; range 2.2 to 3.9%) from 2002 to 2016 and 2.4% (SD 0.22; range 2.2 to 2.9%) from 2009 to 2016. At the 6.8% HbA1c (2012 – 2016) the CV ranged from 2.49% to 2.85% [24].

Studies evaluating the performance of a single POCT site, usually for a short period of time, reported mixed results [22, 24, 25, 30, 31]. One of these studies, Paknikar 2016, reported on the analysis of 3-year data from the CAP proficiency testing programme in the USA. The results showed considerable variation in bias not only for the POCT devices, but also for the laboratory methods. For DCA the bias around the diagnostic threshold ranged from -0.25% to 0.15%. The authors emphasised the limitations of these devices and that the long-term variability in the performance of these methods should be considered when making clinical decisions [30].

The results from the CAP Survey data for 2019 and the first 5 samples in 2020 were based on a median of 349 laboratories and covered the HbA1c range approximately from 5.0% (31 mmol/mol) to 10.0% (86 mmol/mol). Across the four datasets included here, only one of the samples, at 5.6% (38 mmol/mol) HbA1c, exceeded the current CAP acceptance limit of ±6.0% of the target value and showed a positive bias of 0.32%. The between-laboratory CV was ≤3.5% for all samples and ≤3.0% for most of them. DCA Vantage consistently met the quality criteria of CAP and outperformed some of the laboratory methods enrolled in the programme.

Afinion

Hirst 2017

Twelve studies were included in the meta-analysis of studies evaluating Afinion devices. The mean bias ranged from -0.65% (SD 0.40) to 0.43% (SD 0.48) and the SD of mean bias ranged from 0.19% to 0.48%. The pooled mean bias was -0.06 (95%CI -0.21 to 0.09) and the pooled SD of mean bias was 0.31% (95%CI 0.25 to 0.36). Both estimates had very wide prediction intervals ranging from -0.66% to 0.53% and from 0.11% to 0.51%, respectively. Sensitivity analyses in which studies were excluded if the review authors had to make estimates from graphs produced similar result, while greater negative mean bias was reported when 1) studies at high risk of bias were excluded, -0.16% (95%CI -0.47 to 0.15); and 2) when only studies in NGSP or IFCC reference laboratories were included (n=4), -0.14% [95%CI -0.30 to 0.03).

Six studies reported imprecision (in NGSP units) based on replicate analyses of a single sample at low (<6.0%, 42 mmol/mol), medium (6.0-8.0%, 42-64 mmol/mol) or high (>8.0%, 64 mmol/mol) HbA1c: 2.5%, 1.5% and 2.0% CV, respectively, with total CV of 1.9%. Median between-laboratory CV from the 2016 CAP data was 3.1% (slightly higher than DCA 2000 which was 2.1%). The median between-laboratory CV of comparator laboratory instruments from the same survey ranged from 1.6% to 2.6% (table 4 in the paper).

The summary estimates of sensitivity and specificity at the DM diagnostic threshold of 6.5% (48 mmol/mol) (n= 8) were 96.5 (95%CI 91.0-98.7) and 99.7 (95%CI 91.0-100.0), respectively [2]. No subgroup analysis was conducted to investigate the analytical performance or diagnostic accuracy of the assay at the 6% (42 mmol/mol) HbA1c cutoff which differentiates between NDH and normoglycaemia.

Studies comparing Afinion to a laboratory instrument

Sixteen new studies evaluated the accuracy of Afinion devices against a laboratory instrument as reference standard. Of those, 10 used Afinion AS100, one used Afinion2 and 4 did not specify the model (Table 10).

Three papers reported on the accuracy of the new **Afinion Dx assay** which is the only HbA1c POCT test cleared by the USA Food and Drug Administration (FDA) for diagnosis of diabetes. Two of the papers reported on related studies [12, 13] and were combined in a single analysis [12]:

- Mean bias (using capillary blood) was -0.02% (SD 0.18) [13]
- Total CV ranged from 1.3% to 3.0% across different HbA1c levels
- Total error ranged from 2.9% to 4.8% and was 3.7% at 6.5% (48 mmol/mol) HbA1c (below the CAP threshold of ±6.0% of the target value)[12]

Sobolesky 2018 reported similar results with 97.1% of the Afinion Dx and 94.5% of the routine laboratory results falling within the target value of ±6.0% of the NGSP reference method results (across the range of HbA1c values 4.0%–15.0% HbA1c). The mean bias of Afinion Dx at 6.5% HbA1c (diagnostic threshold for DM) was -0.04% while the aggregate of laboratory methods displayed a mean bias of −0.06%. The total CV of the POCT results ranged from 0.7% to 2.1% while that of the routine laboratory methods ranged from 0.8% to 3.2% CV across the range of HbA1c values. The authors of both studies concluded that the accuracy of Afinion Dx is comparable to that of laboratory methods and suitable for diagnosis of DM. We note, however, that the studies reported in the first two papers were sponsored by the manufacturer (Abbott Rapid Diagnostics) and the authors were its employees [12, 13]; and the third study was published as a conference abstract only and reporting of funding may not have been required [19].

Ten studies assessed the mean bias of AS100 and one of Afinion 2 [10]; 7 of those studies reported results on imprecision (Table 10). The mean bias across studies ranged from -0.50% [30] to 0.56% [16]. This is consistent with the prediction interval in Hirst 2017 where 95% of future studies were predicted to report mean bias in the interval between -0.66% and 0.53%. Paknikar 2016 reported considerable fluctuation in the performance of the assay over a 3-year period, the mean bias varying from -0.40% to 0.10% (when the test was compared to BioRad Variant II) and from -0.50% to 0.15% (when compared to Tosoh G8) [30].

Two studies reported diagnostic accuracy data [32, 33]. Jain 2017 was conducted in the UK and included patients screened for DM in the community within a CVD prevention program (Table 7). No sensitivity and specificity estimates were reported in the paper; POCT and the laboratory method picked up 6 and 5 new patients with DM. The 5 patients were the same and the extra patient picked up by the POCT had HbA1c = 49 mmol/mol. The second study, Abbai 2017, was conducted in South Africa and reported sensitivity of 90.9% (95%CI 82.0 to 96.0) and specificity of 92.6% (95%CI 88.0 to 96.0). The cutoff was not explicitly stated and the study included convenience sample from another study. Both studies were at high risk of bias in at least one domain (Table 3).

The authors of 5 studies concluded that the performance of the device was acceptable or comparable to laboratory instruments [10, 18, 32-34]; one of them recommended using it for diagnosis [32]. Two studies concluded that the device was suitable for monitoring, but not for diagnosis [16, 35] and the authors of one study cautioned users about long-term variation in its performance while pointing out that similar variation was observed in the compared laboratory methods [30].

Studies analysing data from EQA programmes

Seven studies reported on the performance of Afinion devices using EQA data (Table 11). Three studies, two from the Netherlands and the EurA1c trial reported that the device met the 2 or 4 sigma criteria [11, 23, 27]. An analysis of data from the Equalis EQA (Sweden) showed that since 2008 between 60.0% and 97.0% of the reported results (>95% in the last two years) met the programme's pass criteria [26]. Similarity, pooled data from 7 surveys (2017 – 2018) from the Noklus EQA in Norway, which included 725 participants, mostly GP practices, achieved a pass rate of 98.2% to 99.7%.

The latter study also investigated a range of factors that might affect the performance of the participants including an instrument and reagent lot number, profession of the operator (in order of most to least skilled in laboratory work: biomedical laboratory scientist, medical secretary, nurse, GP/other), the number of patient samples performed per week (1–10, 11–15, 16–20, >20) and the frequency of running an internal quality control (IQC) (daily/weekly, monthly, when opening a new reagent kit, never). Two of these factors, "frequency of IQC" and "kit reagent lot number" were associated with good performance when the results from the 7 surveys were pooled. The result for the "frequency of IQC" is difficult to explain as monthly IQC was associated with better performance than daily/weekly IQC; and 5 lot numbers were associated with poorer results than the other reagent lots [29].

Paknikar 2016 (see the section on DCA) also reported variation in the device's performance when compared to the USA CAP proficiency testing programme criteria. As mentioned above considerable variation was observed not only for Afinion (and DCA) but also for the two laboratory methods used in the study (BioRad Variant II and Tosoh G8). The mean bias of Afinion ranged considerably at the high HbA1c level but was relatively consistent at low and mid-range HbA1c: from approximately -0.05% to 0.15% and from -0.10% to 0.10%, respectively (read off the graph).

The results from the CAP Survey data for 2019 and the first 5 samples in 2020 were based on a median of 88 laboratories for Afinion AS100 and 22 laboratories for Afinion 2 and covered the HbA1c range from approximately 5.0% (31 mmol/mol) to 10.0% (86 mmol/mol). The mean bias was within the current CAP acceptance limit of ±6.0% of the target value at all HbA1c levels and for both devices; the between-laboratory CV was ≤3.0% for all samples. Both Afinion devices consistently met the quality criteria of CAP and outperformed some of the laboratory methods enrolled in the programme.

A1c Now

Hirst 2017 [2] included 17 studies with mean bias ranging from -0.70% to 0.67% and the SD of the mean bias ranging from 0.21% to 0.96%. The pooled mean bias was -0.05% (95%CI -0.15 to 0.05) and the pooled SD of the mean bias was 0.53 (95%CI 0.44 to 0.62). The respective 95% prediction intervals ranged from -0.50% to 0.39% and from 0.12% to 0.94%, indicating considerable variation in the results of future studies with similar characteristics.

Year of publication did not affect significantly the estimate of the pooled mean bias. The subgroup analysis showed that there was no difference in the mean bias: 1) when the device was used in different settings or by different operators; 2) when venous or capillary blood was used. When studies at high risk of methodological bias (n=4) were excluded from analysis, the pooled mean bias was 0.09% (95%CI -0.25 to 0.43). Including only studies conducted in IFCC and NGSP reference laboratories (n=6) produced mean bias of -0.09% (95%CI -0.33 to 0.15). Total imprecision (based on 3 studies) was 2.9% CV. Summary sensitivity and specificity at 6.5% (48 mmol/mol) HbA1c (n=7) were 93.1% (95%CI 83.5-97.3) and 95.7 (95%CI 87.8-98.5), respectively.

We identified 3 new studies all reporting mean bias ≤0.30% (3.28 mmol/mol) (Table 12); one of them also reported CV = 3.3% [19]. Two of the studies reported on clinical risk agreement when the device used to screen for DM. Both studies were published as conference abstracts only and provided little information on methods and results. In the first study, patients were categories as normoglyceamic, NDH and DM, and change in risk was defined as 1 or 2 categories reclassification (e.g. from DM to NDH would be 1 and from DM to normoglyceamic would be 2 categories change). The clinical risk agreement was 94.0% (p=0.17); none of the changes were 2-categories change and the 1-category change was similar for POCT and laboratory methods. In the second study, the clinical risk agreement was 77.7% (vs 81.7% for the laboratory method, p=0.54).

Quo-Test

Hirst 2017 included 7 studies evaluating Quo-Test. The mean bias ranging from -0.73% to 0.29% and SD of the mean bias ranged from 0.17% to 0.49%. The pooled mean bias was -0.18% (95%CI -0.46 to 0.09) and the pooled SD was 0.30% (95%CI 0.22 to 0.38). The respective 95% prediction intervals ranged from -1.21% to 0.84% and from 0.01% to 0.59%. Total CV was 3.4% (n=2) and the summary estimates of sensitivity and specificity at 6.5% (48 mmol/mol) HbA1c (n=3) were 97.0% (95%CI 79.2-99.6) and 95.0% (95%CI 77.6-99.1), respectively [2].

We identified 3 new studies (Table 12). One reported mean bias of 0.13% (SD 0.59%) [36], the second [37] was a conference abstract which reported only coefficient of determination (r^2) which is not a good measure of agreement [38] and the third reported on the analysis of EQA data [29]. The first two studies concluded that the assay's performance was similar to that of laboratory instruments and should be considered for diagnosis of diabetes. The third study, Stavelin 2019, reported earlier in relation to DCA and Afinion, analysed data from the Noklus EQA programme (Norway). Out of the three POCT devices reported on in the study only Quo-Test failed to meet the study pass criteria. More specifically, all 12 primary care Quo-Test users met the criterion for bias (<0.3% HbA1c) but failed to meet the criteria for precision (<2.0% in NGSP units), with 9 of the 14 samples having a CV in the range of 2.1% and 5.7% [29].

Cobas b101

Hirst 2017 included 5 studies evaluating the accuracy of Cobas b101. Mean bias across studies ranged from -0.50% to 0.09%. The pooled mean bias was -0.13% (95%CI -0.36 to 0.09) with prediction interval from -1.03% to 0.76%. One study reported SD of the mean bias, 0.21% (95%CI 0.16 to 0.26) and one study reported total CV of 1.5% [2].

We identified 4 new studies none of which analysed EQA data [35, 39-41]. It is possible that the device was included in the European HbA1c Trial [11], but since all Roche devices were combined in one group and no results were reported specifically for Cobas b101, we did not included the data here. The mean bias in the 4 studies ranged from -0.20% to 0.10% (Table 12). A study from New Zealand reported on a series of assessments in relation to a faulty batch. After fixing the problem, an assessment of the assay (intended for use in the EPOCH trial, see Wells 2017 [14]) was carried out. The authors reported that 97.5% of POCT readings of 38 mmol/mol or lower would indicate a "real" result of 40 mmol/mol or lower; and 97.5% of POCT readings of 50 mmol/mol or higher would indicate a "real" result of 50 mmol/mol or higher, which was considered acceptable [40].

Precision was variable with two studies reporting higher CV at low HbA1c level: Criel 2016 reported 2.4% and 1.5% CV [39] and Toro-Crespo 2017 [35] 2.1% and 1.9% CV at low and high level HbA1c, respectively (intra-assay CV in NGSP units). Lyon 2017, on the other hand, reported lower imprecision with 1.6%, 1.0% and 1.7% CV at 5.3%, 7.0% and 13.4% HbA1c, respectively [41]. Kenealy 2019 reported 1.9% CV (in IFCC units) in the region of 40 and 51 mmol/mol and concluded that the

assay was suitable for monitoring and diagnosis provided testing is done within stringent quality assurance processes prior to and while in use [40]. Lyon 2017 also concluded that the device's performance is comparable to that of laboratory instruments [41] while Toro-Crespo 2017 recommended it for monitoring but not diagnosis [35].

Other devices

The results for devices evaluated in ≤2 new studies are presented in (Table 12). This included B-analyst, NycoCard, PixoTest, in2it (discontinued), DS5, RC20 (not available in the UK), I-Chroma, Quo-Lab, A1Care, the HbA1c 501, A1C EZ 2.0 and the 2 devices for which certification data were provided by the manufacturers: Allegro Analyser and Fincare HbA1c test. Also, Zhou 2018 [42] (a case-controlled study) reported the sensitivity and specificity and the positive and negative predictive values of A1C EZ 2.0 at different HbA1c levels, including 6% (42 mmol/mol) and 6.5% (48 mmol/mol) which are reported in Table 7.

Comparative studies

The results from studies comparing the performance of two or more devices are summarised in Table 13. We identified 18 such studies, one of which did not report accuracy results in sufficient detail and was excluded from the table [43]. Of the remaining 17 studies, 11 compared DCA and Afinion (3 of them also included Quo-Test, A1CNow+ and in2it, respectively); and only one study did not include either DCA or Afinion. The results were variable, with no clear advantage of DCA or Afinion, even when only the 5 studies reporting on EQA data were considered. One of the EQA studies reported that Quo-Test failed to meet the criteria for imprecision while the users of DCA and Afinion met both the criteria for bias and imprecision [29]. None of the studies mentioned the use of the Afinion HbA1c Dx assay, which is claimed to have superior performance and "*…US FDA clearance for use as an aid in the diagnosis of diabetes and in the identification of people at risk of developing diabetes.*" (p.2) [13].

Studies investigating the benefits and disadvantages of using HbA1c POCT in screening for NDH and DM

In addition to the literature review conducted by Schaffert 2016 [3], we identified 12 studies reporting on various aspects of the implementation and use of the HbA1c POCT in screening programmes (5 of these studies also reported on accuracy). We identified only one RCT which was directly relevant to the current review [14]. The rest of the studies were observational studies of limited applicability and/or at high risk of bias. The results from the literature review are summarised in Table 14 and the methods and results of the new primary studies are reported in Table 15. Below we discuss in more detail two of the studies: the RCT and an NHS DPP programme evaluation conducted by Barron 2019 (based on unpublished manuscript provided by the authors) [44].

The EPOCH trial was a cluster RCT of good methodological quality conducted in New Zealand. It looked at the impact of using HbA1c and lipids POCT (in addition to laboratory testing) on the completion of CVD risk assessment in primary care. Nineteen GP practices were randomised to 'POCT in addition to laboratory' and 'laboratory testing only'. Staff were trained by representatives of the POCT manufacturer (Roche Diagnostics NZ Ltd) and a comprehensive IQC programme was put in place following local guidelines. The primary outcome was completed CVD risk assessments; secondary outcomes included incomplete (but electronically saved) CVD risk assessments and time (in days) to completion. Qualitative interviews were carried out to investigate the expected (control arm) and real (intervention arm) experience of POCT [14].

Soon after its start, the trial was halted following an assay precision problem with the batch of Cobas b101 HbA1c discs (reported in another included study [40]). By the time the POCT was introduced, 90% of the eligible patients had already had a CVD risk assessment and the methodology was reconsidered adopting a non-inferiority hypothesis that the proportion of CVD risk assessment completions in the POCT practices would be no less than that in the control practices.

Having a POCT device within the practices made no difference to the completion of CVD risk assessments, the proportion of incomplete assessments and the time to completion. The nurses in the control arm expressed interest and expected benefits from using a POCT device. The main concern of those in the intervention arm was that "*POC testing was not a good use of their time*" (p.6). Having a POCT device changed the work flow, as nurses had to get a sample at the start of a patient consultation; to do two consecutive tests (lipids and HbA1c) each one taking 5 minutes and to wait for the result by the machine. Other issues included time for monthly quality assurance measures and suitable place for the device, so that it could be accessed by other members of the team. The results highlighted the importance of "*context, specifically usual policy, procedures, staff time and resources*" (p. 8).

Barron and colleagues analysed data from the UK NHS Diabetes Prevention Programme (DPP) which delivers behavioural interventions to prevent or delay the onset of type 2 diabetes in adults diagnosed with NDH. HbA1c was measured at referral using mainly laboratory methods and at the initial assessment using POCT if the referral HbA1c was done >3 months ago. Data from 73,703 participants referred to the programme over a period of >2 years were included in the analysis. The mean (SD) number of days between the two HbA1c measurements was 203 (120) days and the following POCT devices were used: Afinion, DCA Vantage and A1cNow+ [44].

