

Glycated haemoglobin measurements from UK Biobank are different to those in linked primary care records: implications for combining biochemistry data from research studies and routine clinical care

Authors: Katherine G Young,^{1*} Timothy J McDonald^{1,2} and Beverley M Shields¹

¹Institute of Biomedical and Clinical Science, College of Medicine and Health, University of Exeter, Exeter, UK, ²Department of Clinical Biochemistry, Royal Devon and Exeter NHS Foundation Trust, Exeter, UK

*Corresponding Author: Level 3 RILD Building, RD&E Wonford, Barrack Road, Exeter, Devon, EX2 5DW, UK. E-mail: k.young3@exeter.ac.uk

Word count: 649

Biochemical tests of the same individual carried out on different test platforms are often not comparable due to bias in assay method and calibration^{1,2}. Combining measurements from different sources, or diagnoses based on these measurements, is therefore not always valid. We highlight an example using glycated haemoglobin A1c (HbA1c) test results from two different sources in UK Biobank data: HbA1c measurements taken at baseline assessment using a single assay method (the Bio-Rad Variant II Turbo HPLC analyser³); and HbA1c measurements from linked UK primary care records, where assay method was dependent on which NHS laboratory the sample was processed in.

We identified UK Biobank participants with no pre-existing or previous diagnosis of diabetes mellitus (any type), with a primary care HbA1c measurement ≤ 100 days before or after baseline assessment ($n=1,039$; a detailed method is provided in Supplementary Figure S1). In individuals without diabetes, HbA1c should be relatively stable within this short timeframe. We found that UK Biobank baseline measurements were on average lower than primary care measurements with a mean difference of 2 mmol/mol (Figure 1), regardless of whether the primary care measurement was taken before or after baseline assessment.

The difference in measurements from the two sources may be due to a number of factors. These include biological variation within an individual due to the time difference between the primary care measurement and baseline assessment (-100 to +100 days). However, there was not a significant association between this time difference and the difference in the measurement values (Pearson correlation coefficient = -0.025, $P = 0.4$; see Supplementary Figure S2). The difference in HbA1c values is therefore likely to be due to methodological differences. We consider the most probable contributors to be the use of different HbA1c analysers, and differences in sample storage (UK Biobank stored blood samples frozen for 4-10 years prior to analysing⁴). A brief discussion of these methodological differences and their potential contributions can be found in Supplementary Table S2. It should be noted that the UK Biobank and all NHS laboratories are registered with external quality assurance

(EQA) schemes which verify the performance of HbA1c assays³, indicating that measurements from both sources meet the appropriate standards to be used for clinical decision making.

Small differences in measurements from different sources, which may not be clinically significant at an individual patient level, can result in large differences when used to define disease cases in a large cohort. Using HbA1c thresholds to identify cases of pre-diabetes (HbA1c 42.0-47.9 mmol/mol) and type 2 diabetes (HbA1c \geq 48.0 mmol/mol) in this dataset results in under-diagnosis of pre-diabetes and diabetes when using UK Biobank HbA1c measurements compared to primary care HbA1c measurements, due to the differences between these measurements (Figure 2).

Participants who enter UK Biobank with a clinical diagnosis of type 2 diabetes will most likely have received this in primary care on the basis of an HbA1c measurement \geq 48.0 mmol/mol. Additional cases of diabetes identified by a UK Biobank HbA1c \geq 48.0 mmol/mol as in (5) cannot be treated as equivalent, because of the apparent difference between UK Biobank HbA1c measurements and primary care measurements. Re-aligning UK Biobank measurements using the equation of the linear regression line shown in Figure 1 as per Cull et al.⁶ may improve the comparability of these measurements. However, caution should be exercised as the $n=1,039$ primary care measurements in Figure 1 represent a range of different analysers, and only a subset of those used throughout the NHS for diagnostic diabetes testing.

Combining biochemical data from research studies with data from routine clinical care is becoming increasingly commonplace, whether for prospective cohorts such as UK Biobank or for randomised controlled trial participants⁷. Care must be taken to ensure that such data are equivalent and that it is valid to combine them. This is of particular importance when biomarkers are used to determine disease states, as in the case of HbA1c.

Ethics approval

UK Biobank received ethical approval from the North West Multi-centre Research Ethics Committee (REC reference: 11/NW/03820). All participants gave written informed consent.

Author contributions

BMS and KGY conceived the study. KGY conducted the data analysis. All authors offered advice on the study design, analysis and interpretation of the results. KGY wrote the first draft of the manuscript. All authors read, reviewed, revised and approved the final manuscript.

