Performance of point-of-care HbA\textsubscript{1c} test devices: implications for use in clinical practice – a systematic review and meta-analysis


Abstract

Background: Point-of-care (POC) devices could be used to measure hemoglobin A\textsubscript{1c} (HbA\textsubscript{1c}) in the doctors’ office, allowing immediate feedback of results to patients. Reports have raised concerns about the analytical performance of some of these devices. We carried out a systematic review and meta-analysis using a novel approach to compare the accuracy and precision of POC HbA\textsubscript{1c} devices.

Methods: Medline, Embase and Web of Science databases were searched in June 2015 for published reports comparing POC HbA\textsubscript{1c} devices with laboratory methods. Two reviewers screened articles and extracted data on bias, precision and diagnostic accuracy. Mean bias and variability between the POC and laboratory test were combined in a meta-analysis. Study quality was assessed using the QUADAS2 tool.

Results: Two researchers independently reviewed 1739 records for eligibility. Sixty-one studies were included in the meta-analysis of mean bias. Devices evaluated were A1cgear, A1cNow, Afinion, B-analyst, Clover, Cobas b101, DCA 2000/Vantage, HemoCue, Innovastar, Nycocard, Quo-Lab, Quo-Test and SDA1cCare. Nine devices had a negative mean bias which was significant for three devices. There was substantial variability in bias within devices. There was no difference in bias between clinical or laboratory operators in two devices.

Conclusions: This is the first meta-analysis to directly compare performance of POC HbA\textsubscript{1c} devices. Use of a device with a mean negative bias compared to a laboratory method may lead to higher levels of glycemia and a lower risk of hypoglycaemia. The implications of this on clinical decision-making and patient outcomes now need to be tested in a randomized trial.

Keywords: diabetes; HbA\textsubscript{1c}; instrument performance; meta-analysis; point-of-care testing; systematic review.

Introduction

Regular monitoring of glycated hemoglobin subfraction A\textsubscript{1c} (HbA\textsubscript{1c}) in people with diabetes and treatment with glucose-lowering medications to improve glycaemic control can reduce the risk of developing complications [1]. In 2011, a World Health Organization consultation concluded that HbA\textsubscript{1c} at a threshold of 6.5% (48 mmol/mol) can be used as a diagnostic test for diabetes [2]. HbA\textsubscript{1c} monitoring often requires the patient to attend the health center twice: once to have blood taken and then returning to get test results and receive adjustments to medication.

Point-of-care (POC) analysers are bench-top instruments that use a finger-prick blood sample and are designed for use in a treatment room or at the bed-side. They provide a test result within a few minutes allowing clinical decisions and medication changes to take place immediately. The suitability of many of these devices for the accurate measurement of HbA\textsubscript{1c} has been questioned, with some POC HbA\textsubscript{1c} test devices reported not to meet accepted accuracy and precision criteria [3]. Ideal
imprecision goals for HbA\textsubscript{lc} should be coefficient of variation (CV) of <2% for HbA\textsubscript{lc} reported in % units (or <3% in SI units, mmol/mol) [4–6].

Most evaluations of POC HbA\textsubscript{lc} devices have taken place in laboratory settings [7, 8]; fewer studies have assessed device performance in a POC setting or with clinicians performing the tests [9, 10]. The only published review that has attempted to combine data from accuracy studies identified five studies covering three devices and compared correlation coefficients [11]. Systematically reporting and pooling data estimates of bias and precision between POC HbA\textsubscript{lc} devices and laboratory measurements would enable end users to assess which analysers best meet their analytical performance needs. This may be of particular importance for clinicians in primary care settings where much of the management of diabetes patients takes place. The comparison of accuracy between devices over the entire therapeutic range would need to be carried out by combining data on measurement error (bias) between POC and laboratory tests [12].

The aim of this study was to compare accuracy and precision of POC HbA\textsubscript{lc} devices with the local laboratory method based on data from published studies and discuss the clinical implications of the findings.

**Materials and methods**

**Data sources and searches**

The protocol was registered on the PROSPERO database (Registration number CRD42013006678). Searches of MEDLINE (1950–2015), EMBASE (1980–2015) and Web of Science databases (1900–2015) were carried out in June 2015 using search terms for glycated hemoglobin and POC device names, and Medical Subject Headings for POC systems (Medline search strategy in Table 1, Supplementary Data). Hand searches of original articles, reviews and conference abstracts were also carried out. Reports in all languages published in peer-reviewed journals were considered for inclusion.

**Study selection**

Title and abstract review of retrieved references and screening of full texts for eligibility were performed by two reviewers (JH and JM).

Included studies performed both a POC HbA\textsubscript{lc} test (defined as any instrument designed to provide a rapid quantitative measurement of HbA\textsubscript{lc} using capillary blood at the point of care) and laboratory-based methods on each sample and reported mean difference between POC and laboratory HbA\textsubscript{lc} (referred to hereafter as comparator method). The mean difference (or mean bias) from included studies was obtained using methods described by Bland and Altman for assessing agreement between two methods [13]. Studies were excluded if the study did not evaluate the accuracy of HbA\textsubscript{lc} POC devices, if there was no comparator test, if the article was not peer reviewed, data could not be extracted, patients had hemoglobinopathies or if the POC analysers evaluated were not commercially available on 30 June 2015.