The study found that the mean difference between referral and POCT HbA1c was -2.48mmol/mol (-0.23%) ($p<0.001$) with significant differences between devices. The SD of POCT HbA1c was 4.46mmol/mol (0.41%) and measurements in participants who were older, from more deprived areas and from Asian, black and mixed ethnic groups were associated with smaller HbA1c differences.

When only participants who had sufficient time to attend an intervention session (n=46,894) were analysed, 48% of HbA1c values were in the normal HbA1c range, 46% in the NDH range and 6% in the type 2 diabetes range. Participants with a normal POCT HbA1c result had significantly lower subsequent attendance at an intervention session compared to those in the NDH range (58% vs. 67%; $p<0.001$).

Due to the observational nature of the study and the high risk of selection bias, the above results should be interpreted with caution. The reasons for the observed difference between the laboratory and POCT measurements are unclear and, most likely, the result of multiple factors. They should not be attributed entirely to bias in the POCT devices and the contribution of other factors should also be considered, including the possible lack of harmonisation between laboratory methods and POCT, the changes in the condition that could occurred between the measurements and the conditions of use of the POCT devices. The Norfolk Diabetes Prevention Study, which we discuss in more detail later on, used only a single laboratory method and still observed considerable discrepancies in HbA1c measurements done a median of 40 days apart [45]. Nevertheless, the study by Barron and colleagues [44] shows the potential downstream consequences of inconsistent test results and highlights the importance of quality control, harmonisation and close collaboration with the local laboratory, and monitoring the impact of testing on patient outcomes.

The main benefits and concerns reported in the studies could be summarised as follows:

Immediate results: POCT has the potential to save time, reduce the number of visits and minimise the risk of non-attendance, which can be helpful in testing patients less likely to engage in a screening programme [46]. However, these expectations may not be fulfilled if the POCT is not well-integrated in the local practice and no dedicated time and resources are available [14].

Ease of use: Multiple studies reported that ease of use is an important consideration when the device is operated by non-laboratory professionals [3, 18]. A nurse-based evaluation reported that staff felt frustration due to several manual steps and the need of constant attention when using the NycoCard (ref. 35 in the literature review). Zhou 2017 also reported that training had greater impact on NycoCard users, which is a semi-automated device, compared to users of Afinion and DCA Vantage [43]. In another study, users preferred Afinion over DCA 2000+ because of easier sample loading (ref. 30 in the literature review). Yet another study concluded that "*InnovaStar HbA1c instrument requires users with laboratory experience*" (ref. 27 in the literature review) [3].

Relative analysis time: This was also considered an advantage even when the difference between devices was small. For instance, Deobald 2016 considered an advantage the fact that Afinion had faster turnaround time of 3 minutes compared to 6 minutes on DCA [18]. In another study, users preferred Afinion over DCA 2000+ because of faster analysis time [and easier sample loading] (ref. 30 in the literature review) [3]. Again, we found no studies reporting on the actual impact of such difference and its importance relative to other features of the devices.

Patient satisfaction and engagement: A few studies reported that patients were satisfied with a screening intervention or programme in which HbA1c POCT was used. The convenience of having the test done there and then was appreciated by patients, but since no comparison with laboratory testing was carried out, the importance of having a POCT rather than laboratory testing is difficult to determine [47-50].

Ergonomics and practicality: Small footprint and portability were mentioned in some studies as an obvious advantage of the POCT devices [3, 14, 18]. However, as the EPOCH trial shows, such expected advantages might be difficult to realise in real life without proper planning and reorganisation of the existing practice [14]. Also, in some cases, patients will require additional laboratory tests, such as a confirmatory test if the result is in the diabetic range; other blood tests (e.g. a lipid profile, renal function), or laboratory glucose testing.

Maintenance, training and quality assurance: Low maintenance was considered an advantage [3, 18] but, once more, the lessons from the EPOCH trial [14], the DPP programme evaluation [44] and the included accuracy studies highlight the importance of quality assurance that incorporates device maintenance and staff training, which could be more costly for a small number of POCT devices. Without such arrangements even the most accurate device may fail to fulfil its potential and inconsistencies in test results could have an impact on patient outcomes. Another aspect of this is the integration with the laboratory information management system which is not be possible for POCT used in the community.

Health economic evaluations

We found no health economic evaluations directly relevant to the focus of the review. One paper, El-Osta 2017 [46], reported on the cost-effectiveness of the NHS Health Check using a POCT device. Unfortunately, the device was the Alere Cholestech LDX Analyzer POCT device which measures cholesterol and blood glucose levels but not HbA1c. Therefore, the results from the study are not entirely relevant to the current review questions. It was an observational study and mathematical

model with micro-costing approach, which aimed to determine if using the Alere Cholestech LDX Analyzer is less costly than laboratory testing in delivering the NHS Health Check programme. Data were collected from 7 GP practices using POCT and two using a laboratory. The total expected cost of using POCT was lower than the laboratory-led pathway with savings of £29 per 100 invited patients up to the point of CVD risk score presentation. The main driver of these savings was the fact that POCT could deliver the assessment in one visit and minimise the non-attendance rates associated with laboratory testing. As it considered costs up until the initial CVD risk assessment and because it went no further, it implicitly assumed that the accuracy of POCT is equivalent to that of laboratory testing and that any downstream impacts are the same between POCT and laboratory testing [46].

Two US studies also reported on costs, but neither of them carried out a full economic evaluation. Bossart 2016 assessed the feasibility of DM screening by a dental hygienist and reported mean time and direct cost of using an HbA1c POCT device [48]. Lewandrowski 2017 investigated the impact of implementing HbA1c POCT on practice efficiency in an academic primary care practice and reported reduction in letters and calls to patients, a 50% reduction in follow-up tests per visit ($p = 0.044$) and a 38% reduction in follow-up visits due to abnormal test results ($p = 0.178$) with net financial benefit of $11.90 – $14.74 per patient visit [51].

In the EPOCH trial "*the majority of respondents (7/9 POC practices and 5/10 control practices) reported that they would not conduct POC testing if the cost of the consumables was borne by the practice*" (p. 11).

The literature review included two health economic evaluations both of which related to using HbA1c POCT for monitoring and not diagnosis. The authors stated that no such studies are currently available and further research is needed [14].

Guidelines

Table 16 provides a brief summary of the relevant points from the 9 guidelines we identified as relevant to the review [52-60]. Four guidelines (ADA 2009, IDF 2017, DAGDC 2009, NZMoH 2011) state that HbA1c should only be measured in laboratories [53, 56, 58, 59]. Both the PHE NHS Health Check Best Practice Guidance and the NHS DPP Guidance on the use of POCT HbA1c report recommendations on using POCT from the Medicine and Healthcare products Regulatory Agency (MHRA), which are not specifically for HbA1c [55, 60]. The DPP Guidance also recommends close collaboration with the local laboratory and stipulates that:

- "*POCT HbA1c should only be considered where there is evidence for cost effective implementation. There should be an investigation into projected workload, workflow and whether changes to local practice and the established laboratory service can meet these demands to circumvent the need for POCT HbA1c"*
- *"All parameters of the MHRA stipulated quality framework must be in place prior to implementation of the POCT HbA1c device"*
- *"All POCT HbA1c devices must have a clear process for internal quality control and be enrolled into an external quality assessment programme."* (p 7).

NICE guidance (2017) states that HbA1c testing, including POCT, should only be done by trained staff and follow expert consensus reports on use and conform to national quality specifications [54]. The Canadian guidance (DCCPGEC 2018) only recommends HbA1c POCT when laboratory testing is not

available for screening remote Indigenous populations. In these circumstances the testing must be part of a quality control programme where expertise and follow-up is available [57].

Recent American guidance (ADA 2020) does recommend diagnostic use of HbA1c POCT but only "*in settings licensed to perform moderate-to-high complexity tests.*" [59].

Discussion

We reviewed the research evidence pertaining to the accuracy and utility of HbA1c POCT devices when used in a non-laboratory setting to screen for DM and NDH in adults. We also reviewed selected guidelines that make specific recommendations for this. A systematic review of good methodological quality and focus similar to ours, included accuracy studies published before 2016 and reported the results of several meta-analyses [2]. Another study, a literature review conducted by the Diagnostic Evidence Co-operative Oxford and covering the period before 2016, reported on research findings related to the advantages and disadvantages of using such devices [3]. We used these two reviews as our starting point and limited our searches to studies published in English in the period between 1st January 2015 and 10th January 2020. Because of the small number of studies conducted in the relevant setting and patients, we relaxed our inclusion criteria and also included analytical validity studies in which the POCT was used in a laboratory. The meta-analyses reported by Hirst 2017 provide some evidence that the performance of HbA1c POCT devices may not be significantly different between those two settings [2].

We included 63 publications in total, 13 of which were conference abstracts: 45 accuracy studies, 12 studies investigating the benefits and disadvantages of POCT (5 of which also reported on accuracy) and 9 guidelines. Below we address each of the review questions considering both the volume and quality of evidence and the implications for clinical practice and research.

What is the accuracy of HbA1c POCT devices when used in a non-laboratory setting compared to a venous blood sample analysed in a laboratory setting?

Since accuracy encompasses both analytical validity and diagnostic accuracy, we first discuss the evidence pertaining to each of these aspects and then bring them together to consider the overall performance of the evaluated devices. Analytical validity refers to how well an instrument can measure the analyte of interest (in this case, HbA1c) compared to a reference method which, hopefully, has better accuracy. The two main aspects of analytical validity, bias (deviation from the true value) and precision (closeness of agreement between a series of measurements) are, usually, combined and compared to a single reference value, TAE. The current IFCC criteria stipulate that for routine laboratories the TAE is 10% (5 mmol/mol at 50 mmol/mol HbA1c) and >95% of the results should be within this limit. In sigma-metrics, this level of performance is defined as 2-sigma, with a higher level (4-sigma) required for reference laboratories. The criteria for POCT devices and routine laboratory methods are the same. Analytical validity is usually assessed in ideal laboratory conditions and, therefore, likely to overestimate the performance of the device compared to 'real life' clinical practice. Also, such studies used selected blood samples to cover the analytically relevant range of HbA1c values, but do not necessarily represent the distribution of values observed in a specific clinical setting (e.g. the Health Check programme).

Diagnostic accuracy, on the other hand, refers to the ability of the test to discriminate between individuals with and without the target condition (e.g. between patients with and without DM; or, between normoglyceamic, NDH and DM). Diagnostic accuracy studies should be conducted in conditions similar to those in which the test is intended to be used. The most common oucoumes reported by such studies are the true positive rate (sensitivity) and the true negative rate

(specificity). The former is an estimate of the proportion of patients with the target condition that will test positive while the latter is an estimate of the proportion of patients without the target condition that will test negative. Diagnostic accuracy is still a surrogate measure and does not tell us what impact different testing strategies have on patient outcomes and the health care system. Such higher level of evidence is provided by end-to-end studies and health economic evaluations. However, in the absence of the latter, accuracy estimates could be linked to evidence on downstream consequences to estimate the impact of using alternative diagnostic strategies, such as POCT versus laboratory testing.

Analytical validity of HbA1c

The largest proportion of included analytical validity studies evaluated DCA and Afinion, which probably reflects the fact that approximately 90% of the current market is held by these two instruments [61]. The distribution of new studies per device varied slightly from that observed in Hirst 2017, mainly in the fact that much smaller proportion of the new studies evaluated A1cNow compared to Hirst 2017 (6.7% vs. 27.9%, respectively). The number of studies evaluating the rest of the devices were small. No new studies evaluated Innovastar, HemoCue, Clover and A1cgear (included in Hirst 2017) and there was a small number of studies evaluating devices not included in Hirst 2017: PixoTest, A1C EZ 2.0, DS5, RC20 and I-Chroma. We also received information from the manufacturers that the in2it (BioRad) has been discontinued; RC20 (Sekisui Medical) is not available in the UK/Europe, but a new version intended for the European market will soon be available; and that the manufacturers of the Allegro Analyser (Nova Biomedical) and NycoCard (Abbott) have no claims for the use of these devices in diagnosis of DM or NDH.

Most analytical performance studies followed established protocols for evaluation of method comparison (e.g. CLSI EP-9) and precision (e.g. CLSI-EP5, often adapted). However, they used different reference methods, the variable performance of which was most likely the main contributor to the heterogeneity of results observed in both Hirst 2017 and the new studies. Variability in the performance of routine laboratory methods is well documented, both across different instruments and for the same instrument over time (e.g. in the CAP Survey and other EQA programmes). As demonstrated by Hirst 2017, the 'methodological noise' created by such variability makes it difficult to obtain accurate estimates of the analytical performance of HbA1c POCT.

For instruments with a larger number of studies (e.g. DCA, Afinion and A1cNow) Hirst 2017 investigated different sources of variability as well as the impact of the methodological quality of the studies. The results, however, remained contradictory and difficult to interpret. For instance, in the analysis of DCA, excluding studies in which the authors had to make estimates from graphs led to a slight improvement in the pooled mean bias, from -0.25% (95%CI -0.33 to -0.18) to -0.17% (95%CI -0.26 to -0.08); excluding studies at high risk of methodological bias led to even greater negative mean bias, -0.36% (95%CI -0.48 to -0.24); and including only studies conducted in NGSP or IFCC reference laboratories (n=6) led to a pooled mean bias of 0% (95%CI -0.16 to 0.15)! Since most of the studies had to be excluded from these analyses, the estimates are even more difficult to interpret because of the high level of statistical uncertainty, reflected in the wide prediction intervals. For most of the studies such analyses were not even possible, as the number of studies was too small and, in the best case scenario, only one or two studies could be considered at low risk of bias.

Given the above limitations, we decided not to pool the results, but to focus, if possible, on studies using superior reference standard (e.g. SRMs) and data from EQA programmes. The advantages of the latter include standardised procedures, high quality reference standard, data on the

performance of the test over time and direct comparison between POCT and laboratory instruments [61]. The results from such selected analyses indicate that:

- DCA Vantage and Afinion devices have the potential to meet the IFCC criterion of 2-sigma at TAE of 10% which means that >95% of their results will fall within that limit; TAE of 10% is 5 mmol/mol at 50 mmol/mol HbA1c and 4 mmol/mol at 40 mmol/mol HbA1c. A higher level of performance (e.g. 4-sigma at TAE of 10% or 2-sigma at TAE of 6%) was also reported, but was achieved less consistently (by smaller proportion of the evaluated instruments). EQA data showed that these devices perform as well as and, in some cases, better compared to some of the routine laboratory methods. Also, a new Afinion Dx assay has been cleared by the USA FDA for diagnosis of diabetes and initial evaluations suggest improved performance in terms of analytical accuracy and precision.
- Other devices have also shown potential to perform to the above standards, but the results are less conclusive and the evidence from good quality studies and EQA programmes is limited. Among those are new devices, such Allegro and Finacare, which achieved very good analytical performance in the IFCC certification process.

As stated earlier, meeting the analytical performance targets does not guarantee satisfactory diagnostic accuracy. Although EQA programmes in Australia, Sweden and Norway show that HbA1c POCT devices could perform as well as routine laboratory instruments in terms of analytical performance when used in a primary care setting and/or for screening of high risk patients, they do not provide data on the actual error rate at patient level.

Diagnostic accuracy of POCT HbA1c

In contrast to analytical performance, the evidence on the diagnostic accuracy of HbA1c POCT devices is practically non-existent! Hirst 2017 reported pooled sensitivity and specificity estimates for DCA (n=24), Afinion (n=8), A1cNow (n=7), QuoTest (n=3) and NycoCard (n=4). At 48 mmol/mol both sensitivity and specificity were >95.0% except for A1cNow which had sensitivity of 93.0% and the NycoCard which had specificity of 82.0% but sensitivity of 99.0%. Although these seem like very promising results, the analyses have a range of limitations and the estimates should be treated with caution. Firstly, two-by-two data were not reported in the papers, but extracted by the review authors from correlation and Bland-Altman plots, which increases the risk of error. Secondly, and more importantly, these were not diagnostic accuracy but analytical performance studies. Most of them were conducted in a laboratory, by a limited number of highly trained test operators (i.e. ideal conditions), and their results are likely to overestimate the performance of the test in a clinical setting. Thirdly, the distribution of HbA1c values used in such studies is artificially created with the intention to cover the whole analytically relevant range of values and is unlikely to reflect the distribution that will be observed if an unselected cohort of patients is tested. This may affect the proportions of different types of error (e.g. false negatives or false positives) and lead to biased estimates of sensitivity and specificity.

With regards to new evidence, we found only 8 diagnostic accuracy studies all of which were at high risk of bias and with limited applicability to our review. At 48 mmol/mol cutoff the reported sensitivities and specificities were around or even below 90.0%. Two studies reported the accuracy at 42mmol/mol for in2it (discontinued) and A1C EZ 2.0 at 42 mmol/mol. The former had sensitivity and specificity of around 85.0% while the latter had sensitivity of 96.0% and specificity of 77.0%. The study evaluating A1C EZ 2.0 was of relatively good methodological quality, except for the fact that it

was a two-gate (case-controlled) study and, therefore, is likely to overestimate the performance of the test.

In the NHS Health Check programme HbA1c is used for risk stratification and patients are categorised as normoglyceamic, NDH and DM. Therefore, instead of the false positive and false negative results that occur when we have a dichotomise test, we have a wider range of error types, some of which are between neighbouring categories (e.g. DM misclassified as NDH) and some are more extreme (e.g. normoglyceamic misclassified as DM). We found only two studies that attempted to capture this complexity. The first compared DCA Vantage, A1cNow+ and two laboratory instruments, Tosoh G8 and Roche Cobas c513, using Roche Cobas Ingegra 400 plus as a reference method. Clinical risk agreement was 100%, 94% and 98% for DCA, A1cNow+ and the laboratory instruments, respectively, and the difference was not significant (p=0.17).

The second study compared A1cNow+ and two clinical instruments, Roche Cobas Integra and Abbott Architect, to a reference instrument, in this case Tosoh G8. Risk was unchanged in 81.7% of the laboratory results and 77.7% of the A1cNow+ results; there were no 2-category changes (between noromglyceamic and DM) and the difference in 1-category changes between POCT and laboratory instruments was not significant (p=0.54). Both studies were available only as conference abstracts from which important details were missing and their results should be treated with caution.