Data availability statement

UK Biobank data are available through a procedure described at <http://www.ukbiobank.ac.uk/using-the-resource/>.

Supplementary data

Supplementary data are available at IJE online.

Funding

This research was supported by Research England's Expanding Excellence in England (E3) fund.

Acknowledgments

This research has been conducted using the UK Biobank resource under Application Number 9072.

Conflict of Interest

None declared.

References

1. The EurA1c Trial Group. 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. *Clinical Chemistry* 2018; **64**: 1183-92.
2. Wiedmeyer H-M, Polonsky KS, Myers GL, et al. International Comparison of C-Peptide Measurements. *Clinical Chemistry* 2007; **53**: 784-7.
3. Tierney A, Fry D, Almond R, Gordon M, Moffat S. UK Biobank Biomarker Enhancement Project – Companion Document to Accompany HbA1c Biomarker Data V1.0. 2018. https://biobank.ndph.ox.ac.uk/showcase/showcase/docs/serum_hb1ac.pdf. (3 Aug 2021, date last accessed).
4. UK Biobank. Biomarker assay quality procedures: approaches used to minimise systematic and random errors (and the wider epidemiological implications) V1.2. 2019. https://biobank.ctsu.ox.ac.uk/crystal/crystal/docs/biomarker_issues.pdf. (3 Aug 2021, date last accessed).
5. Tavaglione F, De Vincentis A, Jamialahmadi O, et al. Inborn and acquired risk factors for severe liver disease in Europeans with type 2 diabetes from the UK Biobank. *JHEP Reports* 2021; **3**: 100262.
6. Cull CA, Manley SE, Stratton IM, et al. Approach to maintaining comparability of biochemical data during long-term clinical trials. *Clinical Chemistry* 1997; **43**: 1913-8.
7. Mc Cord KA, Al-Shahi Salman R, Treweek S, et al. Routinely collected data for randomized trials: promises, barriers, and implications. *Trials* 2018; **19**: 1-9.

Figure 1

Scatter plot of HbA1c measurements taken at UK Biobank baseline assessment vs those in primary care taken ≤ 100 days before or after baseline assessment for $n=1,039$ individuals with no diagnosis of diabetes. Inset shows 30-50 mmol/mol values in more detail. Solid line: linear regression (equation: $y = 0.9696x + 3.3595$); dashed line: line of equality.