**Data extraction and quality assessment**

Data were extracted by two reviewers on the following: sample type (capillary or venous blood), operator (clinical or laboratory operator), setting (point of care, laboratory or other), number of participants, mean HbA\textsubscript{lc} (in either SI, mmol/mol or % units) and mean (±standard deviation, SD) bias, defined as mean of the difference between POC and comparator method (POC-comparator) for each HbA\textsubscript{lc} sample measured. Reviewers were not blinded to any aspects of the studies during data extraction. Disagreements were resolved by discussion or, if consensus could not be reached, discussion with a third reviewer. If duplicate POC tests were performed using different reagent lot numbers, then both comparisons were included in the meta-analysis.

Instrument-specific data were extracted: POC device make and model, make and model of comparator method used and whether the comparator method was carried out in a NGSP (National Glycohemoglobin Standardization Program) or IFCC (International Federation for Clinical Chemistry and Laboratory Medicine) reference laboratory.

Study quality and risk of bias in data collection and methodology were graded using the QUADAS 2 risk of bias table [14]. Funding by device manufacturer and whether the range in HbA\textsubscript{lc} was sufficiently broad (defined as 5%–11% or 31–97 mmol/mol) were recorded. All studies deemed to be at high risk of bias for at least one criterion were excluded in a sensitivity analysis.

Data on the diagnostic accuracy of devices at a threshold of 6.5% (48 mmol/mol) was requested from study authors. Where data were not available, diagnostic accuracy was estimated from correlation and Bland-Altman graphs by counting the number of points in graph quadrants representative of diagnoses that would be truly negative, falsely positive, truly positive and falsely negative.
Data on imprecision recorded as CV was extracted in % units where reported. We combined both within and between-run precision at low (<6%, 42 mmol/mol), medium (6%–8%, 42–64 mmol/mol), and high (≥8%, 64 mmol/mol) levels of HbA1c, as well as total CV regardless of whether the samples were control samples or patient samples. The mean imprecision was combined by taking the average for each device at each level of HbA1c. Imprecision data on comparator instruments was also combined where this was reported.

Data synthesis and analysis

All analyses were carried out using Stata 12.0SE (StataCorp, College Station, TX, USA). Data on mean bias between POC and laboratory HbA1c and SD mean bias were combined using random effects DerSimonian and Laird meta-analysis [15] as recommended for meta-analyses of method comparison studies with considerable heterogeneity [16]. Data for each POC instrument were separately pooled. The following estimates were made: median difference was approximated to mean bias; if SD or standard error were not reported, then these were imputed averaging the SD from studies of the same device [17]. The SD of the mean bias and the standard error of the SD [18] were pooled in a separate meta-analysis to explore the variability in bias within each device. In a descriptive analysis, we reported bias and SD for each device without pooling. In meta-analyses, we summarized mean bias for each device and SD for each device with 95% confidence intervals. Because performance may vary between different clinical settings, we also show 95% prediction intervals (using the “rfdist” command in Stata) to display the estimated uncertainty about the mean bias, and SD, in any future study [19]. Studies in which estimates were made, those at high risk of bias and those in which HbA1c measurement were carried out more than 24 h apart were excluded in sensitivity analyses. Sensitivity analyses including only reports which used NGSP or IFCC reference laboratories for the comparator HbA1c, and those published in 2014–2015 were also conducted. Subgroup analyses were performed where possible to explore differences between operator (clinical or laboratory), setting (POC or laboratory), adults or children or blood sample type (venous or capillary) within a single device. Meta-regression was used where there were sufficient studies to establish whether different settings (POC or laboratory) or study publication year significantly affected the mean bias. A post hoc meta-regression was carried out to explore the effect of publication year on SD of the bias in the DCA device.

Where there were sufficient data, a bivariate random-effects method was used to estimate average sensitivity and specificity using the “metandi” command in Stata [20]. The average sensitivity and specificity was calculated and hierarchical summary receiver-operating characteristic curves were plotted to display the diagnostic accuracy across studies within a device.

Results

Searches identified 1739 records, a further six records were identified from scanning reference lists. There were 435 duplicates, and 1063 records were excluded after title and abstract review and 181 after full text review, leaving 66 studies; a further six studies were excluded because they included people with hemoglobin variants or data could not be extracted. Sixty studies were included in the analysis of mean bias from 13 devices (Supplementary Data, Figure 1). Forty-five included papers were full text articles, the remainder were conference abstracts, letters or comments. Twenty-five studies were carried out in Europe, 16 in the Americas, nine in Asia, eight in Australia and two unknown. Four articles were non-English and required translation [21–24]. Seven studies were conducted in children [25–31]. Some studies reported results from multiple devices or multiple lot numbers giving 105 comparisions in the meta-analysis. Characteristics of all included studies are shown in Table 1.