Some indirect information on the diagnostic accuracy of HbA1c POCT as an entry test to an UK-based CVD prevention programme is provided by the Norfolk Diabetes Prevention Study (NDPS) [45]. This was an RCT which investigated the efficacy of lifestyle interventions in reducing the risk of transition to type 2 DM for people with NDH. In this study patients at high risk of NDH were identified through primary care databases and had HbA1c and fasting plasma glucose at baseline and a median 40 days later. HbA1c was measured using a high performance laboratory instrument, Affinity HPLC (Hb9210, Menarini Diagnostics) operated by a small number of trained staff in a single laboratory. No POCT was used at any point in the trial! For the purpose of our research question, the study could be considered a natural experiment, in which the near perfect testing conditions provide a benchmark for any POCT testing used in the same setting. In other words, we can safely assume that any HbA1c POCT done in the primary care or community will have inferior performance. The trial included 2208 patients with an elevated fasting plasma glucose or HbA1c at baseline; only in 61% of the patients both values were elevated! Forty days later:

- 24% of initially normoglyceamic patients progressed to NDH
- 21% of those in the NDH category regressed to normoglyceamic
- 7% of those in the NDH category progressed to DM
- 21% of those in the DM category regressed to NDH category.

Given the ideal testing conditions in the study, we can expect that HbA1c POCT devices will have worse performance even if their analytical validity is comparable to that of the instrument used in the NDPS! The authors of the study concluded that:

"*These current data suggest very many people entering national prevention programmes or trials based solely on a single elevated HbA1c are in fact at much lower risk than is assumed. Risk categorization using both fasting plasma glucose and HbA1c data, the use of paired baseline data prior to entry into clinical or research programmes, and awareness of diagnostic imprecision would mitigate some of these difficulties, and avoid overestimation of risk and a lifelong misdiagnosis.*"(p.9) [45]

The DPP programme evaluation reported by Barron and colleagues also showed that even when POCT is used for monitoring participants' progress in a diabetes prevention programme, discrepancies in the HbA1c results between referral and initial assessment might have an impact on participants' behaviour and affect their participation in the programme.

What delivery conditions are required to ensure that POCT devices used in non-laboratory settings are as accurate as possible?

The sub-group analyses reported in Hirst 2017 showed no difference in the analytical performance of the devices when used in primary care and in the laboratory, and between laboratory and clinical operators. However, as already discussed, these were studies of variable quality reporting highly heterogeneous results, and should be treated with caution. Zhou 2017 demonstrated that the effect of training on the analytical performance of the devices depends on their user-friendliness. Training had greater effect on the performance of NycoCard, which is a semi-automated device, but limited effect on Afinion and DCA Vantage.

Across the literature, there is an agreement that HbA1c POCT devices are no different than laboratory instruments and training, maintenance and internal and external quality control are essential for their optimal performance and consistency over time.

What are the benefits of using POCT for HbA1c over lab-based testing?

Only a small number of studies at high risk of bias and of limited applicability investigated the advantages and disadvantages of HbA1c POCT devices when used in screening for DM and NDH. Although the importance of features, such as immediate results, ease of use and relative analysis time were pointed out as obvious advantages, there is very little evidence on the relative importance of such features. The only study that compared HbA1c POCT and laboratory testing for screening in primary care was the EPOCH trial conducted in New Zealand. The study demonstrated the complexity of integrating POCT in the practice workflow and showed that without proper planning, additional resources and consideration of the specific context, many of the expected advantages may not materialise.

Features, such as portability and ease of use, offer clear advantage in settings where no laboratory services are available [22, 24]. However, in places where clinicians could choose between laboratory testing and POCT, a range of other factors, such as the need of IQC and EQA, should be considered and the decision should be based on the results of cost-effectiveness analysis, as recommended by the DPP Guidance [60]. We found no health economic evaluations directly related to our review question. The cost-minimisation study conducted by El-Osta 2017 had important limitations and addressed the use of a POCT device that measures cholesterol and blood glucose, not HbA1c. In the NDPS study discussed earlier 48.1% of participants with an impaired fasting glucose at baseline had normal HbA1c! This suggests that even if the results reported by El-Osta 2017 are valid, they still may not be generalizable to settings using HbA1c POCT. The need of proper cost-effectiveness studies that compare HbA1c POCT to laboratory testing in the relevant setting is clear and was emphasized in the literature review conducted by Schaffert 2016, which failed to find any relevant studies published up until 2016. The DPP programme evaluation by Barron and colleagues also shows how crucial is the collaboration with the local laboratories even when POCT HbA1c is used for monitoring of patients' progress within a prevention programme.

What are the current national and international guidelines on the use of POCT for classifying/ diagnosing non-diabetic hyperglycaemia (NDH) and type 2 Diabetes Mellitus (DM)?

Most of the included national and international guidelines stipulate that HbA1c POCT should not be used for diagnosis of DM, or should be used where no laboratory testing is available (e.g. remote locations) and only in association with a quality control programme. The American Diabetes Association stipulates that "… *point-of-care assays approved for diagnostic purposes should only be considered in settings licensed to perform moderate-to-high complexity tests.*" Both the NHS Health Check Best Practice Guidance and the DPP Guidance refers to the general recommendations for POCT made by the MHRA, which emphasise the importance of training, local protocols, quality control and involvement of the local pathology laboratory. In addition, the DPP Guidance clearly stipulates that POCT HbA1c should only be considered where there is evidence for cost effective implementation and the established laboratory service are not able to meet the demands for HbA1c testing.

Strengths and limitations of the review

We conducted the review following the PHE best practice guidance on rapid review methods and other methodological recommendations, such as those made by the Cochrane Collaboration. A review protocol was developed and agreed with the PHE team prior to undertaking the review. We conducted comprehensive searches and contacted manufacturers for titles that have been missed and unpublished data. Two reviewers screened both titles/abstracts and full tests and a second reviewer checked most of the extracted data for errors and inconsistencies.

The review has several limitations that need to be acknowledged. First, the searches were limited to studies published in English since 1st January 2015. For research evidence published before this date we relied on the results of a systematic review of accuracy studies and a literature review of studies investigating the advantages and disadvantages of HbA1c POCT. No specialist health economic databases were searched. We also limited the guidelines included in the review to international guidelines published in English and national guidelines from a small number of selected countries. Data were not extracted independently by two reviewers, although most of the data were double-checked. We did not contact all manufacturers HbA1c POCT; this was limited to devices evaluated in at least one new study and devices listed in the document provided by the PHE.

Conclusions

The identified research evidence suggests that DCA and Afinion has the potential to meet and, in some cases, exceed the criterion of 2-sigma required for HbA1c determination in routine laboratories, provided adequate training, maintenance, internal and external control arrangements are in place. This means that >95% of the HbA1c results will be within the TAE of 10% (5 mmol/mol at HbA1c level of 50 mmol/mol). Other devices also seem to have the potential to meet this target, but further good quality evidence is needed.

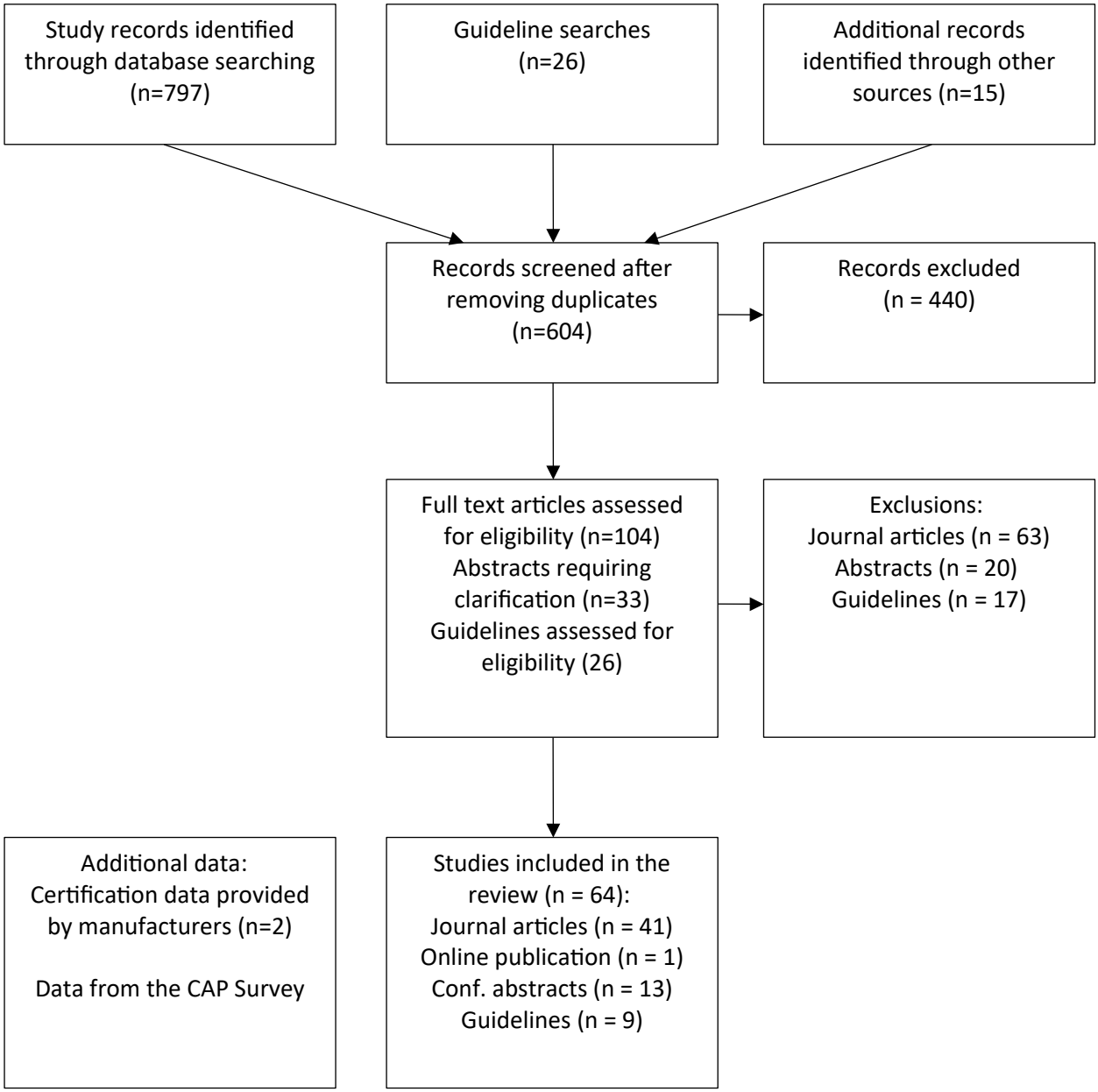
We found no good quality evidence on the diagnostic accuracy of HbA1c POCT when used to screen for DM and NDH, and are unable to report reliable estimates of sensitivity and specificity. However, in the NDPS study a significant proportion of patients were re-classified when two consecutive measurements 40 days apart were made, despite the fact that HbA1c was measured in nearly perfect conditions: using a single high performance laboratory instrument by a small number of highly qualified staff. They also found considerable disagreement between HbA1c and fasting plasma glucose at baseline and suggested that both tests should be used to minimise the proportion of patients misdiagnosed when using a single screening test.

The research evidence pertaining to the advantages and disadvantages of using HbA1c POCT when screening for DM and NDH is also limited, at high risk of bias and with limited applicability to the NHS Health Check programme. An RCT conducted in New Zealand showed that careful planning, additional resources and integration of the POCT in the practice workflow are required for the expected advantages to materialise. None of the included guidelines make specific recommendations for HbA1c POCT when used for screening, but the MHRA recommendations for POCT in general provide some guidance. The DPP Guidance recommends that HbA1c POCT should be considered only if there is evidence for cost effective implementation and when all parameters of the MHRA stipulated quality framework are in place prior to implementation of the POCT HbA1c device.

Future research studies should compare different screening strategies (e.g. HbA1c vs fasting plasma glucose using POCT vs laboratory testing) and aim to go beyond diagnostic accuracy, considering their clinical and cost-effectiveness.

Figures

Figure 1 Flow chart of the selection of studies



Tables

Table 1 POCT HbA1c devices available from the NHS Supply Chain (as provided by PHE)

Suppliers	Product Description	Manufacturer Product Code (MPC)	Brand
Alere now known as Abbott	Afinion Instrument & Power Cord (UK&IE) and Alere Afinion HbA1c Test	SBUK0027 and 1116062	Afinion
BHR	The Eurolyser CUBE and The HbA1c test	BHR#EUR-CA-0100 and BHR#EUR-ST-0110	Eurolyser Cube
PTS/Chek Diagnostics	A1CNow®+	n/a	A1CNow®+
Ciga	Finecare HbA1c Test for use with the Suresign Finecare + Point of Care Analyser	FS-113 and HBA1C-W207(25)	Suresign Finecare
Nova Biomedical	Allegro Analyser and Reagent Cartridge Allegro A1c	566668 and 54215	Allegro
Roche	COBAS B 101 INSTRUMENT and COBAS B 101 HBA1C TEST	06378668190 and 06378676190	Roche
Siemens	DCA Vantage Analyser; DCA HBA1C DIAGNOSTIC CLAIM KIT 1x10 tests; DCA 1x10 test cartridges	10282970 and 06378676190 and 10311480	DCA Vantage

Table 2 Manufacturers who replied to our data request

Analyser (Company)	Reply	Data provided
RC20	RC-W is scheduled to be commercially available in/after October 2020 with function to analyse diabetes and thalassemia. A similar analyser, RC20 is commercially available in Japan with function to analyse diabetes only.	No publications or data on RC-W for the UK. (n=0)
in2it (BioRad)	The analyser has been discontinued and at the moment the company has not POCT devices for HbA1c	(n=0)
Allegro platform	Allegro-HbA1c has no claim for use in diagnosis of non-diabetic hyperglycaemia	IFCC HbA1c Certificate to demonstrate

(NOVA Biomedical UK)	and type 2 diabetes in a clinical (non-laboratory) setting. The platform is currently undergoing several international independent studies both in Europe and in the USA aimed at generating results for publication. Some results are expected to be published by the end of the year, but no results has been published as yet.	traceability to the IFCC Reference Measurement Procedure (n=1)
NycoCard (Abbott Rapid Diagnostics)	The intended use claim is limited to monitoring of diabetes.	(n=0)
Afinion (Abbott Rapid Diagnostics)	A list of the most recent research and EQA data publications provided.	One additional research paper, Torregrosa 2015, not captured by the searches identified and included in the review (n=1)
Pixotest A1c (iXensor Co LTD)	The plan to arrange local trial programs on Pixotest A1c & lipid in England has been delayed due to the Covid-19 outbreak. Not able to provide peer-reviewed studies or EQA data published in English at the moment.	(n=0)
A1Care (i-SENS GmbH)	They have passed the precision, accuracy and bias test of A1Care from ISALA in 2019 and are currently writing a paper on the evaluation of the A1Care. Currently, A1Care system is in development stage and the company is planning to launch the equipment to Europe by next year.	(n=0)
Eurolyser CUBE (EuroLyser)	Not aware of any publications.	(n=0)
Finecare HbA1c	No published European studies available, but provided independent EQA data published in English (National Health Commission Clinical Laboratory, China, 2020) and certificates showing that the test has been certified with the IFCC (sigma diagrams for 2018 and 2020) and NGSP (2017 and 2018, only certificates, no numerical data). Also, no cost-effectiveness studies but provide the cost of the Finecare analyser and tests and supplementary items.	(n=3)

Table 3 Quality assessment of diagnostic accuracy studies

	Assessment of the risk of bias								Assessment of applicability	
Study	**Consecutive or random sample**	**Case controlled design avoided**	**Prospective study design**	**Acceptable reference standard**	**Time between IT and RS**	**All received same RS**	**All included**	**Funding by manufacturer**	**Applicability: Participants**	**Applicability: Index test**
Abbai 2017, South Africa [32]	no	yes	yes	no	yes	yes	yes	no	no	no
Gomez-Peralta 2016, Spain [62]	yes	yes	yes	no	yes	yes	yes	yes	no	unclear
Jain 2017, UK [33]	unclear	yes	unclear	no	yes	yes	yes	no	yes	unclear
Lynn 2018, USA [15]	no	yes	no	no	yes	yes	no	no	unclear	unclear
Moskowitz 2017 (CA), USA [63]	unclear	unclear	unclear	no	unclear	yes	yes	no	unclear	unclear
Szablowski 2018 (CA), USA [19]	unclear	unclear	unclear	no	unclear	yes	yes	no	unclear	unclear
Valdez-Gonzalez 2018, Mexico [21]	unclear	yes	yes	no	unclear	yes	no	no	no	yes
Zhou 2018, China [42]	unclear	no	yes	no	unclear	no	yes	no	yes	no

Table 4 Quality assessment of analytical validity studies

	Assessment of the risk of bias					Assessment of applicability
Study	**Prospective study design**	**HbA1c approx. range 5-11% (31–97mmol/mol)**	**Reference standard: IFCC/NGSP SRM**	**Time between IT and RS <24 hrs**	**Funding by manufacturer**	**Index test**
Abbai 2017 [32]	yes	unclear	no	yes	no	no
Arnold 2019 [12]	yes	yes	yes	unclear	yes	no
Arnold 2019 [13]	yes	yes	yes	unclear	yes	no
Chen 2018 (CA) [64]	unclear	yes	no	unclear	no	unclear
Cheng 2019 [65]	yes	yes	no	No	no	unclear
Criel 2016 [39]	yes	yes	yes	yes	yes	no
Deobald 2016 (CA) [18]	unclear	yes	no	unclear	no	no
Diaz-Garzon 2017 (CA) [17]	unclear	no	no	unclear	no	no
Dubach 2019 [66]	yes	yes	no	yes	no	unclear
Fellows 2019 [67]	no	yes	no	yes	no	yes
Grant 2017 [36]	unclear	yes	no	unclear	no	no
Hamada 2016 (CA) [68]	unclear	unclear	no	unclear	no	unclear
Jain 2017 [33]	unclear	yes	no	yes	no	unclear
Jones 2016 (CA) [37]	unclear	yes	no	unclear	no	unclear
Kenealy 2019 [40]	yes	yes	no	unclear	yes	yes
Lenters-Westra 2018 [10]	yes	yes	yes	yes	yes	no
Lynn 2018 [15]	no	yes	no	yes	no	unclear
Lyon 2017 (CA) [41]	unclear	yes	no	unclear	no	unclear
Manthei 2017 [69]	no	yes	no	yes	no	yes
Moskowitz 2017 (CA) [63]	unclear	unclear	no	unclear	no	unclear
Nathan 2019 [34]	unclear	yes	no	yes	yes	yes
Paknikar 2016 [30]	unclear	yes	no	yes	no	unclear
Razi 2016 [70]	unclear	yes	yes	No	no	unclear