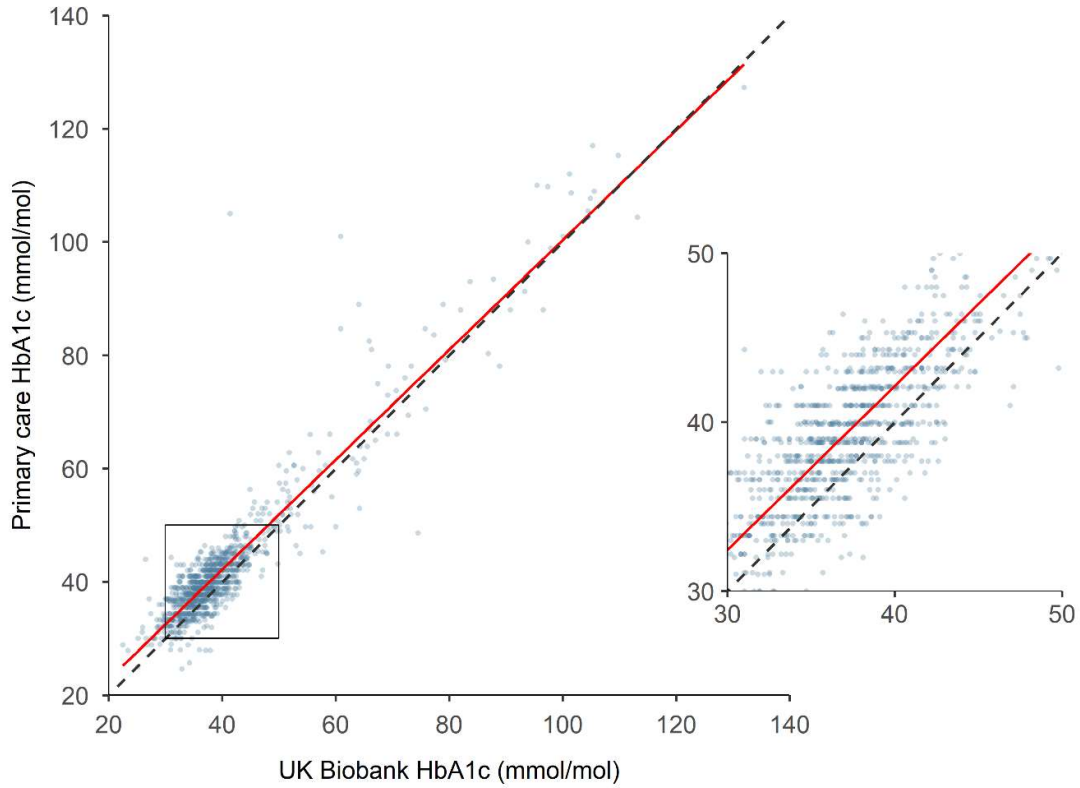
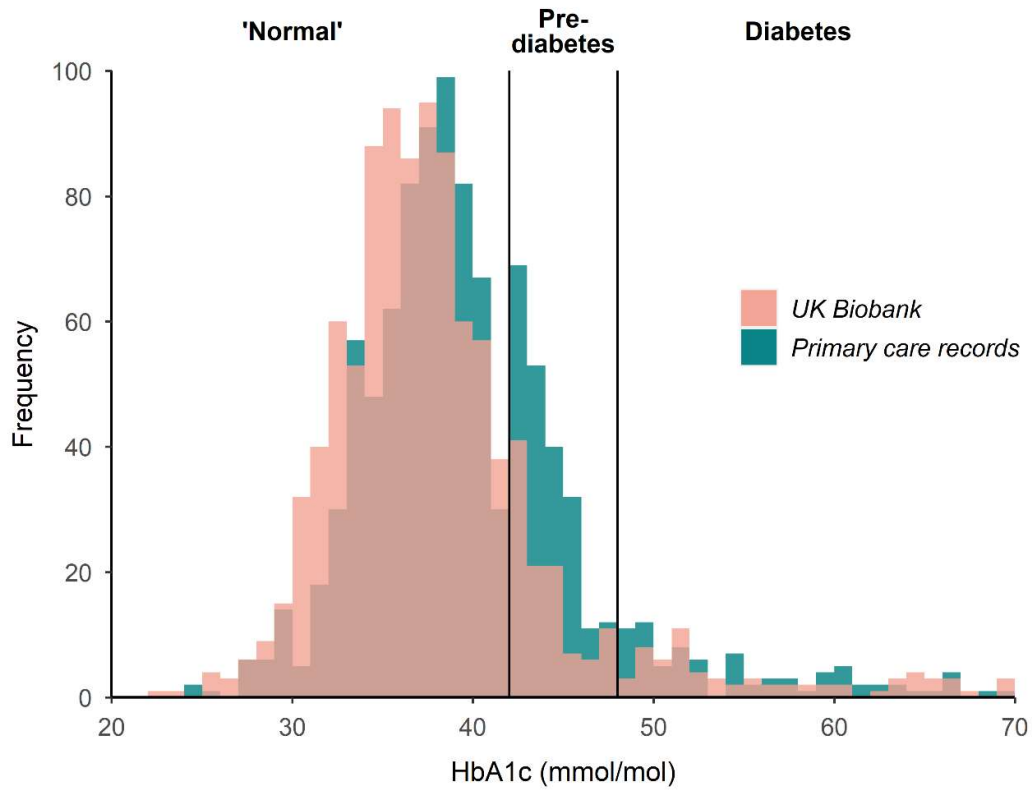