Devices compared in the meta-analysis were A1cGear (n=1), A1cNow (n=17), Afio (n=12), B-analyst (n=5), Clover (n=2), Cobas b101 (n=5), DCA 2000/Vantage (n=39), HemoCue (n=2), InnovaStar (n=5), Nyocard (n=6), Quo-Lab (n=3), Quo-Test (n=7), and SDA1cCare (n=1). The majority of studies reported bias in HbA1c in % units; only two studies reporting in mmol/mol [34, 62]. Nineteen studies were carried out in a laboratory setting by laboratory staff, 21 studies were carried out at point of care and in 12 studies, a clinician operated the POC device.

Mean bias and 2SD range of difference in HbA1c between each POC device and comparator method (Figure 1) shows that there is a large variability in bias for most devices. Pooling data on mean bias (Figure 2) in a meta-analysis shows that nine devices have a negative mean bias; this was significant for three devices (Quo-Lab, DCA and InnovaStar). Four devices had a positive mean bias which was significant for two devices (B-analyst and SDA1cCare). Heterogeneity between studies was large for all devices, with I² ranging from 67.5% to 99.5%. Although the pooled mean bias for some devices
Table 1: Table of included studies.

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>POC device(s)</th>
<th>Comparator method</th>
<th>Comparator instrument and manufacturer</th>
<th>POC sample</th>
<th>Comparator sample</th>
<th>Operator</th>
<th>Setting</th>
<th>Number of patients</th>
<th>Mean HbA1c %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arsie et al. [32]</td>
<td>2000</td>
<td>DCA 2000</td>
<td>Ion exchange</td>
<td>Diamat, Bio-Rad Laboratories Ltd</td>
<td>Venous</td>
<td>Venous</td>
<td>Laboratory</td>
<td>Laboratory</td>
<td>117</td>
<td>7.2</td>
</tr>
<tr>
<td>Azevedo et al. [33]</td>
<td>2011</td>
<td>DCA 2000 Quo-Test</td>
<td>Ion exchange</td>
<td>Biorad D10, Bio-Rad Laboratories Ltd</td>
<td>Capillary</td>
<td>Venous</td>
<td>–</td>
<td>–</td>
<td>52</td>
<td>8.1</td>
</tr>
<tr>
<td>Berry et al. [34]</td>
<td>2014</td>
<td>Cobas b101, HemoCue, DCA Vantage, Afinion</td>
<td>Ion exchange</td>
<td>Menarini HA 8140, A Menarini diagnostics Ltd</td>
<td>Venous</td>
<td>Venous</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Boz et al. [22]</td>
<td>1997</td>
<td>DCA 2000</td>
<td>Ion exchange</td>
<td>Adams Arkray HA8160, Menarini</td>
<td>Venous</td>
<td>Venous</td>
<td>–</td>
<td>Laboratory</td>
<td>100</td>
<td>7.9</td>
</tr>
<tr>
<td>Carrera Font et al. [23]</td>
<td>2011</td>
<td>Afinion</td>
<td>Ion exchange</td>
<td>Menarini HA 8140, A Menarini diagnostics Ltd</td>
<td>Venous</td>
<td>Venous</td>
<td>–</td>
<td>Laboratory</td>
<td>94</td>
<td>7.6</td>
</tr>
<tr>
<td>Chan et al. [25]</td>
<td>2014</td>
<td>DCA Vantage</td>
<td>Ion exchange</td>
<td>Menarini HA 8140, A Menarini diagnostics Ltd</td>
<td>Venous</td>
<td>Venous</td>
<td>–</td>
<td>Laboratory</td>
<td>155</td>
<td>–</td>
</tr>
<tr>
<td>Clark et al. [36]</td>
<td>2005</td>
<td>A1cNow</td>
<td>Ion exchange</td>
<td>Biorad Variant II, Bio-Rad Laboratories Ltd</td>
<td>Venous</td>
<td>Venous</td>
<td>–</td>
<td>Laboratory</td>
<td>120</td>
<td>8.1</td>
</tr>
<tr>
<td>Criel et al. [37]</td>
<td>2015</td>
<td>Cobas b101, Afinion, B-analyst</td>
<td>Ion exchange</td>
<td>Menarini HA 8160, A Menarini diagnostics Ltd</td>
<td>Venous</td>
<td>Venous</td>
<td>–</td>
<td>–</td>
<td>40</td>
<td>–</td>
</tr>
<tr>
<td>Dupuy et al. [31]</td>
<td>2013</td>
<td>DCA Vantage</td>
<td>Ion exchange</td>
<td>Menarini HA 8160, A Menarini diagnostics Ltd</td>
<td>Venous</td>
<td>Venous</td>
<td>–</td>
<td>–</td>
<td>124</td>
<td>–</td>
</tr>
<tr>
<td>Dupuy et al. [38]</td>
<td>2014</td>
<td>Afinion</td>
<td>Ion exchange</td>
<td>Menarini HA 8160, A Menarini diagnostics Ltd</td>
<td>Venous</td>
<td>Venous</td>
<td>–</td>
<td>Laboratory</td>
<td>155</td>
<td>–</td>
</tr>
<tr>
<td>El Arabi et al. [26]</td>
<td>2013</td>
<td>DCA Vantage</td>
<td>Ion exchange</td>
<td>Biorad Variant II, Bio-Rad Laboratories Ltd</td>
<td>Venous</td>
<td>Venous</td>
<td>–</td>
<td>Laboratory</td>
<td>120</td>
<td>–</td>
</tr>
<tr>
<td>Estevez et al. [40]</td>
<td>2014</td>
<td>DCA Vantage</td>
<td>Ion exchange</td>
<td>Biorad Variant II, Bio-Rad Laboratories Ltd</td>
<td>Venous</td>
<td>Venous</td>
<td>–</td>
<td>–</td>
<td>42</td>
<td>–</td>
</tr>
<tr>
<td>Ginde et al. [41]</td>
<td>2008</td>
<td>A1cNow</td>
<td>Immunoassay</td>
<td>Bio-Rad D10, Bio-Rad Laboratories Ltd</td>
<td>Capillary</td>
<td>Venous</td>
<td>Laboratory</td>
<td>POC</td>
<td>161</td>
<td>5.8</td>
</tr>
<tr>
<td>Greaves et al. [27]</td>
<td>2005</td>
<td>DCA 2000</td>
<td>Ion exchange</td>
