METHODOLOGY

Effects of the UK Biobank collection protocol on potential biomarkers in saliva

Rashida Pramanik,¹,² Hayley Thompson,³ James O Kistler,³ William G Wade,³ John Galloway,⁴ Tim Peakman⁵ and Gordon B Proctor²*

¹NIHR Biomedical Research Centre, Guy’s & St Thomas’ Hospital NHS Foundation Trust, London, UK, ²Salivary Research Unit, Department of Clinical and Diagnostic Sciences, King’s College London Dental Institute, London, UK, ³Microbiology Unit, Department of Clinical and Diagnostic Sciences, King’s College London Dental Institute, London, UK, ⁴Dental Teams Study Unit, Joint UCL/UCLH Biomedical Research Centre, University College London Hospital NHS Foundation Trust, London, UK and ⁵UK Biobank, Stockport, UK

*Corresponding author. Salivary Research Unit, Clinical and Diagnostic Sciences, Fl 17 Guy’s Tower, King’s College London Dental Institute, London SE5 9NU, UK. E-mail: gordon.proctor@kcl.ac.uk

Accepted 23 August 2012

Background The UK Biobank (UKB) is a national epidemiological study of the health of 500,000 people, aged 40–69 years, who completed health-related tests and a questionnaire and gave samples of blood and urine. Salivas collected from 120,000 of these subjects were transported at 4°C and were placed in ultra-low temperature archives at up to 24 h after collection. The present study assessed how changes in saliva composition under UKB conditions influence a range of potential biomarkers resulting from holding saliva at 4°C for 24 h.

Methods Unstimulated whole-mouth saliva samples were collected from 23 volunteers aged 45–69 years. Salivas were split into aliquots some of which were immediately frozen at −80°C, whereas others were stored at 4°C for 24 h and then frozen at −80°C, mimicking the UKB protocol.

Results Assessment of mRNA by real-timepolymerase chain reaction revealed no difference between samples that were analysed after the UKB protocol and those that were immediately preserved. Immunochemical analysis showed some loss of β-Actin under UKB conditions, whereas other salivary proteins including cytokines and C-reactive protein appeared to be unaffected. Cortisol and showed no reduction by UKB conditions, but salivary nitrite was reduced by 30%. The oral microbiome, as revealed by sequencing 16S rRNA genes, showed variations between subjects, but paired samples within subjects were very similar.

Conclusions Our results suggest that many salivary components remain little affected under UKB collection and handling protocols, suggesting that the resource of 120,000 samples held in storage will be useful for phenotyping subjects and revealing potential prognostic disease biomarkers.

Keywords Biological markers, body fluid, oral diagnosis, hormones, cytokines, DNA, RNA
Introduction

The routine collection and storage, or biobanking, of saliva, blood and urine is a key element in using biofluids for diagnostic purposes allowing the long-term assessment of biofluids from a large population. The UK Biobank (UKB) is a national epidemiological study in which questionnaire data, blood and urine have been collected from 500 000 people, aged 40–69 years, for approximately a 3.5-year period until July 2010. In the past 12 months of the study, a number of enhancements to subject phenotyping were integrated into the existing UKB protocol, including the collection of saliva, and resulted in the banking of salivas from almost 120 000 subjects. The database of health and lifestyle questionnaire responses and health screening test results along with the bank of biofluids collectively represent a resource for the study of the links between disease development, genetic predisposition and lifestyle. To complete the study within the proposed timeframe, each UKB clinic recruited up to 120 subjects per day. The high throughput of subjects in UKB was achieved by adopting a strategy of installing temporary clinics in different UK cities for the collection and processing of samples from volunteers. The collected samples underwent minimal processing at clinics; most blood samples, urine and saliva samples were stored at 4°C in the clinic and transported at 4°C to a high throughput sample processing centre where they were aliquoted and stored in ultra-low temperature archives. The processing of samples from collection to final storage was completed by no longer than 24 h.

One of the main reasons for the inclusion of saliva in the UKB study is the recognition of its utility as a biofluid with obvious potential for personalized healthcare applications such as monitoring health status, disease onset and progression and treatment outcome, as it can be collected simply and non-invasively. Salivary constituents are derived from glandular salivas, oral surfaces and gingival crevicular fluid, which can be described as a plasma-like exudate containing systemic components. Saliva therefore contains potential biomarkers of both oral and systemic disease; including host cells and microorganisms, a large number of proteins—over 2000—have been catalogued in the salivary proteome; cellular mRNA, short mRNA and miRNA molecules present in exosomes; hormones and metabolites. Salivary disease biomarkers have been identified for oral cancer, acute myocardial infarction, Sjögren syndrome and HIV and other infectious diseases; saliva can be used as a source of DNA for genotyping studies.