Figure 2

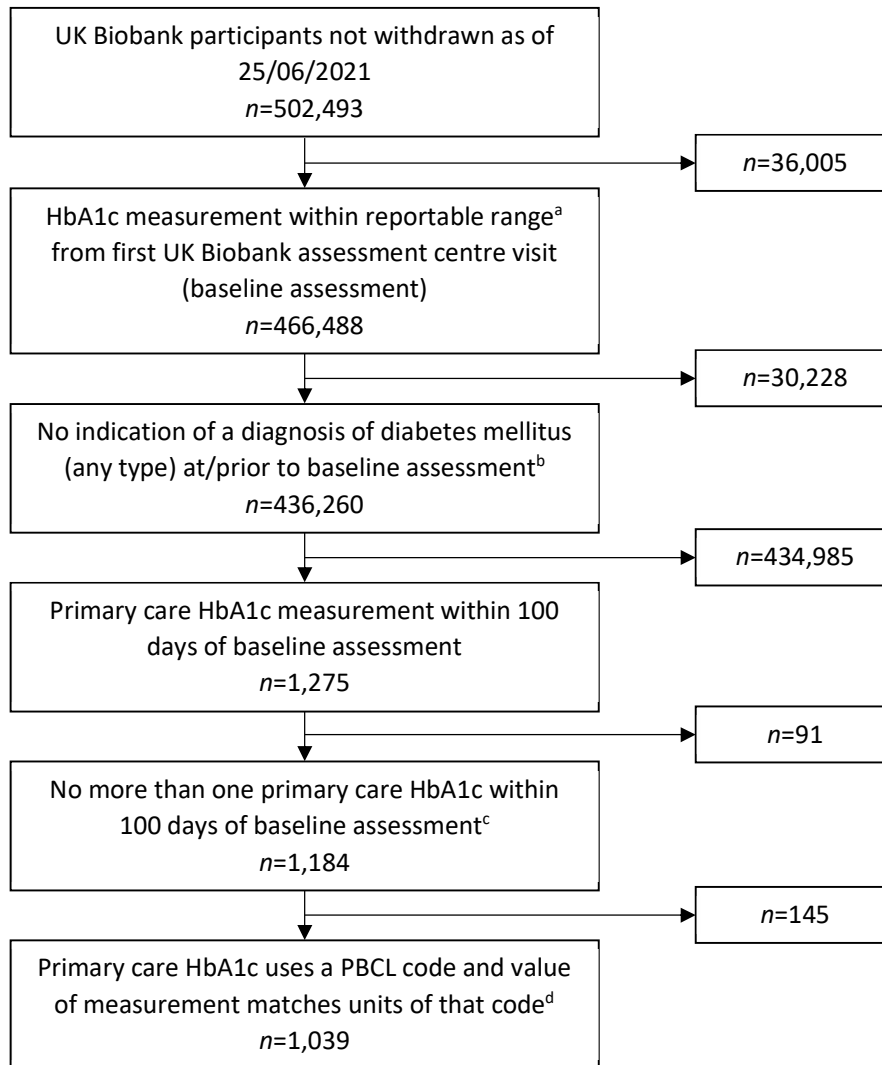
Histogram of HbA1c measurements taken at UK Biobank baseline assessment and in primary care records taken ≤ 100 days before or after baseline assessment for $n=1,039$ individuals with both measurements and no diagnosis of diabetes (x-axis truncated at 70 mmol/mol). HbA1c measurements in the 'normal' (< 42.0 mmol/mol), pre-diabetes (42.0-47.9 mmol/mol), and diabetes (≥ 48.0 mmol/mol) ranges are indicated; counts for these categories are shown underneath.



	'Normal' (HbA1c < 42.0 mmol/mol)	Pre-diabetes (HbA1c 42.0-47.9 mmol/mol)	Diabetes (HbA1c ≥ 48.0 mmol/mol)
<i>UK Biobank</i>	824 (79%)	112 (11%)	103 (10%)
<i>Primary care records</i>	670 (64%)	240 (23%)	129 (12%)

Supplementary Data

Supplementary Figure S1: Method for selecting UK Biobank participants with primary care HbA1c measurements within 100 days of baseline assessment



^a Reportable range for UK Biobank HbA1c measurements was 15-184 mmol/mol (inclusive)³.

^b Participants with any indication of a diagnosis of diabetes mellitus (any type) at/prior to baseline assessment were removed in case they were receiving treatment resulting in rapid changes in HbA1c in the 100 day window between blood draws for the primary care and UK Biobank HbA1c measurement. Indications of a diabetes diagnosis were defined as:

- Diabetes diagnosis code in linked hospital episode statistics (HES) records with an episode start date prior to or on the same day as baseline assessment (admission date used where episode start missing)
- Diabetes code, HbA1c measurement ≥ 48 mmol/mol, or prescription for a glucose-lowering medication/glucagon/glucose testing strips in linked primary care records with a date prior to or on the same day as baseline assessment

- Self-reported diabetes, diabetes-specific complications, or prescription for glucose-lowering medications at first UK Biobank assessment centre visit (baseline assessment), or self-reported diabetes at second or third assessment centre visit where the diagnosis was before or in the same year as first assessment centre visit

Details of HES codes, primary care codes and UK Biobank self-reported variables used can be found in the Appendix of this document.