	Assessment of the risk of bias					Assessment of applicability
Study	**Prospective study design**	**HbA1c approx. range 5-11% (31–97mmol/mol)**	**Reference standard: IFCC/NGSP SRM**	**Time between IT and RS <24 hrs**	**Funding by manufacturer**	**Index test**
Saxton 2018 [16]	yes	n/a	no	unclear	yes	no
Sobolesky 2018 [71]	unclear	yes	yes	no	yes	no
Springer 2016 (CA) [72]	unclear	unclear	no	unclear	no	unclear
Swensen 2016 (CA) [73]	unclear	yes	yes	unclear	no	unclear
Szablowski 2018 (CA)	unclear	yes	no	unclear	no	unclear
The EurA1c Trial 2018 [11]	n/a	yes	yes	yes	no	unclear
Torregrosa 2015 [9]	unclear	yes	no	no	yes	no
Toro-Crespo 2017 [35]	unclear	yes	no	unclear	no	unclear
Valdez-Gonzalez 2018 [21]	yes	yes	no	unclear	no	yes
Vargas 2019 [20]	yes	yes	no	no	no	no
Zhou 2017 [43]	yes	yes	yes	n/a	no	no
Zhou 2018 [42]	yes	yes	no	unclear	no	no

Table 5 Quality assessment of the included guidelines

AGREE II criteria / Authors, Year, Country	ADA, 2009, International[1]	Diabetes Australia Guideline Development Consortium, 2009, [illegible]	Ministry of Health, 2011, New Zealand[3]	IDF, 2017, International[4]	NICE, 2017, UK[5]	Diabetes Canada, 2018, Canada[6]	ADA, 2020, US[7]	NHS DPP[8]
The overall objective(s) of the guideline is (are) specifically described								
The health question(s) covered by the guideline is (are) specifically described								
The population (patients, public, etc.) to whom the guideline is meant to apply is specifically described								
The guideline development group includes individuals from all relevant professional groups								
The views and preferences of the target population (patients, public, etc.) have been sought.								
The target users of the guideline are clearly defined								
Systematic methods were used to search for evidence								
The criteria for selecting the evidence are clearly described								
The strengths and limitations of the body of evidence are clearly described								
The methods for formulating the recommendations are clearly described								
The health benefits, side effects, and risks have been considered in formulating the recommendations								
There is an explicit link between the recommendations and the supporting evidence								
The guideline has been externally reviewed by experts prior to its publication.								
A procedure for updating the guideline is provided								

AGREE II criteria / Authors, Year, Country	ADA, 2009, International[1]	Diabetes Australia Guideline Development Consortium, 2009, [illegible]	Ministry of Health, 2011, New Zealand[3]	IDF, 2017, International[4]	NICE, 2017, UK[5]	Diabetes Canada, 2018, Canada[6]	ADA, 2020, US[7]	NHS DPP[8]
The recommendations are specific and unambiguous								
The different options for management of the condition or health issue are clearly presented								
Key recommendations are easily identifiable								
The guideline describes facilitators and barriers to its application								
The guideline provides advice and/or tools on how the recommendations can be put into practice								
The potential resource implications of applying the recommendations have been considered								
The guideline presents monitoring and/or auditing criteria								
The views of the funding body have not influenced the content of the guideline	FNR			FNR				
Competing interests of guideline development group members have been recorded and addressed		NSF	NSF		NSF		NSF	

Legend: Green – criterion met; Yellow – information not found; Red – criterion not met; FNR – Funder not reported; NSF – No statement found;

[1]International Expert Committee Report on the Role of the A1C Assay in the Diagnosis of Diabetes 2009 https://care.diabetesjournals.org/content/32/7/1327, [2]National Evidence Based Guideline for Case Detection and Diagnosis of Type 2 Diabetes 2009 http://static.diabetesaustralia.com.au/s/fileassets/diabetes-australia/af2389ea-8f61-4c54-82d6-77ab07f03597.pdf, [3]Guidance on the management of Type 2 diabetes 2011 https://www.moh.govt.nz/notebook/nbbooks.nsf/0/60306295DECB0BC6CC257A4F000FC0CB/$file/NZGG-management-of-type-2-diabetes-web.pdf, [4]International Diabetes Federation IDF Clinical Practice Recommendations for managing Type 2 Diabetes in Primary Care 2017 https://idf.org/e-library/guidelines/128-idf-clinical-practice-recommendations-for-managing-type-2-diabetes-in-primary-care.html, [5]Type 2 diabetes: prevention in people at high risk https://www.nice.org.uk/guidance/ph38, [6]Diabetes Canada Clinical Practice Guidelines Expert Committee 2018 Clinical Practice Guidelines http://guidelines.diabetes.ca/docs/CPG-2018-full-EN.pdf, [7]Classification and Diagnosis of Diabetes: Standards of Medical Care in Diabetes https://doi.org/10.2337/dc20-S002; [8] The use of POCT HbA1c devices in the NHS Diabetes Prevention Programme: Recommendations from an expert working group commissioned by NHS England https://www.england.nhs.uk/wp-content/uploads/2016/07/poct-paper.pdf

Table 6 Quality assessment of the included cost-minimisation analysis

Critical appraisal question	**El-Osta 2017 [46]**
Is the chosen time horizon appropriate to include relevant costs and consequences?	Y
Are all important and relevant outcomes for each alternative identified?	Y
Are all important and relevant costs for each alternative identified?	Y
Has any adverse effect of the intervention been captured?	N*
Is the structure of the model consistent with a coherent theory of the health condition under evaluation?	Y
Has the unit of representation been given?	Y
Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	Y
Are all future costs and outcomes discounted appropriately?	NA**
Are all important variables, whose values are uncertain, appropriately subjected to sensitivity analysis?	Y
Do the conclusions follow from the data reported?	Y
Has heterogeneity been dealt with by running the model separately for different sub-groups?	N
Has a discussion about the inclusion/exclusion of assumptions affecting the structure of the model been included?	Y
Have limitations and strengths been acknowledged/discussed?	Y
Have methodological uncertainties been addressed by running alternative versions of the model with different methodological assumptions?	N
Is there evidence that structural uncertainties have been addressed via sensitivity analysis?	Y
Are any counterintuitive results from the model explained and justified?	Y
Have the results of the model been compared with those of previous models and any differences in results explained?	NA***
*No difference in accuracy between the laboratory testing and POCT has been assumed. **The time horizon is <1 year. ***This is the first analysis of this kind.	

Table 7 Diagnostic accuracy studies

Study	Device	Sample size	Cutoff in mmol/mol (in %)	Sensitivity (%)	Specificity (%)	Other measures	Comment
Abbai 2017, South Africa	Afinion AS100	308	Unclear, classified only as 'normal' and 'abnormal'; "using target levels as cutoffs"	90.9 (95%CI 82.0 to 96.0)	92.6 (95%CI 88.0 to 96.0)	92.2% correctly classified	Convenience sample of adults ≥50 participating in the SHIOP study
Jain 2017, UK	Afinion AS100	113	≥48 (≥6.5)	n/a	n/a	POCT and lab methods picked up 6 and 5 new patients with DM, respectively from the CVD screening group; 5 of them was the same; the false positive patient had HbA1c 49 mmol/mol	Directly relevant UK study; only patients screened for DM included here
Lynn 2018	DCA (model not specified)	115	≥48 (≥6.5)	88.6	96.3	10.3% of individuals would be missed if one used the POCT method to diagnose diabetes.	It is also important to recognize that the difference of 0.2% in A1C measurement may be more significant when it is close to the diagnostic cutoff point of 6.5%.
Valdez-Gonzalez 2018	DCA Vantage		≥53 (≥7)	0.96 (95%CI 0.95 to 0.98)	0.89 (95%CI 0.86 to 0.92)	Agreement 93.5% (control of	Patients with type 2 DM in family

						DM2, goal <7%)	medicine units
Szablowski 2018 (CA)	DCA Vantage, A1CNow+	48	HbA1c categories < 5.7%; 5.7-6.4%; ≥ 6.5%)	n/a	n/a	Clinical risk agreement was 100% for the DCA, 94% for A1CNow+ and 98% for lab methods (p = 0.17).	Conference abstract, some details on methods are missing
Moskowitz 2017 (CA)	A1CNow+	94	0, 1 or 2 categories reclassification (based on HbA1c categories < 5.7%; 5.7-6.4%; ≥ 6.5%)	n/a	n/a	Risk unchanged in 81.7% of lab measures and 77.7% of POCT measures resulting in non-statistical category 1 difference (p=0.54)	Conference abstract, some details on methods are missing
Gomez-Peralta 2016	Bio-Rad in2it	187	A capillary HbA1c value ≥42 (≥6.0)	85.7	85.3		Patients at single centre ED; sens and spec for range of cutoff values from HbA1c 3% to 7.15%
			≥48 (≥6.5)	66.7	94.4		
Zhou 2018	A1C EZ 2.0	842	≥42.08	96.5	77.4	PPV 80.2; NPV 95.7	Case-controlled study
			≥47.54	76.1	86.6	PPV 85.0; NPV78.4	

*Table 8 Studies evaluating the analytical validity of DCA 2000/Vantage**

Study and country	Reference instrument (method)	POCT sample	Setting and POCT operator	N of samples (N of patients and diagnosis)	Mean bias (SD and/or LA) in mmol/mol	Mean bias (SD or LA) in % DCCT	Precision / CV (%)	Recommended for screening (other relevant conclusions)
Albeiroti 2018, USA (CA)	n/a (HPLC)	n/a	Laboratory	n/a (127 DM)	9.84	0.09 (higher at higher HbA1c levels)	n/a	n/a (good long term performance)
Deobald 2016, USA (CA)	BioRad Variant II Turbo Hemoglobin A1c assay	capillary	n/a	53 (32 DM and 21 non-DM)	-1.86 (1.75)	-0.17 (0.16)	Low: 2.8% High: 2.5%	n/a
Diaz-Garzon 2017, Spain (CA)	H8180; Arkray, Menarini (HPLC)	n/a	Hospital	NR (Using QC material daily for 3 months)	Low: -8.31 High: -18.47 (average -13.44)	Low (5.5%) - 0.76 High (11.1%) - 1.69 (average –1.23)	Low 2.26% High 2.73%	n/a
Dubach 2019, Switzerland	Tosoh G8 (HPLC)	venous	Outpatient clinic	100 (100 DM)	-2.3 (LA -7.8 to 3.2)	-0.21 (LA – 0.71 to 0.29)	n/a	n/a
Fellows 2019, USA	n/a	n/a	Primary care	42 (40 n/a)	Median 1.5 (range: – 26 to 52)	Median 0.15 (range: -2.4 to 4.8)	n/a	POCT results should be correlated with clinical findings and blood glucose when diagnosing DM
Lynn 2018, USA	Roche Tina-Quant assay (instrument not specified)	capillary	n/a	115 (115 with and without DM)	2.19 (range 0 to 34.97)	0.2 (range: 0 to 3.2)*	n/a	Not recommended for diagnosis. 10.3% of individuals would be missed

Study and country	Reference instrument (method)	POCT sample	Setting and POCT operator	N of samples (N of patients and diagnosis)	Mean bias (SD and/or LA) in mmol/mol	Mean bias (SD or LA) in % DCCT	Precision / CV (%)	Recommended for screening (other relevant conclusions)
								if one used the POC method to diagnose diabetes
Mackenzie-Feder 2016, Canada	Roche Integra 800CTS (TIIS)	capillary	Community-based screening for DM	25 (25 suspected HbA1c ≥5.7%)	-2.08 (SD 2.19)	-0.19 (SD 0.20)	n/a	Useful for screening for DM after adjustment for bias
Manthei 2017, USA	Tosoh G8 (HPLC)	n/a	ambulatory	1843 (n/a, n/a, over 37 months)	-2.62	-0.24 (27% of values outside ±6% of lab result)	n/a	Need to consider such substantial error in interpreting results and determining the suitability of testing in various clinical settings
Paknikar 2016, USA #96	Tosoh G8 or Bio-Rad Variant II (HPLC)	venous	laboratory	Vs Tosoh n=167, Vs Biorad n=449 (unselected, over 3 years period)	1) -6.56 to 1.09 (vs Bio-Rad) 2) -6.56 to –1.09 (vs Toshoh)	Approx. range: 1) -0.6 to 0.1 (vs Bio-Rad) 2) -0.6 to -0.1 (vs Tosoh)	n/a	n/a (All who rely on POC methods as well as on central laboratory measurement of HbA1c must understand the potential limitations of these assays)
Saxton 2018, Peru (USA)	Premier Hb9210 (HPLC)	venous	Laboratory (samples taken	203 (203, undiagnosed	4 (95%CI 3 to 4)	0.32 (95% CI 0.30 to 0.35)	4.01	Imprecision and bias were not low

Study and country	Reference instrument (method)	POCT sample	Setting and POCT operator	N of samples (N of patients and diagnosis)	Mean bias (SD and/or LA) in mmol/mol	Mean bias (SD or LA) in % DCCT	Precision / CV (%)	Recommended for screening (other relevant conclusions)
			in Peruvian Amazon)	12-75 years old)				enough to recommend either POC analyzer for HbA1c determinations in this setting.
Springer 2016, USA (CA)	Sun Tosoh (HPLC)	venous	outpatient setting	50 (n/a, n/a)	1.4 (SD 2.19)	0.128 (SD 0.2)	n/a	n/a
Szablowski 2018 USA (CA)	Roche Cobas Integra 400 plus	capillary	n/a	48 (48, n/a)	-0.44	-0.04 (averaged paired bias with Afinion; 100% clinical risk agreement)	1.7 (NGSP units)	As accurate as the clinical laboratory methods
Torregrosa 2015, Spain	HA 8160 (Menarini Diagnostics)	venous	Laboratory	30 for accuracy and 10 for precision (n/a, n/a)	3.06 (95%CI -3.72 to -2.51)	−0.28 (95%CI −0.34 to −0.23; LA 0.05 to -0.62)	3.74 (NGSP)	Did not meet the NGSP criterion for precision
Vargas 2019, Ecuador (Argentina, USA)	BioRad Variant II Turbo (HPLC)	venous	laboratory	114 (24 T2DM, 90 no T2DM)	-0.2 (3.2)	-0.02 (0.29)	Low: 0.881% High: 1.786% (in mmol/mol)	DCA-Vantage was comparable to HPLC assay
Zhou 2017, China	(D10 [Bio-Rad Laboratories, Inc.], Tosoh G8 HbA1c Variant [Tosoh Corporation], and Premier Ultra 2)	venous	laboratory	5 (5, n/a)	n/a	n/a	n/a	External mathematical calibration and training could improve analytical performance

Study and country	Reference instrument (method)	POCT sample	Setting and POCT operator	N of samples (N of patients and diagnosis)	Mean bias (SD and/or LA) in mmol/mol	Mean bias (SD or LA) in % DCCT	Precision / CV (%)	Recommended for screening (other relevant conclusions)
Valdez-Gonzalez 2018, Mexico	Variant II BioRad (HPLC)	capillary and frozen venous samples	nurses in 25 family medicine units	1103 (1103 DM, 665 [60.3%] fresh samples)	2.95 (8.52)	0.27 (0.78)	n/a	n/a
*DCA Vantage was used in all but one study, Lynn 2018, which did not specify the model								

Table 9 Studies reporting on the analytical validity of DCA Vantage based on EQA data

Study and country	EQA scheme	Setting	Length of evaluation covered	N of POCT sites	Criteria	N (%) of sites meeting criteria	Bias and precision	Conclusions and recommendations
Albeiroti 2018 (CA) [25]	CAP PT	Nurses at hospital's diabetes clinic	3 years	1	CAP PT criteria	n/a	**Mean bias relative to:** NGSP: 0.11% Peer mean: 0.20%	Good performance
Delatour 2019 The Netherlands [27]*	The Instand e.V. EQA scheme	Laboratory	January (1 sample), March (2 samples) and June (1 sample) 2017	Unclear (approx. 450 laboratories in total took part in the study)	IFFC TAE ±5mmol/mol (10%)	n/a	Mean bias using fresh whole blood for 4 HbA1c levels in mmol/mol (% NGSP) (read of graph): 34 (5.3%): approx. 2 49 (6.6%): approx. 1.5 58 (7.4%): close to 0 90 (10.3%): -5 Overall performance >2σ	Non-commutability of EQA materials consistently resulted in a positive shift in bias, i.e. in overestimating bias
Lenters-Westra 2017 The Netherlands [23]	SKML EQAS	Laboratory	March 2016	12	IFFC TAE ±5mmol/mol (TAE 6%, 10%)	**TAE 10%** >2σ: 11 (91.7) >4σ: 5 (41.7)	**Mean σ (range)** **TAE 10%:** 5.1 (1.6 to 15.6)	Very good performance attaining the

Study and country	EQA scheme	Setting	Length of evaluation covered	N of POCT sites	Criteria	N (%) of sites meeting criteria	Bias and precision	Conclusions and recommendations
						TAE 6% >2σ: 5 (41.7)	**TAE 6%:** 2.4 (-0.8 to 7.0)	international guidance target of >2σ at TAE of 10%.
Mackenzie-Feder 2016 Canada [22]*	CEQAL	Mobile diabetes clinic	3 times	1	N/a (study tested the statistical significance of mean bias)	n/a	**Mean bias (SD)** -0.05 (0.17) (p = 0.262)	POC capillary measurement did not perform as well in the field as in the laboratory, but the bias was correctible, and the margin of error was small enough that the authors found the test clinically useful.
Nordin 2018 Sweden [26]**	Equalis	POC laboratories	2008 2017 (12 times per year until 2003; then 10 times per year)	DCA Vantage: 2014 – 2016 >4000, 2017: 3977	TAE of 3.6 mmol/ mol (0.33 NGSP%) at HbA1c level of 50 mmol/mol (6.7 NGSP%)***	Since 2008 varied between approx. 80% and 93% of samples meeting criteria	**Mean bias (±95%CIs)** For the period 2007 to 2017 (in mmol/mol): ranged from -1.44 (0.07) to 2.07 (0.06); the bias was consistently negative until 2016 and 2017 [1.39 (0.06) and 2.07 (0.06), respectively]	Showed fluctuation over time but consistently below the target of 95% of samples meeting the criteria
Orvim Solvik 2018 Norway [28]	Noklus	Pharmacies and GP practices	3 EQA surveys, October	7 pharmacies participated in EQA;	**Trueness:** Target interval ±0.1% HbA1c:	**Pharmacies:** "Very good": 56-100% for	From sample 1 in survey 1 CV% within (95% CI) pharmacies: 1.06 (0.69,	Pharmacies and GP practices had comparable