<td>Bio-Rad D10, Bio-Rad Laboratories Ltd</td>
<td>Capillary</td>
<td>Venous</td>
<td>Laboratory</td>
<td>POC</td>
<td>228</td>
<td>–</td>
</tr>
<tr>
<td>Harris et al. [43]</td>
<td>2000</td>
<td>DCA 2000</td>
<td>Immunoassay</td>
<td>LX20 Beckman Coulter</td>
<td>Capillary</td>
<td>Venous</td>
<td>–</td>
<td>Clinical</td>
<td>102</td>
<td>–</td>
</tr>
<tr>
<td>Heng et al. [45]</td>
<td>2015</td>
<td>Cobas b101</td>
<td>Immunoassay</td>
<td>Roche Tina-quant A1c.gen3, Roche Diagnostics Ltd</td>
<td>Capillary</td>
<td>Venous</td>
<td>–</td>
<td>–</td>
<td>47</td>
<td>–</td>
</tr>
<tr>
<td>Holmes et al. [46]</td>
<td>2008</td>
<td>DCA 2000</td>
<td>Ion exchange</td>
<td>Biorad Variant II, Bio-Rad Laboratories Ltd</td>
<td>Venous</td>
<td>Venous</td>
<td>Laboratory</td>
<td>Laboratory</td>
<td>24, 45, 21</td>
<td>–</td>
</tr>
<tr>
<td>Jou et al. [48]</td>
<td>2013</td>
<td>A1cNow</td>
<td>Ion exchange</td>
<td>Toosh G8, Toosh Bioscience Ltd</td>
<td>Venous</td>
<td>Venous</td>
<td>–</td>
<td>–</td>
<td>57</td>
<td>–</td>
</tr>
<tr>
<td>Karami and Baradaran [49]</td>
<td>2014</td>
<td>Nycocard</td>
<td>Affinity</td>
<td>Nycocard</td>
<td>Venous</td>
<td>Venous</td>
<td>–</td>
<td>–</td>
<td>58</td>
<td>5.8</td>
</tr>
<tr>
<td>Knaebel et al [50]</td>
<td>2013</td>
<td>A1cNow</td>
<td>Ion exchange</td>
<td>Toosh G8, Toosh Bioscience Ltd</td>
<td>Venous</td>
<td>Venous</td>
<td>–</td>
<td>–</td>
<td>40</td>
<td>–</td>
</tr>
<tr>
<td>Leal and Soto-Roven [51]</td>
<td>2009</td>
<td>A1cNow</td>
<td>Ion exchange</td>
<td>Toosh G8, Toosh Bioscience Ltd</td>
<td>Venous</td>
<td>Venous</td>
<td>–</td>
<td>Pharmacist</td>
<td>47</td>
<td>8.1</td>
</tr>
<tr>
<td>Leca et al. [52]</td>
<td>2012</td>
<td>DCA Vantage</td>
<td>Ion exchange</td>
<td>Biorad Variant II, Bio-Rad Laboratories Ltd</td>
<td>Venous</td>
<td>Venous</td>
<td>–</td>
<td>Laboratory</td>
<td>100</td>
<td>7.3</td>
</tr>
<tr>
<td>Lee et al. [53]</td>
<td>2015</td>
<td>SDA1c Care</td>
<td>Ion exchange</td>
<td>Biorad Variant II, Bio-Rad Laboratories Ltd</td>
<td>Venous</td>
<td>Venous</td>
<td>–</td>
<td>Laboratory</td>
<td>150</td>
<td>–</td>
</tr>
<tr>
<td>Author</td>
<td>Year</td>
<td>Method</td>
<td>Comparator Instrument and Manufacturer</td>
<td>POC Sample</td>
<td>Comparator Sample</td>
<td>Operator</td>
<td>Setting</td>
<td>Number of Patients</td>
<td>Mean HbA₁c, %</td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>------</td>
<td>--------</td>
<td>----------------------------------------</td>
<td>-----------</td>
<td>------------------</td>
<td>----------</td>
<td>---------</td>
<td>----------------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>Lenters-Westra et al. [3]</td>
<td>2010a</td>
<td>DCA Vantage</td>
<td>Tosoh G7, Tosoh Bioscience Ltd</td>
<td>Venous</td>
<td>Venous</td>
<td>Laboratory</td>
<td>40</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Lenters-Westra et al. [3]</td>
<td>2010b</td>
<td>Quo-Test</td>
<td>Tosoh G8, Tosoh Bioscience Ltd</td>
<td>Venous</td>
<td>Venous</td>
<td>Laboratory</td>
<td>40</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Lenters-Westra and Slingerland [54]</td>
<td>2010b</td>
<td>Afinion</td>
<td>Tosoh G8, Tosoh Bioscience Ltd</td>
<td>Venous</td>
<td>Venous</td>
<td>Laboratory</td>
<td>40</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Lenters-Westra and Slingerland [7]</td>
<td>2014</td>
<td>DCA Vantage</td>
<td>InnoAfinion</td>
<td>Venous</td>
<td>Venous</td>
<td>Laboratory</td>
<td>40</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Malkani et al. [55]</td>
<td>2013</td>
<td>DCA Vantage</td>
<td>Tosoh G8, Tosoh Bioscience Ltd</td>
<td>Venous</td>
<td>Venous</td>
<td>Laboratory</td>
<td>40</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Manley et al. [56]</td>
<td>2014</td>
<td>DCA 2000</td>
<td>A1cNow</td>
<td>Venous</td>
<td>Venous</td>
<td>Laboratory</td>
<td>128</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Marley et al. [57]</td>
<td>2015</td>
<td>DCA Vantage</td>
<td>A1cNow</td>
<td>Venous</td>
<td>Venous</td>
<td>Laboratory</td>
<td>241</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Martin et al. [58]</td>
<td>2005</td>
<td>DCA 2000</td>
<td>Primer Mini, Biorad Laboratories Ltd</td>
<td>Venous</td>
<td>Venous</td>
<td>–</td>
<td>88</td>
<td>7.1</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>Matthew et al. [59]</td>
<td>2007</td>
<td>A1cNow</td>
<td>Tosoh 2.2, Tosoh Bioscience Ltd</td>
<td>Venous</td>
<td>Venous</td>
<td>Laboratory</td>
<td>34</td>
<td>7.1</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Menendez-Valladares et al. [60]</td>
<td>2015</td>
<td>B-analyst</td>
<td>Menarini HA8180, A Menarini diagnostics Ltd</td>