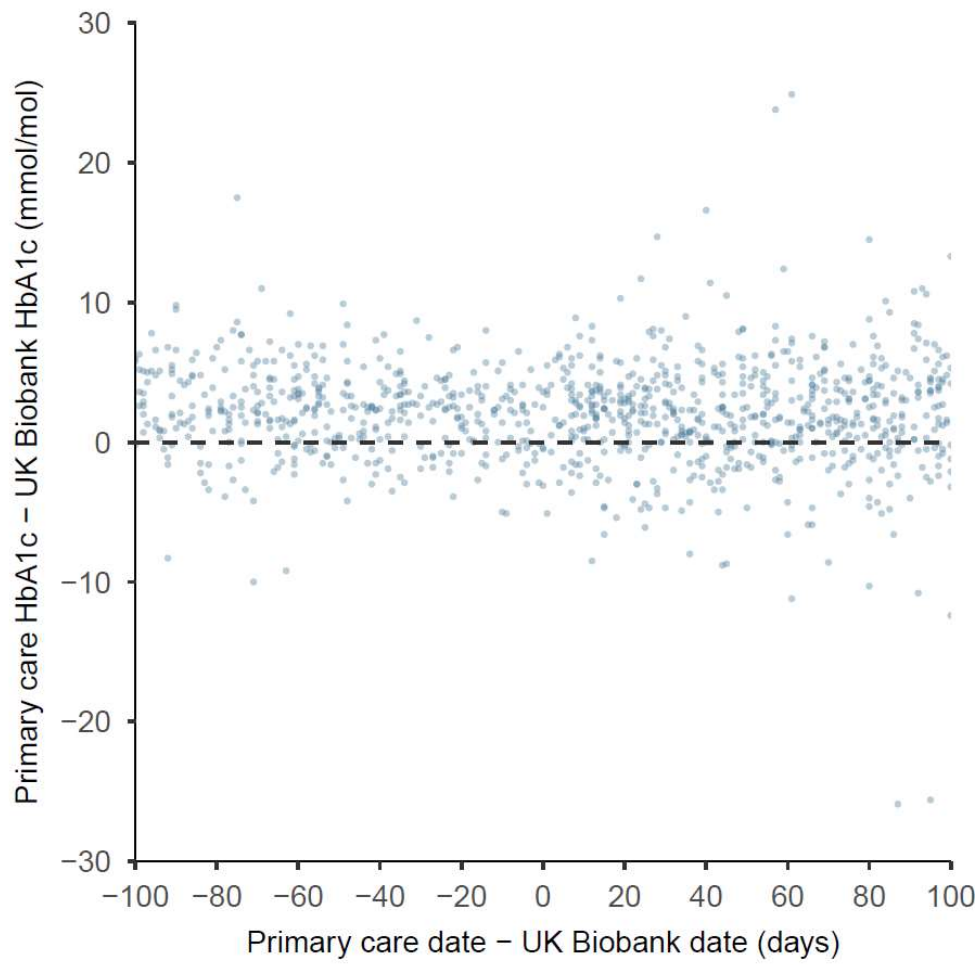
^c Participants with more than one primary care HbA1c measurement within 100 days of baseline assessment were excluded as this close monitoring of HbA1c by their doctor may indicate that the participant was experiencing health issues affecting HbA1c or taking treatment expected to change their HbA1c. This did not include duplicate entries for a single HbA1c measurement (e.g. the same measurement recorded in both percentage and mmol/mol units) where these had the same event date.

^d The Pathology Bounded Code List (PBCL) defines codes used in electronic reporting from pathology labs to general practitioners (GPs). Patient biomarker measurements with these codes may be more reliable than measurements with non-PBCL codes, as the latter are more likely to have been inputted manually by a GP. In order to include only reliable primary care HbA1c measurements in our analysis, we selected measurements with Read 2 and Read CTV3 PBCL codes specifically for HbA1c (not general 'glycosylated haemoglobin' codes or HbA1 codes) from the most recent PBCL release (October 2017; available via the NHS Digital TRUD website). The four codes selected are shown in Supplementary Table S1.

UK Biobank linked primary care data does not reliably include the units of biomarker measurements. We used the PBCL code description (Supplementary Table S1) to infer HbA1c measurement units (DCCT aligned = percentage units; IFCC aligned = mmol/mol units⁸), and then excluded measurements with a value outside of the clinically plausible limits for that unit⁹.

Primary care HbA1c measurements in percentage were converted to mmol/mol using the NGSP/IFCC equation⁸.