Study and country	EQA scheme	Setting	Length of evaluation covered	N of POCT sites	Criteria	N (%) of sites meeting criteria	Bias and precision	Conclusions and recommendations
			2016 to April 2017 (IQA data also reported but not included here)	441, 424, and 402 GP offices participated in the same EQA surveys	Very good ±2%, Acceptable 2% - 5.4%, Poor >5.4% of target interval **Precision:** Very good ≤0.2, Acceptable 0.3 to 0.4, Poor ≥0.5 between duplicate measurements	trueness and 71-100% for precision. No "poor" results **GP practices:** "Very good" 75-87% for trueness and 84-94% for precision. 0.23-1.24% and 0.50-2.04% "poor" for trueness and precision, respectively	1.90) vs GP offices: 2.04 (1.91, 2.18) and from sample 2 in survey 2, CV % between (95% CI) pharmacies: 0.75 (0.48, 1.40) vs GP offices: 1.75 (1.64, 1.88)	performance (following training of pharmacists); compliance was modest but might be due to the short duration of the study
Paknikar 2016 USA [30]	CAP PT	n/a	3 years	n/a (but probably 1)	NGSP	n/a	**Mean bias range:** Low: -0.15 to 0.1 Mid.: -0.25% to 0.15% High: - 0.20 to 0.4 (read off the graph)	Healthcare professionals should understand the limitations of these assays
Schimenes 2019 (CA) USA [31]	EQA using NGSP samples	n/a	n/a	Unclear but probably 1	<3.5% bias to NGSP samples	1	Mean bias 1.59%	n/a (study conducted by manufacturer)
Shephard 2017 Australia [24]; DCA 2000 and Vantage	QAAMS PT and QC	Manage glycaemic control and diagnosis of	15 years	Steady rise in POC devices from 45 in 1999 to 200 in	PT: 0.4% up to 6.7% Hba1C and 6% at concentrations>	See next column	**Bias:** PT: Averaged 89.5% (SD 5.5; range 77–96%) of results from 2002 to 2016 and 94.0%	HbA1c POC testing in QAAMS has remained analytically sound,

Study and country	EQA scheme	Setting	Length of evaluation covered	N of POCT sites	Criteria	N (%) of sites meeting criteria	Bias and precision	Conclusions and recommendations
		DM in Indigenous people		2016	6.7% HbA1c		(SD 1.3; range 92–96%) from 2009 to 2016 when the DCA Vantage was introduced **Median imprecision** across device operators averaged 2.81% (SD 0.50; range 2.2 to 3.9%) from 2002 to 2016 and 2.44% (SD 0.22; range 2.2 to 2.9%) from 2009 to 2016 **Min and max %CV at 6.8% HbA1c (2012 – 2016)** 2.49% to 2.85%	matched the quality achieved by Australasian laboratories and met profession-derived analytical goals for 15 years
The EurA1c Trial Group 2018 (Europe) [11]; DCA 2000 and Vantage combined	European HbA1c trial	Laboratories	n/a	158 devices using fresh whole blood; 6 using lyophilized material	5 mmol/mol (0.46%) at the 2σ level at 50 mmol/mol (6.7%) HbA1c	n/a	**Fresh whole blood:** Bias (between-lab CV) 0.6 mmol/mol (3.6%) [0.06% (2.4%) in NGSP units]; Overall performance >2σ	Met the 2σ when fresh whole blood was used but failed in lyophilized material
Stavelin 2019 Norway [29]	Noklus	n/a (but probably primary care)	January 2017-October 2018	459 primary care participants	Mean bias: <0.3% HbA1c (CAP criterion); within-lab CV <2%; between-lab CV <3.5% (NACB recommendations for diagnosis of	Bias: 100% Precision: 100%	Bias: <0.3% for each of the 14 samples. Pooled within-laboratory CVs were <2%	Performance similar to laboratory instruments. Suitable for diagnosing DM

Study and country	EQA scheme	Setting	Length of evaluation covered	N of POCT sites	Criteria	N (%) of sites meeting criteria	Bias and precision	Conclusions and recommendations
					DM)			
*DCA 2000 used in the study, ** Data also reported for DCA 2000 for 2005 to 2011 but not included here, ***95% of results reported from one participant, or for a method group over time (usually a year), should be within "1.5 mmol/mol ± 1.65 × 0.025 × HbA1c mmol/mol EQA externa quality assurance, SKML Stichting Kwaliteitsbewaking Medische Laboratoria Canadian External Quality Assessment Laboratory (CEQAL)								

Table 10 Studies evaluating the analytical validity of Afinion devices

Study and country	Reference instrument (method)	POC Instrument (if different from Afinion AS100) and sample	Setting and POCT operator	N of samples (N of patients and diagnosis)	Mean bias (SD and/or LA) in mmol/mol [other measures]	Mean bias (SD or LA) in % DCCT [other measures]	Precision / CV (%)	Recommended for screening (other relevant conclusions)
HbA1c Dx assay								
Arnold 2019, USA	Tosoh G8 (HPLC); Secondary reference laboratory (SLR)	Afinion AS100 (HbA1c Dx), capillary (and venous but results not included here)	Laboratory professionals at diabetes centre	The first 120 patients to fill a predetermined distribution were included in the assessment of bias and 170 in the assessment of imprecision	Mean bias (across HbA1c range): -0.23 (SD 2); Lower bound: -0.59, upper bound: -0.13	**Mean bias** across HbA1c range: -0.021 (SD 0.183); 95% lower confidence bound: -0.054; 95% upper bound: -0.012 **Relative bias** at HbA1c levels: 5%: -0.80 6.5%: -0.615 8%: -0.50	Across HbA1c levels repeatability ranged from 0.62% to 1.52% CV and between-lot component ranged from 0.00% to 1.20% CV. The total CV ranged from 0.62% to 1.93% (between-	≥97% of the results across sites fell within ±6% and 95% within ±5% of the reference method; none of the 3 cartridge lots produced less than 96% of results within the above criterion; meet the NGSP and NACB standards for accuracy and

Study and country	Reference instrument (method)	POC Instrument (if different from Afinion AS100) and sample	Setting and POCT operator	N of samples (N of patients and diagnosis)	Mean bias (SD and/or LA) in mmol/mol [other measures]	Mean bias (SD or LA) in % DCCT [other measures]	Precision / CV (%)	Recommended for screening (other relevant conclusions)
						12%: -0.333	operator or between-instrument components of variance assessed)	precision and suitable for diagnosis of DM (study funded by manufacturer)
Arnold 2019, USA Norway? (combined data from 3 studies: study A was Arnold 2019 #8; studies B and C assessed various components of variability)	See Arnold 2019 #8	Afinion AS100 (HbA1c Dx), capillary (and venous but results not included here)	Study B: POC sites with laboratories operated by trained medical professionals; Study C was done at the manufacturer's site	Study A (as above); Study B: 61 prospectively enrolled patients; study C: 4 venous blood samples (results not reported here)		HbA1c level: TE %* 5%: 4.75 6.5%: 3.69 8%: 3.4 12%: 2.87	%HbA1c level (mmol/mol): Total CV 4.00 – 5.99% (20.2 – 42.0): 2.03; 6.00-6.99% (42.1-52.9): 1.58; 7.00-9.99% (53.0-85.7): 1.49; ≥10% (85.8): 1.30	TE below the FDA criterion of ≤6% At the diagnostic cutoff of 6.5% (47.5 mmol/mol) HbA1c total imprecision was <2% CV and TE <4%
Sobolesky 2018, USA	Tosoh G8 (HPLC), SRL method	Afinion HbA1c Dx test, venous?	Laboratory	618 EDTA whole blood excess patient specimens with clinically indicated HbA1c testing		Total relative bias of −0.6% (SD −0.04%); Bias at 5% HbA1c −0.9% [−1.38%, −0.45%]; at 6.5% HbA1c −0.6% [−0.86%,	Total CV (data combined between 5 sites) at mean HbA1c 5.3%: 1.46% 6.5%: 1.35%, 9.8%: 0.85%	Accuracy and precision of the Afinion POC HbA1c was comparable to the laboratory HbA1c methods supporting the FDA's recent approval of the

Study and country	Reference instrument (method)	POC Instrument (if different from Afinion AS100) and sample	Setting and POCT operator	N of samples (N of patients and diagnosis)	Mean bias (SD and/or LA) in mmol/mol [other measures]	Mean bias (SD or LA) in % DCCT [other measures]	Precision / CV (%)	Recommended for screening (other relevant conclusions)
						−0.39%]: at 8% HbA1c −0.5% [−0.76%, −0.15%]		Afinion HbA1c Dx device for use in the diagnosis of diabetes
Other versions of the assay								
Abbai 2017, South Africa	ABX Pentra 400	Afinion AS100, venous	Trained study nurse and medical technologists at research clinic (SHIOP study)	308	2.45 (9.18) BALA: -15.53 to 20.44???	0.224 (0.84) BALA: –1.421 to 1.870 Sensitivity 90.9% Specificity 92.6%	n/a	Supports the use of Afinion AS100 for diagnostic use (no funding from manufacturer)
Criel 2016, Belgium	Adams Arkray HA-8160 (ion-exchange chromatography)	venous???	Outpatient diabetes clinic	40 (40)	Mean bias: -2.2 (-3.38; -1.05) Total error: -9.28 (-11.29; -7.26)	Mean bias: -0.17 (95% CI: -0.27; -0.06) Total error: -0.81 (95%CI: -0.99; -0.62)	IFCC units: Low: 3.7, High: 2.3 NGSP units: Low: 2.5 High: 1.7	Failed to meet precision criterion at low HbA1c level
Deobald 2016, USA (CA)	BioRad Variant II Turbo	capillary	Laboratory?	53 (32 DM and 21 non-DM)	-0.66 (1.86)	-0.06 (0.17)	Low: 1.8 High: 1.3	Acceptable precision, all

Study and country	Reference instrument (method)	POC Instrument (if different from Afinion AS100) and sample	Setting and POCT operator	N of samples (N of patients and diagnosis)	Mean bias (SD and/or LA) in mmol/mol [other measures]	Mean bias (SD or LA) in % DCCT [other measures]	Precision / CV (%)	Recommended for screening (other relevant conclusions)
								samples within 0.5% of target values
Dubach 2019, Switzerland	Tosoh G8 (HPLC)	venous	Diabetes clinic	100 (100 DM)	−2.9 [BALA -9.1 to 3.2]	-0.27 (BALA -0.84 to 0.30	n/a	n/a
Jain 2017, UK	ADAMSTM A1C HA8180V	capillary	Community CVD screening and diabetes clinic	255 (113 CVD screen, 142 diabetes clinic)	1st tertile: 2.1 (-1.95, 6.11) 2nd tertile: 0.6 (-3.52, 4.79) 3rd tertile: -0.1 (-4.9, 4.72)	1st: 0.19 (-0.18 to 0.56) 2nd: 0.06 (-0.32 to 0.44) 3rd: 0.01 (-0.45 to 0.43)	1.8 and 1.6 for within and between batch measurements, respectively (in NGSP units)	Compares well with laboratory-based methods
Lenters-Westra 2018, UK, The Netherlands	1)Roche Tina-quant Gen.3 HbA1c on Cobas c513 (immunoassay), 2)Premier Hb9210 (affinity chromatography HPLC) 3)Tosoh G8 (cation-exchange HPLC), 4)Abbott Enzymatic method on Architect c4000	n/a	Laboratory	40 (n/a)	≤2 mmol/mol at 48 and 75 mmol/mol, σ = 5.8	≤0.2% at 6.5 and 9.0%	≤1.7 in IFCC units (≤1.2 NGSP)	Excellent performance, passed the IFCC Task Force on HbA1c Standardization ≥ 2 sigma at 50 mmol/mol
Nathan 2019, USA (2 studies)	Premier Affinity (affinity	1) venous, 2) capillary	1) laboratory technician,	1) 300 (300), 2) 402 (402)	Study 1	Study 1: Mean bias: 0.01	**Study 1:** Low (5 to 6.9%):	n/a (performed acceptably under

Study and country	Reference instrument (method)	POC Instrument (if different from Afinion AS100) and sample	Setting and POCT operator	N of samples (N of patients and diagnosis)	Mean bias (SD and/or LA) in mmol/mol [other measures]	Mean bias (SD or LA) in % DCCT [other measures]	Precision / CV (%)	Recommended for screening (other relevant conclusions)
	chromatography HPLC)		2) 1 medical assistant, 5 nurses and 3 physicians at diabetes centre		Mean bias: 0.11 Study 2: Mean bias: 2.19	(relative bias 2.1%) Study 2: Mean bias: 0.2 (relative bias 3.14%)	1.00% Medium (7 to 8.9%): 1.18% High (≥9%): 0.78% **Study 2:** Low: 1.39% Medium: 1.42% High: 1.54%	realistic clinical conditions)
Paknikar 2016, USA	Tosoh G8 HPLC Analyzer or the Bio-Rad Variant II	venous	Unselected samples from patients participating in various clinical research protocols	606 (compared to BioRad), 198 (compared to Tosoh) over 3 years period	-4.37 to 1.09 (vs BioRad) and from -5.46 to 1.64 (vs Tosoh)	Ranged from -0.4 to 0.1 (vs BioRad) and from -0.5 to 0.15 (vs Tosoh)	n/a	n/a (long-term variability of both POC and laboratory methods)
Saxton 2018, Peru (USA)	Premier Hb9210 (HPLC)	venous	Laboratory (samples taken in Peruvian Amazon)	203 (203, undiagnosed 12-75 years old)	6 (95% CI 6, 6]) BALA: 2 to 11 mmol/mol),	0.56 (95% CI 0.53% to 0.59%), BALA: 0.16% to 0.97%	1.75	Imprecision and bias were not low enough to recommend either POC analyzer for HbA1c determinations in this setting
Springer 2016, USA (CA)	Sun Tosoh (HPLC)	venous	outpatient setting	50 (n/a, n/a)	0.33 (2.19)	-0.03 (0.2)	n/a	n/a
Toro-Crespo	Tosoh G8, SRL	venous	Laboratory	100 samples	0.19 (-0.61,	0.0178 (-	Inter-assay:	Considered suitable

Study and country	Reference instrument (method)	POC Instrument (if different from Afinion AS100) and sample	Setting and POCT operator	N of samples (N of patients and diagnosis)	Mean bias (SD and/or LA) in mmol/mol [other measures]	Mean bias (SD or LA) in % DCCT [other measures]	Precision / CV (%)	Recommended for screening (other relevant conclusions)
2017, Spain				from the routine laboratory workload	1.00)	0.0561, 0.0917)	2.13% Intra-assay: Low: 1.13% High: 1.97%	for the control, but not for the diagnosis of diabetes
Torregrosa 2015, Spain	HA 8160 (Menarini Diagnostics)	venous	Laboratory	30 for accuracy and 10 for precision (n/a, n/a)	-0.44 (95%CI -1.09 to 0.22)	−0.04 (95%CI −0.10; 0.02; LA - - 0.41 to 0.34)	1.8 (NGSP)	n/a (only states that the test met all NGSP performance criteria)
Zhou 2017, China	Bio-Rad D10, Tosoh G8 HbA1c Variant, and Premier Ultra 2	venous	laboratory	5 (5, n/a)	n/a	n/a	n/a	External mathematical calibration and training could improve analytical performa nce
*Total error %TE = % Bias + 1.96 x %CV (1 + %Bias/100)								

Table 11 Studies reporting on the analytical validity of Afinion devices based on EQA data

Study and country	EQA scheme	Setting	Length of evaluation covered	N of patient sites	Criteria	N (%) of sites/ results meeting criteria	Bias and precision	Conclusions and recommendations
Delatour 2019 The Netherlands [27]*	The Instand e.V. EQA scheme	Laboratory	January, March and June 2017; 1	Unclear (approx. 450 laboratories in	IFFC TAE ±5mmol/mol (10%)	n/a	Mean bias using fresh whole blood for 4 HbAic levels in mmol/mol (%	The Abott-Afinion System was the only assay to meet the

Study and country	EQA scheme	Setting	Length of evaluation covered	N of patient sites	Criteria	N (%) of sites/ results meeting criteria	Bias and precision	Conclusions and recommendations
			sample each	total took part in the study)			NGSP) (read off graph): 34 (5.3%): <1 49 (6.6%): <-1 58 (7.4%): <1 90 (10.3%): approx. -1.5 Overall performance: met the 4σ criterion	desirable performance criterion.
Lenters-Westra 2017 The Netherlands [23]	SKML EQAS	Laboratory	March 2016	3	IFFC TAE ±5mmol/mol (TAE 6%, 10%)	**TAE 10%** >2σ: 3 (100) >4σ: 1 (33.3) **TAE 6%** >2σ: 2 (66.7)	**Mean σ (range)** **TAE 10%:** 12.4 (3.1–30.3) **TAE 6%:** 6.6 (1.0–16.5)	In strict evaluation conditions, point-of-care test devices can perform as well as routine laboratory analyzers and may in future be considered suitable for use in the diagnosis of diabetes.
Nordin 2018 Sweden [26]	Equalis	POC laboratories	12 times per year until 2003; then 10 times per year	Alere Afinion: 2014 – 2016 from 2047 to 3825	TAE of 3.6 mmol/mol (0.33 NGSP%) at HbA1c level of 50 mmol/mol (6.7 NGSP%)*	Since 2008 varied between approx. 60% and 97% of results meeting criteria	**Mean bias (±95%CIs)** From 2005 to 2017 in mmol/mol, range: -3.07 (0.21) to 0.12 (0.10); negative bias except for 2015; improvement in the last 3 years, range: -0.18 to 0.13	Fluctuation over time but > 95% of samples meeting the criteria in the last 2 years