<td>Venous</td>
<td>Venous</td>
<td>Laboratory</td>
<td>120</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Nam et al. [61]</td>
<td>2011</td>
<td>A1cNow</td>
<td>Ns</td>
<td>Capillary</td>
<td>Venous</td>
<td>Laboratory</td>
<td>83</td>
<td>9.1</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Petersen et al. [8]</td>
<td>2010</td>
<td>DCA Vantage</td>
<td>Biorad Variant II, Bio-Rad Laboratories Ltd</td>
<td>Venous</td>
<td>Venous</td>
<td>Laboratory</td>
<td>128</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Phillips et al. [62]</td>
<td>2014</td>
<td>Quo-Test</td>
<td>A1cNow</td>
<td>Venous</td>
<td>Venous</td>
<td>Clinical</td>
<td>45</td>
<td>7.9</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Page et al. [63]</td>
<td>2007</td>
<td>DCA 2000</td>
<td>Immunoassay</td>
<td>Venous</td>
<td>Venous</td>
<td>POC</td>
<td>39–80</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Page et al. [64]</td>
<td>2007</td>
<td>DCA 2000</td>
<td>Immunoassay</td>
<td>Venous</td>
<td>Venous</td>
<td>POC</td>
<td>39–80</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Page et al. [65]</td>
<td>2007</td>
<td>A1cNow</td>
<td>Biorad Variant II, Bio-Rad Laboratories Ltd</td>
<td>Venous</td>
<td>Venous</td>
<td>Clinical</td>
<td>453</td>
<td>8.6</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Page et al. [66]</td>
<td>2007</td>
<td>A1cNow</td>
<td>Biorad Variant II, Bio-Rad Laboratories Ltd</td>
<td>Venous</td>
<td>Venous</td>
<td>Clinical</td>
<td>453</td>
<td>8.6</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Schenker and Whiting [68]</td>
<td>2006</td>
<td>DCA 2000</td>
<td>Affinity</td>
<td>Venous</td>
<td>Venous</td>
<td>POC</td>
<td>117</td>
<td>5.9</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Shepherd and Whiting [69]</td>
<td>2006</td>
<td>DCA 2000</td>
<td>Affinity</td>
<td>Venous</td>
<td>Venous</td>
<td>POC</td>
<td>117</td>
<td>5.9</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Shemesh et al. [67]</td>
<td>2006</td>
<td>DCA 2000</td>
<td>Affinity</td>
<td>Venous</td>
<td>Venous</td>
<td>POC</td>
<td>117</td>
<td>5.9</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Shemesh et al. [67]</td>
<td>2006</td>
<td>DCA 2000</td>
<td>Affinity</td>
<td>Venous</td>
<td>Venous</td>
<td>POC</td>
<td>117</td>
<td>5.9</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Shimoda et al. [69]</td>
<td>2013</td>
<td>A1cNow</td>
<td>Tosoh G8, Tosoh Bioscience Ltd</td>
<td>Venous</td>
<td>Venous</td>
<td>Clinical</td>
<td>144</td>
<td>7.3</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>St John et al. [9]</td>
<td>2013</td>
<td>A1cNow</td>
<td>Tosoh G8, Tosoh Bioscience Ltd</td>
<td>Venous</td>
<td>Venous</td>
<td>Clinical</td>
<td>144</td>
<td>7.3</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>
was small across all studies, bias within a single device could vary substantially between studies. For example, the mean bias across studies was $-0.05\%$ HbA$_1c$ for the A1cNow device, but it varied from $-0.70\%$ (95\% confidence interval, CI $-0.82$ to $0.58\%$) $[56]$ to $+0.67\%$ (95\% CI $0.52$–$0.82\%$) $[21]$. Similarly for the DCA device, mean bias varied from $-0.96\%$ to $+0.28\%$. Even when the mean bias was small, individual differences in measurements within a single study could be larger, as shown by the two SD bars (Figure 1), which typically extended over a range of $1\%$–$1.5\%$ HbA$_1c$ ($10.9$–$16.4$ mmol/mol). Moreover, the width of the prediction intervals in Figure 2 indicate that for most devices the mean bias of a future study could vary widely. Summary estimates of the SD of the differences (Figure 3) ranged from $0.21\%$ to $0.53\%$ but again individual SDs could be larger and varied between studies. Meta-regression did not find that year of publication significantly affected the mean bias for any of the devices but there was a trend towards a less negative bias over time in the DCA device (coefficient $0.014\%$ HbA$_1c$ per year, $p=0.081$). A post hoc meta-regression found that the standard deviation of the bias in the DCA device significantly decreased over time ($-0.01\%$ HbA$_1c$ per year, 95\% CI $-0.015$ to $-0.006$, $p<0.0001$), suggesting that the variability in measurements contributing to the mean bias in the DCA device is decreasing over time. A subgroup analysis to compare the size of the SD in studies published prior to 2006 with those published between 2006 and 2015 found that those published in the last 10 years had a mean SD of $0.33\%$ (95\% CI $0.30$–$0.37$) HbA$_1c$ compared with a SD of $0.45\%$ (95\% CI $0.38$–$0.52$) HbA$_1c$ before 2006.