However, like other biofluids, saliva is subject to compositional change after collection, and for this reason, it has been recommended that it be frozen immediately at −80°C and/or centrifuged to remove the cellular compartment before freezing, as it is recognized that oral bacteria can metabolize constituents. Given the potential degradation of salivary components by bacteria and enzymes, the present study examined a number of salivary biomarker molecules (see Table 1) and aimed to determine the extent of change resulting from the collection and transportation conditions to which salivas were subjected in the UKB study. It is envisaged that such data will enable potential users of UKB saliva samples to make an informed decision as to the likely utility of the samples to their research questions.

Materials and methods

Subjects and saliva collection

The study was approved by the Human Research Ethics (South East London REC 1) Committee as ethics number 10/H0804/046. Saliva samples were collected at King’s College London from 23 volunteers aged between 45 and 69 years. The study population was composed of 11 males and 12 females, with an average age of 54.2 ± 1.4 years.

Unstimulated whole-mouth saliva was collected between 10 am and noon or 2 pm and 4 pm by passive drool into a sterile universal tube to a final volume of 5 ml. The duration of the collection was recorded, and salivary flow rate was calculated. Saliva samples were immediately stored on ice at 4°C, and 0.5 ml aliquots were prepared in 1.5 ml polypropylene tubes and split into two batches. One batch was frozen at −80°C, and the remaining batch was initially stored at 4°C for 24 h and then frozen at −80°C (UKB protocol). A range of analyses were performed on both sets of saliva samples. An additional aliquot of saliva was preserved with ribonuclease inhibitor SuperaseIn (Ambion, Invitrogen) to serve as a control for quantification of mRNA transcripts. SuperaseIn was added to 0.5 ml of saliva to a final concentration of 100 U and immediately frozen at −80°C.

Analysis of salivas

A summary of the analyses undertaken is given in Table 1, and brief details are given below. Full descriptions of the methods used are given in the Supplementary Materials and methods document (available as Supplementary data at IJE online).

RNA and DNA isolation and assessment

Stored whole-mouth saliva was thawed on ice, vortexed and homogenized, and for isolation of RNA, 0.5 ml saliva was used, and for DNA, a separate 0.1 ml aliquot. RNA and DNA in samples were bound to different silica membrane columns. Salivary RNA was eluted from columns with 50 μl nuclese free water, and a small aliquot was quantified by Ribogreen fluorescent dye-binding. DNA was eluted from columns in 100 μl buffer, and a small
aliquot was quantified with Picogreen fluorescence assay. The quality of isolated total RNA was assessed by electrophoresis through a polymer gel and separation according to size of RNA molecule, using a Bioanalyzer system. One microlitre of isolated RNA from saliva was heat denatured at 70°C for 2 min for electrophoresis, and separated RNA was compared with a standard RNA ladder. RNA extracted from oral epithelial (TR146) cells was compared with salivary RNA.

Expression of mRNA gene products was undertaken by thermal cycling after reverse transcription of mRNA sequences between specific primers [real-time polymerase chain reaction (RT-PCR)] and measurement of SYBR green dye fluorescence performed according to manufacturer’s instructions. All RT-PCR reactions were performed in duplicate, and an efficient run was confirmed by a single peak at the correct melt temperature on the melt analysis. This was also confirmed by agarose gel electrophoresis.

Salivary microbiome profiling
Saliva samples (0.5 ml) were centrifuged at 13 000 g for 5 min and resuspended in a lysozyme solution (200 µl, 45 mg/ml). Bacterial genomic DNA was extracted using a protocol optimized for the extraction from Gram-positive bacteria, according to the manufacturer’s instructions; 16S rRNA genes were amplified by PCR using specific primer together with unique ‘barcode’ sequences, as described by the Human Microbiome Project (http://www.hmpdacc.org/). The amplicons were purified using a silica spin column, checked for size and purity by electrophoresis and quantified by Picogreen fluorescent dye-binding. Emulsion polymerase chain reaction (emPCR) and unidirectional sequencing was performed using the Lib-L kit and Roche 454 GS-FLX Titanium sequencer by the Centre for Haemato-Oncology, Bart’s Cancer Institute, Queen Mary University of London, London, UK.

Gel electrophoresis of proteins
Saliva samples were thawed on ice and prepared for sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS-PAGE) under reducing conditions. The equivalent of 10 µl of the heat-denatured reduced saliva samples was applied to each lane of precast 4–12% gradient Bis-Tris gels (Invitrogen, Paisley, UK). Proteins in gels were stained with 0.2% Coomassie Brilliant Blue R250, and glycoproteins in gels were stained with periodic acid Schiff (PAS).