Study and country	EQA scheme	Setting	Length of evaluation covered	N of patient sites	Criteria	N (%) of sites/ results meeting criteria	Bias and precision	Conclusions and recommendations
Paknikar 2016 USA [30]	CAP PT	n/a	3 years	n/a	NGSP	n/a	**Mean bias range:** Low: -0.05 to 0.15% Mid.: -0.1 to 0.1% High: -0.45 to 0.1% (read off the graph)	Significant differences in measured values which are variable over time and should be considered in clinical decision making
Sobolesky 2018 [71] (Afinion HbA1c Dx test)	NGSP	5 clinical sites	June 2016 to June 2017	5	NGSP NGSP: TAE of ±6% across the measurement range; manufacturer certification requires that 37/40 (92.5%) of results must be within ±6% of the SRL	Over the range of 4.0%–15% HbA1c, 97.1% of POC results (and 94.5% laboratory results) fell within ±6% of the NGSP reference method	**Bias** (relative % of the measured value [95%CI]): At 5% HbA1c −0.9% [−1.38%, −0.45%] At 6.5% HbA1c −0.6% [−0.86%, −0.39%] At 8% HbA1c −0.5% [−0.76%, −0.15%] **Total CV:** Low: 1.46% Medium: 1.35% High: 0.85%	The accuracy and precision of the Afinion POC HbA1c method was comparable to the laboratory HbA1c methods supporting the FDA's recent approval of the Afinion HbA1c Dx device for use in the diagnosis of diabetes. (Study funded by manufacturer)
Stavelin 2019 Norway [29]	Noklus	Primary care, 95% of all were GP practices)	January 2017-October 2018 (7 surveys)	725 (90% response rate)	Within ±6% from the target value; Bias <0.3% (approx. 3 mmol/mol) Within-laboratory CV <2% (approx.	Participant pass rates for each survey varied from 98.2% to 99.7%	Bias varied between −0.17 and −0.01 %HbA1c in all surveys. The pooled within-laboratory CV varied from 1.3% to 1.5%, the between-laboratory CV varied from 1.5% to 2.1%	Afinion HbA1c fulfilled the analytical performance specifications and is robust in the hands of the users. It can therefore be used both in diagnosing and monitoring persons

Study and country	EQA scheme	Setting	Length of evaluation covered	N of patient sites	Criteria	N (%) of sites/ results meeting criteria	Bias and precision	Conclusions and recommendations
					3% mmol/mol), between-laboratory CV <3.5%			with diabetes mellitus, given that the instrument is monitored by an EQA system. Reagent lot was the only independent factor to predict good participant performance
The EurA1c Trial Group 2018 (Europe) [11]; DCA 2000 and Vantage combined	European HbA1c trial	Laboratories	n/a	76 devices (fresh whole blood)	5 mmol/mol (0.46%) at the 2σ level at 50 mmol/mol (6.7%) HbA1c	n/a	**Fresh whole blood:** IFCC bias (between-lab CV): -0.7 mmol/mol (3.4%) NGSP bias (between-lab CV): -0.06% (2.2%) Overall performance >2σ	Met the 2σ criteria
CA conference abstract, *95% of results reported from one participant, or for a method group over time (usually a year), should be within "1.5 mmol/mol ± 1.65 × 0.025 × HbA1c mmol/mol EQA externa quality assurance, SKML Stichting Kwaliteitsbewaking Medische Laboratoria Canadian External Quality Assessment Laboratory (CEQAL)								

Table 12 Studies evaluating other devices

Study and country	Reference instrument (method)	POCT sample	Setting and POCT operator	N of samples (N of patients and diagnosis)	Mean bias (SD and/or LA) in mmol/mol	Mean bias (SD or LA) in % DCCT	Precision / CV (%)	Recommended for screening (other relevant conclusions)
A1CNow+								
Moskowitz 2017, USA	Tosoh G8 (HPLC)	capillary	3 wellness centres	94 (94)	3.28	Mean bias: 0.3, Clinical risk unchanged in 77.7% POC measures vs 81.7% of lab measures (p=0.54)	n/a	At least as accurate as the 2 laboratory analysers and well within the 6% CAP guideline for bias. Risk stratification revealed no differences between the lab methods and A1CNow+ in classifying the patient state.
Springer 2016, USA (CA)	Sun Tosoh (HPLC)	venous	outpatient setting	50 (n/a, n/a)	0.33 (SD 2.19)	-0.03 (SD 0.2)	n/a	n/a
Szablowski 2018 USA (CA)	Roche Cobas Integra 400 plus	capillary	n/a	48 (48, n/a)	0.44	-0.04 (averaged paired bias with Afinion; 94% clinical risk agreement (p=0.17)	3.3	As accurate as the clinical laboratory methods

Study and country	Reference instrument (method)	POCT sample	Setting and POCT operator	N of samples (N of patients and diagnosis)	Mean bias (SD and/or LA) in mmol/mol	Mean bias (SD or LA) in % DCCT	Precision / CV (%)	Recommended for screening (other relevant conclusions)
Cobas b101								
Criel 2016, Belgium	Adams Arkray HA-8160 (ion-exchange chromatography)	venous???	Outpatient diabetes clinic	40 (40)	-2.2	-0.20	IFCC units: Low: 3.9, High: 2.0 NGSP units: Low: 2.4 High: 1.5	Failed to meet precision criterion at low HbA1c level
Kenealy 2019, New Zealand	Variant II Turbo, Cobas Integra 800	capillary	National diabetes and cardiovascular risk screening programme (at GP practices)	Study 1) 2 x 50 samples (1 laboratory x 2 batches) Study 2) 2 x 20 (2 practices) Study 3) batch 311021-01	In the range of 40–50 mmol/mol HbA1c: **Study 1:** Bias ranged from 4 and -9 mmol/mol (mean bias not reported but according to graphs was negative and <5mmol/mol; **Study 2:** approx. 1 mmol/mol (range -8 to 5) **Study 3:** sample EPOCH 1408, target 32 mmol/mol, mean 31 mmol/mol	In the range of 3.66% and 4.58% HbA1c level: **Study 1:** Bias ranged from approx. -0.82% to 0.37% (mean not reported but negative and <46%); **Study 2:** Mean bias 0.09% (range -0.73 to 0.46) **Study 3:** sample EPOCH 1408, target 2.93%, mean 2.84% (range 2.65% to 3.11%); sample EPOCH 1405, target	**IFCC units:** Study 1: 1.9% at 40 and at 51 mmol/mol Study 2: n/a Study 3: Low (mean 33 mmol/mol): 5.2%; High (mean 79 mmol/mol): 4.2%.	POC testing for HbA1c can be sufficiently accurate for screening and diagnosis of diabetes If testing is done within stringent quality assurance processes prior to and while in use (see also the EPOCH trial)

Study and country	Reference instrument (method)	POCT sample	Setting and POCT operator	N of samples (N of patients and diagnosis)	Mean bias (SD and/or LA) in mmol/mol	Mean bias (SD or LA) in % DCCT	Precision / CV (%)	Recommended for screening (other relevant conclusions)
					(range 29–34 mmol/mol); sample EPOCH 1405, target 76 mmol/mol, mean 75 mmol/mol (range 73–78 mmol/mol).	6.95%, mean 6.86% (range 6.68% to 7.14%).		
Lyon 2017, Canada (CA)	Cobas Tina-quant Gen. 3 on a c501 analyzer	Venous?	n/a	47 (47)	1.09 (95%CI 0.13 to 1.96)	0.10% (95% CI 0.012 to 0.179).	At HbA1c levels: 5.3%: 1.6%, 7.0%: 1.0%, 13.4%: 1.7%	excellent precision and accuracy relative to the Cobas Tina-quant Gen. 3 NGSP certified HbA1c method
Toro-Crespo 2017, Spain	Tosoh G8, SRL	venous	Laboratory	100 samples from the routine laboratory workload	1.08 (????)	-0.0985 (95%CI: -0.0171 to -0.0264)	Inter-assay: 1.92% Intra-assay: Low: 2.06% High: 1.87%	Considered suitable for the control, but not for the diagnosis of diabetes
Quo-Test								
Grant 2017, UK	BioRad D10 (HPLC)	Venous? (Whole blood EDTA samples)	n/a	100 (with and without diabetes)	1.4 (6.4)	0.13 (0.59)	Intra-assay: 3.5% (0.0–6.7%), Inter-assay: 2.7% (0.7–5.1%)	Similar performance to that of a laboratory HPLC; it should be considered for diagnostic

Study and country	Reference instrument (method)	POCT sample	Setting and POCT operator	N of samples (N of patients and diagnosis)	Mean bias (SD and/or LA) in mmol/mol	Mean bias (SD or LA) in % DCCT	Precision / CV (%)	Recommended for screening (other relevant conclusions)
								purposes (no funding from manufacturer reported)
Jones 2016, (CA) UK	BioRad D10 (HPLC)	Venous (Whole blood EDTA samples)	n/a	150 (two occasions 2 years apart batch 1, n = 50, batch 2, n = 100; subjects with and without diabetes)	n/a (batch 1: r^2 = 0.952, p < 0.0001; batch 2: r^2 = 0.969, p < 0.0001)	n/a	**Batch 1:** Intra-assay: 1.0-5.3%, Inter-assay 1.4%. **Batch 2:** Intra-assay 0.0-4.9%, Inter-assay 1.2%.	n/a (Equivalent performance to a laboratory based HPLC method over a 2 year period)
Stavelin 2019	Noklus EQA, criteria for bias: NGSP 0.3% (approx. 3 mmol/mol), for CV: 2% (approx. 3% mmol/mol)	capillary	n/a (but probably primary care)	13 POC sites	n/a	Fulfilled the bias recommendation of <0.3% HbA1c for each of the 14 samples.	9 out of 14 samples had CVs in the range of 2.1-5.7% (failing the <2% criterion)	Failed to meet the programme's criteria
B-analyst								
Hirst 2017	Mean bias range (n=4): 0.05% to 0.19% Pooled mean bias: 0.13% (95%CI 0.08 to 0.19) 95% Prediction interval: -0.10% to 0.37% Pooled SD of the mean bias (n=3): 0.22% (95% CI 0.19 to 0.24) 95% Prediction interval of the pooled SD: 0.08% to 0.35% Total CV (n=1): 1.9%							
Criel 2016,	Adams Arkray	venous???	Outpatient	40 (40)	Mean bias:	Mean bias:	IFCC units:	Met the criteria

Study and country	Reference instrument (method)	POCT sample	Setting and POCT operator	N of samples (N of patients and diagnosis)	Mean bias (SD and/or LA) in mmol/mol	Mean bias (SD or LA) in % DCCT	Precision / CV (%)	Recommended for screening (other relevant conclusions)
Belgium	HA-8160 (ion-exchange chromatography)		diabetes clinic		0.5 (0.12; 0.93) Total error: 3.02 (2.32; 3.73)	0.05 (0.01; 0.09) Total error: 0.28 (0.2; 0.34)	Low: 2.3, High: 2.9 NGSP units: Low: 1.1 High: 1.9	for precision but significant bias????
Toro-Crespo 2017, Spain	Tosoh G8, SRL	venous	Laboratory	100 samples from the routine laboratory workload	1.36 (95%CI 0.93 to 1.77)	0.124 (95%CI 0.0851 to 0.162).	Inter-assay: 1.34% Intra-assay: Low: 1.79% High: 3.17%	Considered suitable for the control, but not for the diagnosis of diabetes
PixoTest								
Chen 2018, Taiwan (CA)	Roche Cobas c111	n/a	n/a	60 (60, n/a)	n/a	353 of 360 (98%) values with bias within ±10%	n/a	n/a (comply to the criteria)
Cheng 2019, Taiwan	TOSOH G7	capillary (results for venous also reported)	3 clinical sites	120 (120, healthy subjects, outpatients and inpatients)	n/a	349 of 360 (96.9%) with bias within ±10%	n/a	n/a (similar accuracy to TOSOH G7 for management of DM)
NycoCard								
Zhou 2017, China	D10, Tosoh G8 and Premier Ultra 2	n/a	n/a	5 (5, n/a)	Range 1 to 12 mmol/mol pre-training; and 0 to 1 mmol/mol post-training	n/a	Reported in a diagram only	n/a
Razi 2016, Iran	D10, Cobas INTEGRA 400	n/a	n/a	154 (n/a, patients with DM)	-7.54 (6.23) vs D10; -8.74 (6.12)	-0.69 (0.57) vs D10; -0.80 (0.56) vs INGEGRA	At HbA1c levels: 5.5%: 3.12 7.5%: 2.32	n/a

Study and country	Reference instrument (method)	POCT sample	Setting and POCT operator	N of samples (N of patients and diagnosis)	Mean bias (SD and/or LA) in mmol/mol	Mean bias (SD or LA) in % DCCT	Precision / CV (%)	Recommended for screening (other relevant conclusions)
							9%: 2.62	
A1C EZ 2.0								
Zhou 2018, China	Tosoh G8	venous	n/a	40 (n/a, n/a)	2 (at 36 mmol/ mol); 3 (at 107 mmol/ mol); Relative bias at 48 mmol/mol was 0.8%	0.2 (at 5.4% HbA1c); 0.3% (at 11.9% HbA1c)	At HbA1c levels: 5.4%: 3.7 11.9%: 2.7	It has a reasonably high discriminative value for the diagnosis of diabetes
DS5								
Razi 2016, Iran	D10, Cobas INTEGRA 400	n/a	n/a	154 (n/a, patients with DM)		-0.87 (0.46) vs D10; -0.98 (0.52) vs INTEGRA	At HbA1c levels: 5.5%: 1.94 7.5%: 3.35 9%: 2.08	n/a
I-Chroma								
Vargas 2019, Ecuador (Argentina, USA)	BioRad Variant II Turbo (HPLC)	venous	laboratory	114 (24 T2DM, 90 no T2DM)	−5.5 (17.7)	−0.50 ± 1.62	Low: 1.65 High: 1.17% (in mmol/mol)	I-Chroma was precise but inaccurate.
Quo-Lab								
Lenters-Westra 2018 (the Netherlands & UK)	4 IFCC and/or NGSP certified secondary reference methods	n/a	laboratory	40 (n/a, n/a)	≤2 (at 48 and 75 mmol/mol)	≤0.2 (at 6.5% and 9.0%)	2.4 (at 46 and 71 mmol/mol), 1.6 (at 6.4%) and 1.8 (at 8.6%)	Excellent analytical performance (sigma = 4)
A1Care								
Lenters-Westra	4 IFCC and/or	n/a	laboratory	40 (n/a, n/a)	≤2 (at 48 and 75	≤0.2 (at 6.5%	6.2 (at 47	Sigma = 1.4;

Study and country	Reference instrument (method)	POCT sample	Setting and POCT operator	N of samples (N of patients and diagnosis)	Mean bias (SD and/or LA) in mmol/mol	Mean bias (SD or LA) in % DCCT	Precision / CV (%)	Recommended for screening (other relevant conclusions)
2018 (the Netherlands & UK)	NGSP certified secondary reference methods				mmol/mol)	and 9.0%)	mmol/mol), 4.1 (at 71 mmol/mol), 4.1 (at 6.4%) and 2.9 (at 8.7%)	failed all criteria mainly due to a high CV
HbA1c 501								
Lenters-Westra 2018 (the Netherlands & UK)	4 IFCC and/or NGSP certified secondary reference methods	n/a	laboratory	40 (n/a, n/a)	≤2 (at 48 and 75 mmol/mol) except for the second instrument at 75 mmol/mol which has bias = 2.4 mmol/mol	≤0.2 (at 6.5% and 9.0%) except for the second instrument at 9.0% which has bias = 0.2%	3.4 (at 46 mmol/mol), 2.7 (at 72 mmol/mol), 2.1 (at 6.3%) and 1.7 (at 8.7%)	Sigma = 2.1 acceptable performance
Allegro								
IFCC certification data	*Data provided by manufacturer* IFCC certification program, date of certification 01.06.2020 Total error = 2.2 mmol/mol; Bias = +0.4 mmol/mol; Imprecision = 1.8%; Grade = Bronze Bias at different levels of HbA1c: at 30 mmol/mol = -0.4 mmol/mol; at 50 mmol/mol = 0.4 mmol/mol; at 70 mmol/mol = 1.3 mmol/mol							
Finacare								
IFCC certification data	*Data provided by manufacturer* IFCC certification program, date of certification 01.01.2018 Total error = 2.3 mmol/mol; Bias = - 1.5 mmol/mol; Imprecision = 0.8%; Grade = Bronze IFCC certification program, date of certification 01.01.2020 Total error = 2.1 mmol/mol; Bias = - 0.6 mmol/mol; Imprecision = 1.5%; Grade = Silver							
NHCCLC, China	*Data provided by manufacturer*							

Study and country	Reference instrument (method)	POCT sample	Setting and POCT operator	N of samples (N of patients and diagnosis)	Mean bias (SD and/or LA) in mmol/mol	Mean bias (SD or LA) in % DCCT	Precision / CV (%)	Recommended for screening (other relevant conclusions)
	National Health Commission Clinical Laboratory Centre (NHCCLC) certification data, China, date of certification 17.01.2020 Bias (5 samples) at different HbA1c levels (%): at 5.49% = 0.18%, at 4.89% = -1.84%, at 9.71% = 0.93%, at 6.7% = 1.49%, at 7.66% = 0.52%; passed the criteria at all levels							
Other devices								
RC20	Device not available in Europe. Similar device, RC-W, with function to analyse DM and thalassemia, is scheduled to be commercially available after October 2020.							
In2it (BioRad)	Device discontinued							
*Diagnostic accuracy (2x2) data also reported CA conference abstract								

Table 13 Studies reporting direct comparison of two or more devices

Study	POC HbA1c devices	Results concerning comparative accuracy
Deobald 2016 (CA), USA	Afinion AS100, DCA Vantage	Both devices had acceptable precision and bias within 0.5% of the reference value, but Afinion showed less systematic bias: -0.06 (SD 0.17) vs -0.17 (SD 0.16) for DCA (p=0.0007)
Dubach 2019, Switzerland	Afinion AS100, DCA Vantage	Mean bias −0.27% and −0.21% for Afinion and DCA Vantage, respectively
Paknikar 2016, USA (EQA data)	Afinion AS100, DCA Vantage	Variable results over the 3 year period; DCA results were usually higher than the Afinion values except in the last 6 months
Saxton 2018, Peru	Afinion AS100, DCA Vantage	Afinion had CV 1.75% and mean bias 0.56% [6 mmol/mol]; DCA Vantage had CV 4.01% and mean bias 0.32% [4 mmol/mol]
Delatour 2019, The Netherlands (EQA data)	Afinion AS100, DCA Vantage	The overall performance of DCA Vantage met the 2σ criteria while that of Afinion met the 4σ criteria
Lenters-Westra 2017, The Netherlands (1) (EQA data)	Afinion AS100, DCA Vantage	**Afinion:** all 3 sites included in the study met the 2σ and 1 site met the 4σ criteria at TAE 10%; 2 sites met the 2σ at TAE 6%; **DCA Vantage:** 11 out of 12 sites met the 2σ and 5 the 4σ criteria at TAE 10%; 5 met the 2σ at TAE 6%;