Sensitivity analyses excluding studies in which approximations were necessary during data extraction and those at highest risk of bias (Supplementary Data, Figures 2 and 3) gave broadly similar results. Nine studies which used NGSP or IFCC reference laboratories for the comparator method no longer had data for three devices (A1cGear, HemoCue and SDA1cCare). There was no longer a significant mean bias for the DCA device and variability was lower (mean bias $0.00\%$, 95\% CI $-0.16$ to $0.15$), whereas the Quo-Test had a significantly negative bias (mean bias $-0.27\%$, 95\% CI $-0.50$ to $-0.03$) (Supplementary Data, Figure 4). Studies published in 2014–2015 also had a non-significant mean bias for the DCA device ($-0.08\%$, 95\% CI $-0.21$ to $0.06$).

Subgroup analyses for the DCA device comparing different settings found no difference in mean bias in studies carried out in a point of care setting $-0.27\%$ (95\% CI $-0.38$ to $-0.16$) compared with a laboratory setting $-0.30\%$ (95\% CI $-0.47$ to $-0.14$). Similarly, mean bias in studies carried out by a laboratory operator was not...
Figure 1: Forest plot showing mean bias (in % HbA1c)±2SD of HbA1c measured using point-of-care devices compared with laboratory tests, by POC device, ordered by mean bias.

significantly different from that in those carried out by a clinical operator (−0.30%, 95% CI −0.47 to −0.14 vs. −0.30%, 95% CI −0.42 to −0.18, p=0.956) (Supplementary Data, Figure 5). Sensitivity analysis gave similar results (data not shown). Subgroup analyses carried out for the A1cNow device also found no differences in mean bias.
for different operators or settings (results not shown). There was no difference in bias in six studies carried out in children compared with 30 studies carried out in adults on the DCA device (–0.24%, 95% CI –0.33 to –0.16 in adults compared with –0.30%, –0.50 to –0.10 in children). There was no difference in mean bias between venous or capillary blood in the AlcNow or DCA (Supplementary Data, Figures 6 and 7).