Supplementary Figure S2: Scatter plot of time difference between blood draws for primary care HbA1c and UK Biobank HbA1c measurements vs difference in value of these HbA1c measurements (n=1,039)



Pearson correlation coefficient = -0.025, $P = 0.4$

Supplementary Table S1: Pathology Bounded Code List (PBCL) codes used to identify reliable primary care HbA1c measurements. Measurements with values outside of the plausible limits were excluded.

PBCL code	Read 2 / Read CTV3	Code description	Unit^a	Plausible limits of measurement value based on unit³
42W4.	Read 2	HbA1c level (DCCT aligned)	percentage	≥ 3.9 and ≤ 20
42W5.	Read 2	HbA1c level (IFCC standardised)	mmol/mol	> 20 and ≤ 195
XaERp	Read CTV3	HbA1c level (DCCT aligned)	percentage	≥ 3.9 and ≤ 20
XaPbt	Read CTV3	HbA1c level (IFCC standardised)	mmol/mol	> 20 and ≤ 195

^a inferred from code descriptions: DCCT aligned = percentage units; IFCC aligned = mmol/mol units⁸

Supplementary Table S2: Discussion of possible methodological differences between primary care and UK Biobank HbA1c measurements and their potential contributions to the differences observed in measurement values.

Note: UK Biobank HbA1c measurements used venous blood samples. We expect the vast majority of HbA1c measurements in the primary care records to also be venous samples; although point-of-care testing (POCT) of capillary samples is available in some general practices, it is not recommended for diagnosis of diabetes¹⁰. Since our study population consisted of participants without a diagnosis of diabetes, we assume that the reason for their having an HbA1c test in primary care was screening for diabetes, and so POCT would not have been used. In addition, we have only used HbA1c measurements with Pathology Bounded Code List (PBCL) Read codes (Supplementary Figure S1), which are primarily from NHS laboratories, which only deal with HbA1c tests on venous samples.

Possible methodological difference	UK Biobank method	Primary care (NHS) method	Expected impact of difference	Contribution to observed differences ^a
Storage of blood samples prior to HbA1c measurement	Whole blood sample in EDTA collection tubes stored at 4°C for < 36 hours and then in liquid nitrogen ⁴ for 4-10 years (median time for whole cohort of 482,331 with UK Biobank HbA1c measurement = 6.7 years) ^b	Whole blood sample in EDTA collection tubes stored at ambient temperatures for ≤ 72 hours and 4°C for ≤ 14 days ¹¹	Several studies have looked at the impact of freezing whole blood samples at -70 - -80°C on HbA1c measurement. Of the studies which stored samples for at least 12 months, both reductions ¹²⁻¹⁴ and increases ¹³ in HbA1c of 0.1-3.3 mmol/mol have been reported. The reason for these changes is unclear.	The small reductions in HbA1c value after storage at -70 to -80°C observed in previous studies ¹²⁻¹⁴ are comparable to the difference between UK Biobank and primary care measurements observed in this study (2.1 mmol/mol). Different storage methods could therefore feasibly account for this difference.
HbA1c analyser used	Five Bio-Rad Variant II Turbo HPLC analysers ³	NHS laboratories use a range of analysers similar to that used by UK Biobank	External Quality Assurance (EQA) data from 2019-2020 ¹⁵ show that Bio-Rad Variant II analysers, on average, obtain <u>higher</u> readings than the majority of analysers tested in the same EQA cycle, including the Arkray/Adams/ Menarini A1c HA-8000 and TOSOH HLC723/G7/G8/GX, which were the two most popular analysers.	Given the EQA results, we would expect the Bio-Rad Variant II Turbo used by UK Biobank to give higher readings than those seen in primary care, the opposite to what was observed in this study. We therefore do not think that analyser differences contributed to the observed differences.