Nordin 2018, Sweden (EQA data)	Afinion, DCA Vantage	**Afinion:** Mean bias ranged from -3.07 (0.21) to 0.12 (0.10) mmol/mol for the period 2005-17 and was negative except for 2015; it improved in the last 3 years, range: -0.18 to 0.13; the number of results meeting the criteria varied between approx. 60% and 97% (since 2008) **DCA Vantage:** Mean bias ranged from -1.44 (SD 0.07) to 2.07 (SD 0.06) mmol/mol in the period 2007-17 and was consistently negative until 2016 and 2017 [1.39 (0.06) and 2.07 (0.06) mmol/mol, respectively; the number of results meeting the criteria varied between approx. 80% and 93% (since 2008)
The EurA1c Trial Group, 2018 (EQA data)	Afinion, DCA Vantage	Both Afinion (76 devices) and DCA (158 devices) met the 2σ criterion (using fresh blood)
Stavelin 2019, Norway (EQA data)	Afinion AS100, DCA 2000/Vantage, Quo-Test	Afinion and DCA fulfilled the programme's criteria; Quo-Test meet the criterion for bias but not for imprecision
Springer 2016 (CA), USA	Afinion, DCA Vantage, A1CNow+	**Afinion:** mean bias -0.03 (SD 0.2) **DCA Vantage:** mean bias 0.128 (SD 0.2) **A1CNow+:** mean bias -0.03 (SD 0.2) Afinion and A1CNow+ showed better performance.
Torregrosa 2015, Spain	Afinion, DCA Vantage, In2It	**Afinion:** mean bias -0.04 (95%CI -0.10 to 0.02); CV 1.8% **DCA Vantage:** -0.28 (95% CI -0.34 to -0.23); CV 3.74% **In2it:** 0.06 (-0.14 to 0.26); CV 7.14%
Toro-Crespo 2017, Spain	Afinion, Cobas b101, B-analyst	**Afinion:** mean bias 0.0178 (95%CI: -0.0561, 0.0917); inter-assay CV 2.13%; intra-assay CV 1.13% (low), 1.97% (high) **Cobas b101:** -0.0985 (95%CI: -0.0171, -0.0264); inter-assay CV 1.92%; intra-assay CV 2.06% (low), 1.87% (high) **B-analyst:** 0.124 (95%CI:0.0851, 0.162); inter-assay CV 1.34%; intra-assay CV: 1.79% (low), 3.17% (high) All three recommended for monitoring but not diagnosis.
Criel 2016, Belgium	Afinion, Cobas b101, B-Analyst	**Afinion:** mean bias -2.2 mmol/mol; CV 2.3 – 3.7% **Cobas b101:** -2.2 mmol/mol; CV 2 - 3.9% **B-analyst:** 0.5 mmol/mol; CV 2.3 – 2.9% Only B-analyst met the quality specifications for precision.
Lenters-Westra 2018, The	Afinion2,	Sigma for Afinion 2 was 5.8 and for the Quo-Lab 4.0. Both POC devices

Netherlands	Quo-Lab, The A1Care, The HbA1c 501	passed the NGSP criteria with the 2 instruments used in the study. The HbA1c 501 had sigma = 2.1 and passed with 2 instruments except for one of the reference methods. The A1Care had a sigma of 1.4 and failed all criteria mainly due to high CV.
Szablowski 2018 (CA), USA	DCA Vantage, A1CNow+	**DCA Vantage:** mean bias -0.04%; Total CV 1.7%; clinical risk agreement; 100% **A1CNow+:** mean bias -0.04%; Total CV 3.3%; clinical risk agreement 94% The performance of both POC devices was similar to the laboratory methods
Vargas 2019, Ecuador	DCA Vantage, I-Chroma	**DCA Vantage:** -0.02% (SD 0.29); -0.2 (SD 3.2) mmol/mol **I-Chroma:** -0.50% (SD 1.62); -5.5 (SD 17.7) mmol/mol
Razi 2016, Iran	DS5, NycoCard	Both methods had negative bias within approximately -0.20% and CV 3.12%, 2.32% and 2.62% for NycoCard and 1.94%, 3.35% and 2.08% for DS5 at 5.5%, 7.5% and 9% HbA1c levels, respectively

Table 14 Studies investigating the benefits and disadvantages of HbA1c POCT included in the Horizon Scanning Report [3]

Device	**Excerpts from the report (**references are from the report and not from the bibliography included here)
Afinion vs DCA 2000+	"Operators experienced the device to be easy to use and identified, in comparison with the DCA 2000+, faster analysis time and easier sample loading as a considerable advantage of the Afinion (30)."
DCA 2000	"Instrument was used without difficulty by four different operators (50); DCA was simplest instrument to maintain (49)"

DCA Vantage	"Lot number-dependent performance (20); device is small and can be installed on a bench or on a table; user-friendly, with good ergonomics (23)"
A1cNow	"Lot-to-lot variation for the methods is a concern (32); simple to be operated by untrained patient users who can obtain performances equivalent to that obtained by trained medical professional users (43)
A1cNow+	"Device is wearable and can be used anywhere (14); provides a significant cost advantage to a patient who is responsible for fee-for-service and to primary care clinics that use the device for haemoglobin A1c determination (11); accessible, accurate and easy to use (10); A1CNow+ is a simple, portable, handheld device that is Clinical Laboratory Improvement Amendments waived, requires no calibration, and reagents need no refrigeration if used within 4 months (13)"
Cobas b101	**"**It is recommended not to use the Cobas b101 analyser in regions where the prevalence of Hb AE variants is high, due to possible interference (16)."
A1cGear	"The presence of the S haemoglobin trait in some of the analysed samples did not appear to negatively impact the measurements (18)."
Nycocard	"A nurse-based evaluation comparing performance of the analyser when handled by laboratory trained vs non-laboratory trained professionals reports on frustration felt by the staff due to several manual steps and the need of constant attention, as well as several error messages which lead to erroneous data (35)."
InnovaStar	"Needs to be calibrated and certified with fresh patient samples instead of frozen material (16); users were satisfied with the user manual; InnovaStar HbA1c instrument requires users with laboratory experience (27)"

Table 15 New studies investigating the benefits and disadvantages of HbA1c POCT

Study	Study design and POCT	Summary of methods	Summary of relevant results
Alzubaidi	Observational	Aimed to develop community-based screening model for	115 participants were screened; 92.3% of the screening

2019 United Arab Emirates [47]	study; Roche Cobas b101	diabetes and CVD. Pharmacists from 12 community pharmacies screened participants ≥40 years of age who had not been diagnosed with diabetes or CVD. Both cholesterol and HbA1c were measured using the same device.	was completed in a single visit. Mean duration was 27 min. At-risk individuals (57.4%) were referred to their physicians for further testing, while 94.5% of participants were at least satisfied with their screening experience.
Barron 2020 UK [44]	Retrospective observational study; DCA Vantage, Afinion, A1C Now+	Retrospective analysis of prospectively collected data from the NHS DPP from individuals referred to the programme between June 2016 and October 2018 with an HbA1c result indicating NDH. The study investigated the mean difference in HbA1c values from the referral (mainly laboratory) and initial assessment (POCT), the SD of the POCT HbA1c, the association of various factors with the mean HbA1c difference and the association between the POCT values and subsequent attendance on the programme.	Data from 73,703 participants were analysed. The mean (SD) number of days between HbA1c measurements was 203 (120) days. There was a significant mean difference between referral (mainly laboratory) and POCT HbA1c of -2.48mmol/mol (-0.23%) ($p<0.001$) with significant differences in mean HbA1c between devices ($p=<0.001$). The SD of POCT HbA1c was 4.46mmol/mol (0.41%) with significant differences in SDs between devices ($p<0.001$). Measurements in participants who were older, from more deprived areas and from Asian, black and mixed ethnic groups were associated with smaller HbA1c differences. In participants who had sufficient time to attend an intervention session (n=46,894), normoglycaemic POCT HbA1c vs. NDH POCT HbA1c values were associated with lower subsequent attendance at behavioural interventions (58% vs. 67%,$p<0.001$).
Bossart 2016 USA [48]	Observational study; POCT not specified	Aimed to assess diabetes screening by a dental hygienist in patients with periodontitis, without diagnosis of diabetes, using a diabetes risk questionnaire, periodontal findings and POCT HbA1c analyser.	32% (n = 16) presented HbA1c values indicating prediabetes and one patient DM2. Direct cost for each HbA1c was $9US. Mean screening time including patient education was 14 min (SD = 6.2); 53% (n = 9 of 17) of participants with elevated HbA1c contacted their primary healthcare provider within 2 weeks as recommended.
Deobald 2016 USA [18]	Observational study; DCA Vantage, Afinion AS100	Compared assay performance and ease of use of two POCT devices; the method of the latter was not specified.	Compared to DCA, Afinion had a smaller footprint, no required maintenance, and a faster turnaround time of 3 minutes compared to 6 minutes on DCA
Karmali 2017	Observational	This study tested the preliminary feasibility, acceptability	18 patients were recruited. After the intervention, 83% of

USA [49]	(pre-post) study; DCA Vantage	and efficacy of POCT and quantitative CVD risk assessment in high-risk adults to increase guideline-recommended statin use in primary prevention.	participants discussed CVD risk with their PCP, 47% received a statin recommendation, and 29% received a new statin prescription during the PCP visit. Participants reported high levels of satisfaction with the intervention.
Lewandrowski 2017 USA [51]	Observational (retrospective comparison) study; Afinion	Investigated the impact of implementing POCT on practice efficiency in an academic primary care practice.	In the patient cohort that received POCT there was a 99% reduction in letters to patients ($p < 0.001$), a 75% decrease in calls to patients (not significant due to small numbers), a 50% reduction in follow-up tests per visit ($p = 0.044$) and a 38% reduction in follow-up visits due to abnormal test results ($p = 0.178$). Financial analysis including testing costs, revenues and efficiency gains to the practice demonstrated a net financial benefit of $11.90–14.74 per patient visit
Orvim Solvik 2018 Norway [28]	Analysis of data from the Noklus EQA; DCA Vantage	The aim of the study was to describe the implementation of quality control of the HbA1c POCT instruments and investigate the performance in IQC and EQA for HbA1c POCT in Norwegian community pharmacies	Community pharmacies were able to implement procedures to carry out IQC and EQA on HbA1c POCT instruments, and the performance was comparable with that of GP offices. The compliance in the EQA surveys was modest, which the authors explained with the short duration of the study (making it difficult to implement all procedures)
Shephard 2017 Australia [24]	Analysis of data from the QAAMS EQA; DCA Vantage	The paper reports on the past 15 years of quality testing in QAAMS and examines the performance of HbA1c POCT	The QAAMS programme provides training, technical support, a quality assurance programme and a consultation programme to support POCT at participating Aboriginal and Torres Strait Islander health care sites. The QAAMS quality management framework includes monthly testing of quality control and EQA samples, with specific procedures in place to increase the effectiveness of the process. Key performance indicators of quality include imprecision and percentage acceptable results.
Stavelin 2019 Norway [29]	Analysis of data from the Noklus EQA;	In addition to evaluation of the analytical performance of POCT, the study investigated which of the following variables predict good participant performance: instrument and reagent lot numbers, the profession of the operator,	Reagent lot was the only independent factor to predict good participant performance.

	Afinion	the number of patient samples performed per week and the frequency of running IQC.	
Zhou 2017 China [43]	Pre-post study; Afinion, NycoCard and DCA Vantage	The study assessed the usefulness of commutable secondary reference materials with IFCC-assigned values for improving the accuracy of HbA1c determinations and the impact of training of operators on improving precision, especially for semi-automated POCT devices.	Inter-laboratory CV was reduced significantly after standardized on-site training of operators. CVs in the NycoCard group decreased from 12% to 4%; the reduction was less pronounced in the DCA Vantage and Afinion groups, suggesting that training is more effective with semi-automated devices.
Wells 2017 New Zealand [14]	The EPOCH trial (Evaluating a POC device in Heart healthcare): Pragmatic, cluster RCT; Cobas b101	It assessed the effect of POCT for lipids and HbA1c in addition to testing by community laboratory (usual practice) on the completion of CVD risk assessments in general practice. 20 GP practices stratified by size and rurality were randomised to POC device plus usual practice or usual practice alone. Patients aged 35–79 years were eligible if they met national guideline criteria for CVD risk assessment. The primary outcome was the proportion of completed CVD risk assessments. The emphasis was on undertaking a CVD risk assessment for all patients in the practice population within five years, rather than ensuring population groups with high CVD risk had the recommended frequency of monitoring.	A CVD risk was recorded for 7421 patients in 10 POCT practices and 6217 patients in 10 control practices; 99.5% of CVD risk assessments had complete data in both. There was an interruption in the trial due to changes in HbA1c test performance related to a technical fault; 90% of the eligible patients had already had a CVD risk assessment by the time the POC device was re-introduced. Having a POCT device within the practices made no difference to the completion of CVD risk assessments, incomplete assessments and time [in days] to complete the assessment. Performance incentives and external influences were more powerful modifiers of practice behaviours than the POCT device which was viewed by most as an additional tool rather than as an opportunity to review practice work flow and leverage the immediate test results for patient education and CVD risk management discussions. Practices that made systemic changes to incorporate POCT found it useful. POCT was viewed as more suitable for monitoring than screening, but most participants stated they would not be using it if they had to pay for consumables.
Whitley 2017 USA [50]	Observational (prospective longitudinal) study;	The study compared diabetes screenings between standard practices vs systematically offered POC HbA1c tests in patients aged 45 years or older	Systematic screening (n = 164) identified 63% (n = 104) with unknown hyperglycemia and 53% (n = 88) in prediabetes. Standard practice (n = 324) screened 22% (n = 73), usually by blood glucose (96%); 8% (n = 6) and 33% (n = 24) were

	POCT not specified		found to have diabetes and prediabetes, respectively. The association between screening outcome and screening method was statistically significant (P = 0.005)

Table 16 Guidelines relevant to the use of HbA1c POCT for diagnosis of NDH and DM

Guideline title	Authors	Country	Year	Recommendations
International Expert Committee Report on the Role of the A1C Assay in the Diagnosis of Diabetes 2009 [52]	ADA	International	2009	POCT for HbA1c "have not yet been shown to be sufficiently accurate or precise for diagnosing diabetes" HbA1C test for diagnosis should be performed using laboratory equipment
International Diabetes Federation IDF Clinical Practice Recommendations for managing Type 2 Diabetes in Primary Care 2017 [53]	IDF	International	2017	No recommendations for HbA1c POCT - HbA1c test should be performed in a laboratory using a NGSP-certified method standardized to the DCCT assay
Type 2 diabetes: prevention in people at high risk [54]	NICE	UK	2017 updated	HbA1c tests – which include POCTs - should "conform to expert consensus reports on appropriate use and national quality specifications" and only be carried out by trained staff
NHS Health Check Best practice guidance [55]	PHE	UK	2019	Recommendations for POCT from MHRA (not specifically HbA1c): Local pathology laboratory is involved, POCT co-ordinator identified; only staff trained and competent should use POCT; instructions must be read; awareness of situations where device should not be used; develop standard operating procedures; quality control is implemented and recorded; consider which staff review results and note this on patient records; record test results, strip lot number and operator identify; essential to maintain device according to manufacturer's guidance
The use of POCT HbA1c devices in	NHS DPP	UK	2019	Recommendations for POCT from MHRA and related HbA1c guidelines.

Guideline title	Authors	Country	Year	Recommendations
the NHS Diabetes Prevention Programme: Recommendations from an expert working group commissioned by NHS England				More specifically, it recommends that: 1) POCT HbA1c should only be considered where there is evidence for cost effective implementation; 2) Procurement of HbA1c POCT devices should only be considered in collaboration with the local UKAS accredited pathology laboratory with the involvement of a local POCT committee; 3) A service level agreement with the local laboratory should be made to ensure adequate support at every stage during device selection, procurement, evaluation, implementation and thereafter to help initiate, advise on and maintain the quality framework; 4) A designated member of a POCT committee (POCT coordinator) should act as a liaison between the laboratory and the user; 5) Advice on which device to purchase should be obtained with local laboratory support and expertise, relating to the minimum analytical performance criteria, published studies on device performance, local laboratory experience and external quality assessment data; 6) All parameters of the MHRA stipulated quality framework must be in place prior to implementation of the POCT HbA1c device; 7) All POCT HbA1c devices must have a clear process for internal quality control and be enrolled into an external quality assessment programme. There are also pre- and post-analytic considerations In addition, of which those relating directly to POCT HbA1c are as follows: 1) An asymptomatic patient with a value >48 mmol/mol should have a confirmatory HbA1c test to confirm a diagnosis of diabetes, using a methodology validated for diagnosis of type 2 diabetes; 2) Individuals identified on the basis of a laboratory HbA1c result should have a baseline POCT HbA1c, if POCT is to be used to monitor progress during the DPP; 3) Once enrolled in the NHS DPP testing of HbA1c should be by the same modality that is if enrolled on basis of a POCT result, all subsequent testing

Guideline title	Authors	Country	Year	Recommendations
				should be via the same POCT methodology.
National Evidence Based Guideline for Case Detection and Diagnosis of Type 2 Diabetes 2009 [56]	Diabetes Australia Guideline Development Consortium	Australia	2009	No recommendation for HbA1c POCT. HbA1c is an option for screening undiagnosed type 2 diabetes
Diabetes Canada Clinical Practice Guidelines Expert Committee 2018 Clinical Practice Guidelines [57]	Diabetes Canada	Canada	2018	Advantages of HbA1c POCT acknowledged: Rapid test results to expedite medical decision-making, convenience for people with diabetes, potential improved health system efficiency and improved access to testing for underserved populations. However, no POCT HbA1C analyzers approved in Canada for the diagnosis of diabetes due to lack of evidence on: impact on medication use, clinical decision-making and participants' outcomes, and economic evaluation Only where access to standard laboratory testing for remote Indigenous populations is not available, HbA1C POCT may be considered "where testing is associated with a quality control program; and interpretation and follow-up expertise is available" HbA1C must be "measured using a validated assay standardized to the NGSP-DCCT reference"
Guidance on the management of Type 2 diabetes 2011 [58]	Ministry of Health	New Zealand	2011	POCT HbA1c are not "sufficiently accurate" for diagnosis. HbA1c should be measured in an accredited laboratory
Classification and Diagnosis of Diabetes: Standards of Medical Care in Diabetes [59]	ADA	USA	2020	"Although point-of-care A1C assays may be NGSP certified or U.S. Food and Drug Administration approved for diagnosis, proficiency testing is not always mandated for performing the test. Therefore, point-of-care assays

Guideline title	Authors	Country	Year	Recommendations
				approved for diagnostic purposes should only be considered in settings licensed to perform moderate-to-high complexity tests.”