Figure 2: Forest plot showing mean bias (in % HbA1c) and 95% CI and prediction intervals, by POC device, ordered by mean bias.
There were sufficient data to carry out meta-analysis on diagnostic accuracy for five devices (Afinion, DCA, A1cNow, Quo-Test and Nycocard). Sensitivity across all the devices was similar, although specificity varied more, with the Afinion and DCA having the highest specificity at a cutoff of 6.5% HbA1c (48 mmol/mol) (Supplementary Figure 3: Forest plot showing standard deviation of differences, 95% confidence intervals of SD and prediction intervals for included studies, by POC device, ordered by SD.)
Data, Table 2 and Figure 8). Repeating this analysis only in studies published from 2006 to 2015 did not change the results.

Imprecision measured using replicate analyses of a single sample for each analyzer at low, medium and high HbA1c are presented as mean CV in the Supplementary Data (Table 3). All devices bar one which only had a single evaluation had mean CVs above 2% at low HbA1c values (lower than 6% HbA1c, 42 mmol/mol). At high HbA1c (above 8% HbA1c, 64 mmol/mol), mean CV was <2% for 4 devices, all of which only included data from one study. Five devices had a total CV <2%. Only two devices had all measured CVs below 2% (Cobas b101 and A1cgear) but both of these were only based on a single evaluation. Imprecision of comparator methods is shown in Table 4 in Supplementary Data. Each CV was from a single study, in some cases using multiple data points from repeat analysis of different samples. CVs ranged from 0.8% for Tosoh (model not specified) to 2.3% for the Diamat instrument. Between-laboratory precision reported in the 2016 CAP survey [77] is shown for the comparator methods and Afinion, DCA 2000 and DCA Vantage devices in Table 4 in Supplementary Data which range from 1.6% for the Tosoh G8 to 3.1% for Afinion.

Study quality/risk of bias

Sixteen studies reported receiving funding from the POC device manufacturers [7, 9, 27, 30, 35, 36, 39, 41, 47, 50, 51, 56, 58, 59, 65, 70]. In 18 of the included studies, the instrument was provided free of charge by the manufacturers [3, 7, 9, 24, 27, 31, 36, 38, 39, 44, 50, 53, 54, 56, 59, 65, 69, 70]. In eight studies, it was unclear what laboratory methodology was used in the comparator method [22, 33, 35, 41, 43, 51, 61, 74]. Another seven studies did not use a wide range of HbA1c concentrations to assess mean bias across all HbA1c concentrations [24, 25, 41, 50, 57, 71, 73]. Nine studies used an NGSP- or IFCC-certified laboratory for the comparator method [3, 7, 29, 41, 50, 54, 56, 59, 60].

Discussion

This analysis has shown that the majority of POC HbA1c devices have a mean negative bias compared to laboratory methods. Devices with the most comparisons were found to have higher specificity, rather than consistently high sensitivity, for detecting HbA1c above the diagnostic threshold. Unexplained variation in the mean bias was also observed for most devices, as indicated by large variations in SDs of the bias as well as high imprecision from the repeated analysis of the same sample. This is the first meta-analysis to compare performance between HbA1c measured using POC devices and laboratory methods drawn from published data. We have used a novel approach to pool data on mean bias, variability, device precision and sensitivity and specificity. Our analysis has demonstrated that differences between POC HbA1c and comparator methods can vary considerably within a single device across all the included studies with POC values ranging from as much as 1.5% HbA1c below to 1.5% HbA1c above the comparator method HbA1c across all devices.

Many evaluations included in our review compared capillary blood on the POC device with venous blood on the comparator method. Although there may be slight differences in HbA1c in different sample types [78, 79] we have been able to eliminate this as a source of bias in two devices (A1cNow and DCA). We did not find evidence that changes in calibration over time resulted in significantly lower levels of bias in device; however, these data suggest that variability in the DCA device may be improving over time, suggesting that its precision may also have improved over time. The diagnostic accuracy of devices was not affected when only the most recent studies were included.

Acceptable limits for bias and imprecision vary between organizations; the College of American Pathologists (CAP) criteria state that acceptable limits of bias are ±6% [80], whereas a within-laboratory CV of <2% (in % units) is widely recommended [4, 80]. More recently, the IFCC Task Force on HbA1c Standardization have recommended using a sigma metrics approach to evaluating analytical error. This takes into account both bias and imprecision in one model based on total allowable error, where bias and imprecision are not treated as separate performance measures [81]. The CAP survey [80] includes data from HbA1c POC devices and shows comparable performance to a number of routine laboratory analysers. However, the survey is aimed at clinical laboratories and participation is not mandatory for instruments used in non-laboratory settings. Those that enter the survey may therefore be a self-selecting group and under-represent the true extent of the use of POC devices in the USA.

Limitations

Sensitivity and specificity at the diagnostic threshold (6.5% HbA1c, 48 mmol/mol) were not reported in the included papers. However, we have been able to combine
data from plots and unpublished data from authors to compare sensitivity and specificity of five devices at the diagnostic threshold.