			Different analysers are affected differently by haemoglobin variants ¹⁶ .	The prevalence of haemoglobin variants in the UK population is relatively low (<5% ¹⁷) and so is very unlikely to account for the differences observed in this study.
User differences (person performing HbA1c assays)	Insufficient information	Insufficient information	Could result in higher or lower readings	Insufficient information
Quality assurance/quality control protocols used	Registered with UK NEQAS EQA scheme ³ . ISO 17025:2005 Quality Accreditation ³ .	NHS laboratories are required to be registered with an EQA scheme such as NEQAS or RIQAS.	No difference expected	No difference expected
	Used algorithms to ensure that the order in which samples were tested was random in regards to geography, dates or time of day that the sample was collected ⁴ . This allowed identification and correction of day-to-day assay variation (correction was based on average assay values for each day) ⁴ .	Corrections for assay drift based on average assay result are not part of the routine quality control protocols used in NHS laboratories.	The date-of-assay corrections applied by UK Biobank were mostly very small (IQR -0.9 to +0.9 mmol/mol, median 0.0 mmol/mol). They had the overall effect of reducing higher values and increasing lower values, lowering the overall variability.	The UK Biobank date-of-assay corrections do not explain the difference between the UK Biobank and primary care HbA1c measurements. In the cohort used in this study ($n=1,039$), the raw (uncorrected) UK Biobank HbA1c measurements are on average 2.1 mmol/mol lower than the primary care HbA1c measurements (very similar to the 2.1 mmol/mol difference observed between the corrected values and the primary care measurements) ^f .

^a UK Biobank HbA1c measurements were lower by an average of 2.1 mmol/mol than those in primary care for a cohort for $n=1,039$ participants with a primary care HbA1c within 100 days of baseline assessment at UK Biobank.

^b Uncorrected UK Biobank HbA1c measurements and dates of assays were from the Biomarker Assay Extended Dataset (Return 1602) provided by UK Biobank

Appendix: Diabetes mellitus codes and variables

Hospital Episode Statistics (HES) and primary care codes

All HES and primary care codes, together with notes on how they are implemented are available from the following public Github repository: https://github.com/drkgyoung/UK_Biobank_codelists.

Self-reported UK Biobank fields

Field	Field description	Which assessment centre visit	Values used to indicate a diagnosis of diabetes prior to/at first assessment centre
2443	Diabetes diagnosed by a doctor	First	Any value other than 0 No (1 Yes, -1 Do not know, -3 Prefer not to answer, or missing) was treated as a potential indication of a diagnosis of diabetes
6148	Eye problems/ disorders	First	1 Diabetes related eye disease
6153	Medication for cholesterol, blood pressure, diabetes, or take exogenous hormones	First	3 Insulin
6177	Medication for cholesterol, blood pressure or diabetes	First	3 Insulin
20002	Non-cancer illness code, self-reported	First	1220 diabetes 1221 gestational diabetes 1222 type 1 diabetes 1223 type 2 diabetes 1276 diabetic eye disease 1468 diabetic neuropathy/ulcers 1607 diabetic nephropathy
20002	Non-cancer illness code, self-reported	Second and third	Any of the values shown above (1220/1221/1222/1223/1276/1468/1607) where year of diagnosis was prior to or the same as the year of the first assessment visit
20003	Treatment/ medication code	First	1140857494 glibornuride 1140857496 glutral 25mg tablet 1140857500 glymidine 1140857502 gondafon 500mg tablet 1140857506 pramidex 500mg tablet 1140857584 acetohexamide 1140857586 dimelor 500mg tablet 1140857590 libanil 2.5mg tablet 1140874646 glipizide 1140874650 glibenese 5mg tablet 1140874652 minodiab 2.5mg tablet 1140874658 gliquidone 1140874660 glurenorm 30mg tablet 1140874664 tolazamide 1140874666 tolanase 100mg tablet 1140874674 tolbutamide 1140874678 glyconon 500mg tablet 1140874680 rastinon 500mg tablet 1140874686 glucophage 500mg tablet 1140874690 orabet 500mg tablet 1140874706 chlorpropamide 1140874712 diabinese 100mg tablet 1140874716 glymese 250mg tablet 1140874718 glibenclamide

		1140874724 daonil 5mg tablet 1140874726 semi-daonil 2.5mg tablet 1140874728 euglucon 2.5mg tablet 1140874732 malix 2.5mg tablet 1140874736 diabetamide 2.5mg tablet 1140874740 calabren 2.5mg tablet 1140874744 gliclazide 1140874746 diamicon 80mg tablet 1140883066 insulin product 1140884600 metformin 1140921964 glucamet 500 tablet 1141152590 glimepiride 1141153254 troglitazone 1141153262 romozin 200mg tablet 1141156984 amaryl 1mg tablet 1141157284 glipizide product 1141168660 repaglinide 1141168668 novonorm 0.5mg tablet 1141169504 diaglyk 80mg tablet 1141171508 vivazide 80mg tablet 1141171646 pioglitazone 1141171652 actos 15mg tablet 1141173786 starlix 60mg tablet 1141173882 nateglinide 1141177600 rosiglitazone 1141177606 avandia 4mg tablet 1141189090 rosiglitazone 1mg / metformin 500mg tablet 1141189094 avandamet 1mg / 500mg tablet
--	--	---