References

1. *NHS Health Check*. [cited 2020 3 April]; Available from: https://www.nhs.uk/conditions/nhs-health-check/.
2. Hirst, J.A., et al., *Performance of point-of-care HbA<inf>1c</inf> test devices: Implications for use in clinical practice - A systematic review and meta-analysis.* Clinical Chemistry and Laboratory Medicine, 2017. **55**(2): p. 167-180.
3. Schaffert, L.-N., et al., *Point-of-care HbA1c tests - diagnosis of diabetes*. Horizon Scan Report 0044. 2016: NIHR: Diagnostic Evidence Co-operative.
4. Whiting, P.F., et al., *QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies.* Ann Intern Med, 2011. **155**(8): p. 529-36.
5. Shea, B.J., et al., *AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both.* BMJ, 2017. **358**: p. j4008.
6. *Assessing risk of bias in a randomized trial*, in *Cochrane Handbook for Systematic Reviews of Interventions*. p. 205-228.
7. Brouwers, M.C., et al., *AGREE II: advancing guideline development, reporting and evaluation in health care.* Cmaj, 2010. **182**(18): p. E839-42.
8. Peñaloza Ramos, M.C., et al., *A Systematic Review of Research Guidelines in Decision-Analytic Modeling.* Value in Health, 2015. **18**(4): p. 512-529.
9. Torregrosa, M.E., et al., *Accuracy of three hemoglobin A1c point-of-care systems for glucose monitoring in patients with diabetes mellitus.* Endocrinol Nutr, 2015. **62**(10): p. 478-84.
10. Lenters-Westra, E. and E. English, *Evaluation of Four HbA1c Point-of-Care Devices Using International Quality Targets: Are They Fit for the Purpose?* Journal of Diabetes Science and Technology, 2018. **12**(4): p. 762-770.
11. Group, T.E.T., *EurA1c: The European HbA1c Trial to Investigate the Performance of HbA1c Assays in 2166 Laboratories across 17 Countries and 24 Manufacturers by Use of the IFCC Model for Quality Targets.* Clin Chem, 2018. **64**(8): p. 1183-1192.
12. Arnold, W.D., et al., *Fingerstick Precision and Total Error of a Point-of-Care HbA1c Test.* Journal of Diabetes Science and Technology., 2019.
13. Arnold, W.D., et al., *Accuracy and Precision of a Point-of-Care HbA1c Test.* Journal of Diabetes Science and Technology., 2019.
14. Wells, S., et al., *The impact of a point-of-care testing device on CVD risk assessment completion in New Zealand primary-care practice: A cluster randomised controlled trial and qualitative investigation.* PLoS ONE, 2017. **12 (4) (no pagination)**(e0174504).
15. Lynn, L., et al., *Accuracy and the Proper Use of the Point-of-Care Analysis in Hemoglobin A1C Measurement.* Metabolic Syndrome and Related Disorders, 2018. **16**(3): p. 117-121.
16. Saxton, A.T., et al., *Assessment of two diabetes point-of-care analyzers measuring hemoglobin A1c in the peruvian amazon.* Annals of Global Health, 2018. **84**(4): p. 618-624.
17. Diaz-Garzon, J., et al., *Assessment of the reliability of poct vs HPLC methods for HbA1c measurement by analytical performance monitoring.* Clinical Chemistry and Laboratory Medicine, 2017. **55 (Supplement 1)**: p. S581.
18. Deobald, G.R., et al., *Comparison of siemens DCA vantage and alere afi nion point of care analyzers for hemoglobin A1c determination.* Clinical Chemistry, 2016. **62 (10 Supplement 1)**: p. S228.
19. Szablowski, C.J., et al., *Point-of-care HbA1c-a case for diabetes screening and diagnosis.* Diabetes, 2018. **67 (Supplement 1)**: p. A406.
20. Vargas, M.G., et al., *Assessment of two glycated hemoglobin immunoassays.* Endocrinologia, Diabetes y Nutricion., 2019.
21. Valdez-Gonzalez, L.A., et al., *Agreement between the 'point of care' tests for microalbuminuria and HbA1c performed in mexican family medicine units and the results of*

standard laboratory tests. Scandinavian Journal of Clinical and Laboratory Investigation, 2018. **78**(1-2): p. 87-93.

22. Mackenzie-Feder, J., et al., *Agreement of Point-of-Care Capillary Glycated Hemoglobin Levels with Conventional Screening Tests for Diabetes Mellitus in a Canadian First Nations Population.* Canadian Journal of Diabetes, 2016. **40**(3): p. 242-246.
23. Lenters-Westra, E. and E. English, *Understanding the Use of Sigma Metrics in Hemoglobin A1c Analysis.* Clinics in Laboratory Medicine, 2017. **37**(1): p. 57-71.
24. Shephard, M., et al., *Results from 15 years of quality surveillance for a National Indigenous Point-of-Care Testing Program for diabetes.* Clinical Biochemistry, 2017. **50**(18): p. 1159-1163.
25. Albeiroti, S., et al., *Long-term performance of the DCA vantage analyzer for HbA1C measurement and initiating HbA1C profciency testing in the ambulatory setting.* American Journal of Clinical Pathology, 2018. **149 (Supplement 1)**: p. S201.
26. Nordin, G., *Accuracy of HbA1c as Monitored by External Quality Assessment and Compared With Patient Mean Values.* Journal of Diabetes Science & Technology, 2018. **12**(4): p. 771-779.
27. Delatour, V., et al., *Trueness assessment of HbA1c routine assays: are processed EQA materials up to the job?* Clin Chem Lab Med, 2019. **57**(10): p. 1623-1631.
28. Orvim Solvik, U., et al., *Quality Control of Norwegian Pharmacy HbA1c Testing: A Modest Beginning.* Journal of Diabetes Science & Technology, 2018. **12**(4): p. 753-761.
29. Stavelin, A., et al., *Performance of Afinion HbA1c measurements in general practice as judged by external quality assurance data.* Clinical chemistry and laboratory medicine., 2019. **11**.
30. Paknikar, S., et al., *Long-Term Performance of Point-of-Care Hemoglobin A1c Assays.* Journal of Diabetes Science & Technology, 2016. **10**(6): p. 1308-1315.
31. Schimenes, T.V., et al., *Diagnosing and monitoring diabetic patients with a Siemens Healthineers end-to-end solution for HbA1c testing*.* Clinica Chimica Acta, 2019. **493 (Supplement 1)**: p. S269-S270.
32. Abbai, N.S., et al., *Good correlation between the Afinion AS100 analyser and the ABX Pentra 400 analyser for the measurement of glycosylated haemoglobin and lipid levels in older adults in Durban, South Africa.* South African medical journal = Suid-Afrikaanse tydskrif vir geneeskunde, 2017. **108**(1): p. 50-55.
33. Jain, A., et al., *Evaluation of the point of care Afinion AS100 analyser in a community setting.* Annals of Clinical Biochemistry, 2017. **54**(3): p. 331-341.
34. Nathan, D.M., et al., *Accuracy of a Point-of-Care Hemoglobin A1c Assay.* Journal of Diabetes Science and Technology., 2019.
35. Toro-Crespo, M., et al., *Evaluation of 3 Hemoglobin A1c Point of Care Instruments. Point of Care Testing for HbA1c: Evaluation of Cobas b101, B-Analyst and AfinionTM.* Clinical Laboratory, 2017. **63**(7): p. 1107-1112.
36. Grant, D.A., et al., *Comparison of a point-of-care analyser for the determination of HbA1c with HPLC method.* Practical Laboratory Medicine, 2017. **8**: p. 26-29.
37. Jones, D.A., et al., *Comparison of a point-of-care analyser for the determination of HbA1c against a laboratory based analyser.* Diabetic Medicine, 2016. **1)**: p. 66-67.
38. Giavarina, D., *Understanding Bland Altman analysis.* Biochemia medica, 2015. **25**(2): p. 141-151.
39. Criel, M., S. Jonckheere, and M. Langlois, *Evaluation of three hemoglobin A1C point-of-care instruments.* Clinical Laboratory, 2016. **62**(3): p. 285-291.
40. Kenealy, T., et al., *HbA<inf>1c</inf> screening in the community: Lessons for safety and quality management of a point of care programme.* Primary Care Diabetes, 2019. **13**(2): p. 170-175.

41. Lyon, M.E. and A.W. Lyon, *Analytic evaluation of the b101 hemoglobin A1c point of care testing method from roche diagnostics.* Clinical Chemistry, 2017. **63 (Supplement 1)**: p. S267-S268.
42. Zhou, R., et al., *Evaluation of a new hemoglobin A1c analyzer for point-of-care testing.* Journal of Clinical Laboratory Analysis, 2018. **32 (1) (no pagination)**(e22172).
43. Zhou, R., et al., *Improvement in the Quality of HbA1c Determination by Using Commutable Specimens With IFCC-Assigned Values.* Laboratory Medicine, 2017. **48**(2): p. 148-153.
44. Barron E, M.S., English E, John W G, Sampson M, Bachmann M, Barth J H, Oliver N, Alberti G, Bakhai C, O'Neill S, Young B, Wareham N, Khunti K, Jebb S, Smith S, Valabhji J, *Use of Point-of-Care HbA1c testing in the English National Health Service Diabetes Prevention Programme*. 2020.
45. Sampson, M., et al., *Discordance in glycemic categories and regression to normality at baseline in 10,000 people in a Type 2 diabetes prevention trial.* Scientific reports, 2018. **8**(1): p. 6240-6240.
46. El-Osta, A., et al., *Does use of point-of-care testing improve cost-effectiveness of the NHS Health Check programme in the primary care setting? A cost-minimisation analysis.* BMJ Open, 2017. **7 (8) (no pagination)**(e015494).
47. Alzubaidi, H.T., et al., *Diabetes and cardiovascular disease risk screening model in community pharmacies in a developing primary healthcare system: A feasibility study.* BMJ Open, 2019. **9 (11) (no pagination)**(e031246).
48. Bossart, M., et al., *A pilot study of an HbA1c chairside screening protocol for diabetes in patients with chronic periodontitis: the dental hygienist's role.* International journal of dental hygiene, 2016. **14**(2): p. 98-107.
49. Karmali, K.N., et al., *Point-of-care testing to promote cardiovascular disease risk assessment: A proof of concept study.* Preventive Medicine Reports, 2017. **7**: p. 136-139.
50. Whitley, H.P., C. Hanson, and J.M. Parton, *Systematic Diabetes Screening Using Point-of-Care HbA1c Testing Facilitates Identification of Prediabetes.* Annals of family medicine, 2017. **15**(2): p. 162-164.
51. Lewandrowski, E.L., et al., *Implementation of point-of-care testing in a general internal medicine practice: A confirmation study.* Clinica Chimica Acta, 2017. **473**: p. 71-74.
52. *International Expert Committee Report on the Role of the A1C Assay in the Diagnosis of Diabetes.* Diabetes Care, 2009. **32**(7): p. 1327-1334.
53. IFD. *IDF Clinical Practice Recommendations for Managing Type 2 Diabetes in Primary Care*. 2017 [cited 2020 3 April]; Available from: https://www.idf.org/e-library/guidelines/128-idf-clinical-practice-recommendations-for-managing-type-2-diabetes-in-primary-care.html.
54. NICE, *Type 2 diabetes: prevention in people at high risk. Public health guideline [PH38]*, N.I.f.H.a.C. Excellence, Editor. 2017.
55. England, P.H., *NHS Health Check. Best practice guidance. For commissioners and providers*, P.H. England, Editor. 2019.
56. Colagiuri S, D.D., Girgis S, Colagiuri R, *National Evidence Based Guideline for Case Detection and Diagnosis of Type 2 Diabetes*, D.A.a.t. NHMRC, Editor. 2009: Canberra.
57. Committee, D.C.C.P.G.E., *Diabetes Canada 2018 Clinical Practice Guidelines for the Prevention and Management of Diabetes in Canada.* Can J Diabetes, 2018. **42**(Suppl 1): p. S1-S325.
58. Group, N.Z.G., *Guidance on the Management of type 2 diabetes 2011*. 2011, New Zealand Guidelines Group: Wellington.
59. *2. Classification and Diagnosis of Diabetes: Standards of Medical Care in Diabetes—2020.* Diabetes Care, 2020. **43**(Supplement 1): p. S14-S31.
60. DPP, *The use of POCT HbA1c devices in the NHS Diabetes Prevention Programme: Recommendations from an expert working group commissioned by NHS England*, N.D.P. Programme, Editor. 2019, NHS.

61. Lenters-Westra, E. and E. English, *Analysis: Investigating the quality of POCT devices for HbA1c, what are our next steps?* Journal of Diabetes Science and Technology, 2019. **13**(6): p. 1154-1157.
62. Gomez-Peralta, F., et al., *Point-of-care capillary HbA1c measurement in the emergency department: a useful tool to detect unrecognized and uncontrolled diabetes.* International Journal of Emergency Medicine, 2016. **9**(1): p. 1-6.
63. Moskowitz, K.A., et al., *Multicenter accuracy assessment of A1c now+: A disposable system for monitoring hemoglobin A1c.* Diabetes, 2017. **66 (Supplement 1)**: p. A231.
64. Chen, C.H. and S.Y. Cheng, *The performance of a novel smartphonebase HBA1C measurement.* Diabetes Technology and Therapeutics, 2018. **20 (Supplement 1)**: p. A11.
65. Cheng, S.Y., et al., *Mobile Health for Diagnosis-PixoTest HbA1c POCT System.* Journal of Medical and Biological Engineering., 2019.
66. Dubach, I.L., E.R. Christ, and P. Diem, *HbA<inf>1c</inf>-testing: Evaluation of two point-of-care analysers.* Primary Care Diabetes, 2019. **13**(6): p. 583-587.
67. Fellows, S.E. and G.C. Cipriano, *Variation between Point-of-care and Laboratory HbA1c testing in Clinical Practice.* Puerto Rico health sciences journal, 2019. **38**(3): p. 189-191.
68. Hamada, E. and M. Maekawa, *Basic evaluation of a novel glycohemoglobin analyzer RC20 for POCT.* Clinical Chemistry, 2016. **62 (10 Supplement 1)**: p. S231.
69. Manthei, D.M., et al., *Retrospective accuracy study of point-of-care hemoglobin A<inf>1c</inf> in field conditions.* Clinical Chemistry, 2017. **63**(3): p. 780-781.
70. Razi, F., et al., *Comparative analytical performance of various HbA1c assays in Iran.* Archives of Iranian Medicine, 2016. **19**(6): p. 414-419.
71. Sobolesky, P.M., et al., *Multicenter assessment of a hemoglobin A1c point-of-care device for diagnosis of diabetes mellitus.* Clinical Biochemistry, 2018. **61**: p. 18-22.
72. Springer, L.B., *Performance evaluation of three NGSP-certified point-of-care HbA1c devices.* Annals of Clinical and Laboratory Science, 2016. **46 (3)**: p. 323.
73. Swensen, M.H., et al., *Evaluation of the Alere Afinion HBa1c DX test: Can point of care HBa1c testing deliver the performance needed for diagnostic use?* Diabetes, 2016. **65 (Supplement 1)**: p. A380-A381.

Appendix 1 Search strategies

Database: Embase <1974 to 2020 January 09>

1 HbA1c.mp. (70607)
2 Haemoglobin A1c.mp. (1970)
3 hemoglobin A1c.mp. (106451)
4 glycated haemoglobin.mp. (4751)
5 glycated hemoglobin.mp. (10881)
6 glycosylated hemoglobin.mp. (27841)
7 glycosylated haemoglobin.mp. (3255)
8 Hb A1c.mp. (757)
9 A1c.mp. (112474)
10 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 (136909)
11 ((immediate* or rapid* or same time or same visit or near patient or instant* or portable or bedside or bed-side or handheld or hand-held) adj3 (test* or turnaround or analys* or analyz* or measure* or assay* or monitor*)).tw. (102078)
12 (poc or poct or "point of care").tw. (29297)
13 Point-of-Care Systems/ (1842)
14 A1cNow*.mp. (96)
15 Afinion.mp. (87)
16 NycoCard.mp. (146)
17 DCA Vantage.mp. (111)

18 Cobas b.mp. (40)
19 Allegro analyser.mp. (0)
20 suresign finecare.mp. (1)
21 Eurolyser.mp. (10)
22 in2it.mp. (28)
23 quo test.mp. (16)
24 or/11-23 (128897)
25 10 and 24 (1268)
26 limit 25 to yr="2016 -Current" (493)

****

Database: Cochrane Library
Date Run: 10/01/2020 15:16:29

#1 HbA1c:ti,ab 15892
#2 Haemoglobin A1c:ti,ab 10162
#3 hemoglobin A1c:ti,ab 10162
#4 glycated haemoglobin:ti,ab 4846
#5 glycated hemoglobin:ti,ab 4846
#6 glycosylated hemoglobin:ti,ab 3111
#7 glycosylated haemoglobin:ti,ab 3110
#8 Hb A1c:ti,ab 16084
#9 A1c:ti,ab 19341
#10 #1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 22258
#11 ((immediate* or rapid* or "same time" or "same visit" or "near patient" or instant* or portable or bedside or "bed side" or handheld or "hand held") near/3 (test* or turnaround or analys* or analyz* or measure* or assay* or monitor*)):ti,ab 8378
#12 (poc or poct or "point of care"):ti,ab 3032
#13 MeSH descriptor: [Point-of-Care Systems] explode all trees 472
#14 A1cNow*:ti,ab 4
#15 Afinion:ti,ab 5
#16 NycoCard:ti,ab 8
#17 "DCA Vantage":ti,ab 6
#18 "Cobas b":ti,ab 1
#19 "Allegro analyser":ti,ab 2
#20 "suresign finecare":ti,ab 0
#21 Eurolyser:ti,ab 0
#22 in2it:ti,ab 0
#23 "quo test":ti,ab 1
#24 #11 or #12 or #13 or #14 or #15 or #16 or #17 or #18 or #19 or #20 or #21 or #22 or #23 11366
#25 #10 and #24 with Publication Year from 2016 to 2020, with Cochrane Library publication date Between Jan 2016 and Jan 2020, in Trials 76
#10 and #24 with Cochrane Library publication date Between Jan 2016 and Jan 2020, in Cochrane Reviews 6

Appendix 2: Search log

Database	Date searched	Number of hits
Embase 1974 to 2020 January 09 via OvidSp	10/01/20	493
Ovid MEDLINE(R) ALL 1946 to January 09, 2020	10/01/20	228
Cochrane Library (CDSR)	10/01/20	6

Cochrane Library (CENTRAL)	10/01/20	76

Total records = 797
Duplicates = 223
For Ti/Ab Screening = 574