Some laboratory instruments have been reported to have systematic biases against reference laboratories [46] which may have led to under- or over-estimation of biases in some studies in our review. Only nine studies reported to use an NGSP or IFCC reference laboratory for the comparator method which generally resulted in lower variability. The actual number of studies using reference laboratories may have been higher as many of the papers included in this review were conference abstracts which were limited by word count.

Even though the intended end-users of POC devices are clinicians or healthcare workers, the majority of the accuracy studies were carried out in a laboratory setting by laboratory staff. The translatability of the evaluations included in the meta-analysis to the clinical end-user is therefore a major limitation of this study. We have been able to examine the impact of operator and setting on bias for two devices, DCA and A1cNow, both of which showed no significant difference in performance between a clinical and laboratory operator. It is important to note that practices and regulations in POC testing vary between countries and therefore POC testing may impact differently on patient management between geographic locations.

Mean values from duplicate POC tests were used by some researchers [64], and this may lead to the reported variability being lower than actual variability. In some studies, it was not clear whether results were reported as relative mean bias or absolute mean bias; however, we have excluded studies with unclear reporting in a sensitivity analysis and found similar results.

Clinical implications

Because the majority of devices had an overall negative bias as well as large standard deviations, their use in a clinical setting may result in differences in treatment decisions compared with results from laboratory comparator methods. POC HbA1c results which are lower than laboratory methods may result in undertreatment in some cases. The implications of this on overall patient care may result in fewer short-term hypoglycemic episodes, however, the long-term use may leave some patients exposed to higher than optimal levels of glycemia. The long-term effects of this are not known, but studies have shown that even small increases in HbA1c increase the risk of micro- and macrovascular complications [1]. This may be outweighed by the other benefits of POC testing including improved ability of the clinicians to make appropriate and timely treatment decisions [82, 83]. There was insufficient published data to enable comparison of performance between settings and operators for most devices; further evaluations in a clinical setting are now needed.

As well as bias, good reproducibility or precision is crucial in device performance; however, imprecision in nearly all devices was poor with the majority of devices having mean CVs >2% at low HbA1c levels. In contrast, only one of the comparator instruments had a mean CV over 2%. These imprecision data are based on repeated analysis of a single sample; our review has shown that there is also considerable variability in HbA1c results in most devices in evaluations over a range of HbA1c concentrations which may better represent clinical situations. These high levels of uncertainty in results may mean that some real changes in HbA1c may not be picked up and could result in different treatment decisions. The combination of instrument performance and biological variability in HbA1c [84] could substantially affect the numbers of patients diagnosed with diabetes and impact on clinical decision making for those with established diabetes. These high levels of imprecision and the large variability is a key finding of this review and may be a crucial factor influencing the adoption of these technologies.

This analysis contributes to the information available to clinicians who are considering whether to use POC HbA1c testing. The decision to select an analyzer must be carefully considered based on clinical needs and what is deemed to be acceptable bias and precision of the analyzer. The NGSP recommends that for method certification 37 out of 40 HbA1c tests should be within ±6% relative to the standard reference laboratory measurement [80].

Our analysis shows that the mean measurement bias of two devices was >6% relative to the comparator method, however, because of the large variability observed within devices it is probable that many of the devices will give some measurements greater or <6% of the patient’s true HbA1c. It is therefore likely that some of the devices would fail to meet the requirements of 37 of 40 test results within 6% of reference laboratory measurements.

The impact of using POC HbA1c testing on medication use, clinical decision making and patient outcomes now needs to be evaluated in a randomized trial with full economic evaluation.

Acknowledgments: This report is independent research supported by the National Institute for Health Research (NIHR) (Doctoral Research Fellowship, Ms Jennifer Hirst, DRF-2013-06-086). The views expressed in this publication are those of the author(s) and not necessarily those of
the NHS, NIHR or the Department of Health. We are grateful to Dr. Erna Lenters-Westra, Samir Malkani, Bernard Vialettes and Katrina Ruedy on behalf of Jamie Wood for providing us with unpublished data. We also thank Ms. Nia Roberts for her assistance in optimizing the search strategy.

Author contributions: All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission. J.H. designed the study, researched data, conducted the analysis and wrote manuscript, J.M. researched data and reviewed/edited manuscript. C.P. reviewed/editd manuscript and contributed to discussion. E.E. reviewed/editd manuscript and contributed to discussion. B.F. researched data, contributed to the analysis and reviewed/editd manuscript. R.S. contributed to the analysis, reviewed/editd manuscript and contributed to discussion. A.F. reviewed/editd manuscript and contributed to discussion.

Research funding: None declared.

Employment or leadership: None declared.

Honorarium: None declared.

Competing interests: C.P. is a member of the NIHR Diagnostic Evidence Co-operative Oxford. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health. The funding organization(s) played no role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the report for publication. The researchers retained complete independence in the conduct of this study. A.F. is an NIHR Senior Investigator.

References


22. Boz M, Gerard P, Scheen AJ, Lefebvre PJ, Castillo MJ. Rapid determination by immunoassay of glycosylated hemoglobin in capillary blood compared to an affinity method for boronate and


45. Hirst et al.: A systematic review and meta-analysis — 179