References

1. The EurA1c Trial Group. 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. *Clinical Chemistry* 2018; **64**: 1183-92.
2. Wiedmeyer H-M, Polonsky KS, Myers GL, et al. International Comparison of C-Peptide Measurements. *Clinical Chemistry* 2007; **53**: 784-7.
3. Tierney A, Fry D, Almond R, Gordon M, Moffat S. UK Biobank Biomarker Enhancement Project – Companion Document to Accompany HbA1c Biomarker Data V1.0. 2018. https://biobank.ndph.ox.ac.uk/showcase/showcase/docs/serum_hb1ac.pdf. (3 Aug 2021, date last accessed).
4. UK Biobank. Biomarker assay quality procedures: approaches used to minimise systematic and random errors (and the wider epidemiological implications) V1.2. 2019. https://biobank.ctsu.ox.ac.uk/crystal/crystal/docs/biomarker_issues.pdf. (3 Aug 2021, date last accessed).
5. Tavaglione F, De Vincentis A, Jamialahmadi O, et al. Inborn and acquired risk factors for severe liver disease in Europeans with type 2 diabetes from the UK Biobank. *JHEP Reports* 2021; **3**: 100262.
6. Cull CA, Manley SE, Stratton IM, et al. Approach to maintaining comparability of biochemical data during long-term clinical trials. *Clinical Chemistry* 1997; **43**: 1913-8.
7. Mc Cord KA, Al-Shahi Salman R, Treweek S, et al. Routinely collected data for randomized trials: promises, barriers, and implications. *Trials* 2018; **19**: 1-9.
8. Zhelev Z, Peters J, Rogers M, Andrews R, McDonald T, Hyde C. Accuracy and validity of HbA1c Point of Care Testing: A review of the scientific evidence and guidelines. 2020. <https://www.healthcheck.nhs.uk/commissioners-and-providers/evidence/>. (3 Aug 2021, date last accessed).
9. Rodgers LR, Weedon MN, Henley WE, Hattersley AT, Shields BM. Cohort profile for the MASTERMIND study: using the Clinical Practice Research Datalink (CPRD) to investigate stratification of response to treatment in patients with type 2 diabetes. *BMJ Open* 2017; **7**: e017989.
10. National Institute for Health and Care Excellence (NICE). Type 2 diabetes: prevention in people at high risk. 1.3 Risk identification (stage 1). 2012. <https://www.nice.org.uk/guidance/ph38/chapter/Recommendations#risk-identification-stage-1>. (3 Aug 2021, date last accessed).
11. York and Scarborough Teaching Hospitals NHS Foundation Trust. Laboratory Medicine Test Directory / HbA1c. 2021. <https://www.yorkhospitals.nhs.uk/our-services/a-z-of-services/lab-med/test-directory/clinical-biochemistry/hba1c/?rit>. (15 October 2021, date last accessed).
12. Liotta L, Di Franco A, Pazzagli M, Luconi M. Glycated hemoglobin (HbA1c) measurement in frozen whole blood depends on baseline values of fresh samples. *Analytical and Bioanalytical Chemistry* 2013; **405**: 429-34.
13. Rolandsson O, Marklund SL, Norberg M, Ågren Å, Hägg E. Hemoglobin A1c can be analyzed in blood kept frozen at 0°C and is not commonly affected by hemolysis in the general population. *Metabolism - Clinical and Experimental* 2004; **53**: 1496-9.
14. Jones W, Scott J, Leary S, et al. Stability of Whole Blood at -70°C for Measurement of Hemoglobin A1c in Healthy Individuals. *Clinical Chemistry* 2004; **50**: 2460-1.
15. RIQAS. Monthly reports provided to Royal Devon & Exeter NHS Trust Hospital as part of the RIQAS Glycated Haemoglobin (HbA1c) EQA programme.
16. National Glycohemoglobin Standardization Program (NGSP). Factors that Interfere with HbA1c Test Results. 2019. <http://www.ngsp.org/factors.asp>. (3 Aug 2021, date last accessed).
17. Heaton C, Witt M, Cole L, et al. Detection and mapping of haemoglobin variants in blood fingermarks by MALDI MS for suspect “profiling”. *Analyst* 2021; **146**: 4290-302.