Some electrophoresis gels were not stained but instead electroblotted onto nitrocellulose membranes for probing with specific antibodies to mucin (MUC5B), β-Actin or cystatin S. MUC5B was detected with a mouse monoclonal antibody EU-MUC5Bb (EU Consortium, gift of Prof. Dallas Swallow, University College London, UK)12; β-Actin was detected with mouse monoclonal anti-β-Actin (Sigma) diluted 1:250, and Cystatin S with mouse monoclonal anti-Cystatin S (R&D Systems, Abingdon, UK) diluted 1:250. Positive bands were detected using peroxidase-labelled secondary antibodies and followed by a chemiluminescent substrate in the presence of hydrogen peroxide.

Salivary components assayed

<table>
<thead>
<tr>
<th>Salivary component</th>
<th>Clinical significance</th>
<th>Analytical method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall protein composition</td>
<td>Salivary proteins are significant in oral and systemic physiology and disease</td>
<td>SDS electrophoresis western blotting</td>
</tr>
<tr>
<td>Cytokine profile—IL-1α and β, –2, –4, –6, –8, –10, IFN-γ, MCP-1, TNF-α, VEGF and EGF</td>
<td>Mucosal and systemic inflammation</td>
<td>Multiplex biochip immunoassay</td>
</tr>
<tr>
<td>Cortisol and DHEA (dihydroepiandrosterone)</td>
<td>These and other steroid hormones provide markers of stress, menopause and so forth.</td>
<td>ELISA immunoassay</td>
</tr>
<tr>
<td>C-reactive protein</td>
<td>Mucosal and systemic inflammation including diabetes and cardiovascular disease</td>
<td>ELISA immunoassay</td>
</tr>
<tr>
<td>Nitrite</td>
<td>Dietary intake of fresh vegetables and marker of cardiovascular health</td>
<td>Griess reaction</td>
</tr>
<tr>
<td>Bacterial metagenomics</td>
<td>Mucosal health and disease</td>
<td>454 sequencing of 16S rRNA gene (Roche platform)</td>
</tr>
<tr>
<td>RNA—total and integrity mRNA-selected transcripts -β-Actin, IL-8 and GAPDH</td>
<td>Mucosal and systemic disease</td>
<td>Ribogreen, Bioanalyzer and RT-qPCR</td>
</tr>
</tbody>
</table>

Table 1 Summary of salivary components assayed

<table>
<thead>
<tr>
<th>Clinical significance</th>
<th>Analytical method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Salivary protein profile</td>
<td>SDS electrophoresis western blotting</td>
</tr>
<tr>
<td>Cytokine profile—IL-1α and β, –2, –4, –6, –8, –10, IFN-γ, MCP-1, TNF-α, VEGF and EGF</td>
<td>Multiplex biochip immunoassay</td>
</tr>
<tr>
<td>Cortisol and DHEA (dihydroepiandrosterone)</td>
<td>ELISA immunoassay</td>
</tr>
<tr>
<td>C-reactive protein</td>
<td>ELISA immunoassay</td>
</tr>
<tr>
<td>Nitrite</td>
<td>Griess reaction</td>
</tr>
<tr>
<td>Bacterial metagenomics</td>
<td>454 sequencing of 16S rRNA gene (Roche platform)</td>
</tr>
<tr>
<td>RNA—total and integrity mRNA-selected transcripts -β-Actin, IL-8 and GAPDH</td>
<td>Ribogreen, Bioanalyzer and RT-qPCR</td>
</tr>
</tbody>
</table>

Enzyme-linked immunosorbant assays for steroid hormones and C-reactive protein
Salivary cortisol, dehydroepiandrosterone (DHEA) and C-reactive protein (CRP) were determined by enzyme-linked immunosorbant assay (ELISA) with commercial kits (Salimetrics Europe, UK), according
to manufacturer’s instructions. Saliva samples were thawed, vortexed and then centrifuged for 15 min at 3000 r.p.m. For cortisol and DHEA ELISA, salivas were used neat, and for the CRP, ELISA was diluted (1:10) in a phosphate-buffered solution. Saliva samples, as well as hormone standards and controls (25–50 µl), were then transferred onto 96-well microtitre plates pre-coated with anti-CRP, anti-DHEA or anti-cortisol capture antibodies. All unknown samples, standards and controls were run in duplicate. A detecting antibody-enzyme (horseradish peroxidase) conjugate specific to each hormone was diluted in phosphate buffer and added to each well and binding determined using the substrate tetramethylbenzidine in the presence of hydrogen peroxide. A standard curve was generated with purified antigens, and the concentration of unknowns and controls was calculated and expressed at the appropriate concentration.

Cytokine and peptide growth factor analysis
A high sensitivity multiplex assay, which used Biochip Array Technology (BAT; Randox Laboratories Limited, County Antrim, UK), incorporating capture and detecting antibodies was used to determine the levels of 10 cytokines: interleukin (IL)-2, IL-4, IL-6, IL-8, IL-10, epidermal growth factor (EGF), interferon (IFN) gamma, monocyte chemoattractant protein 1 (MCP-1), tumour necrosis factor (TNF) alpha and vascular endothelial growth factor (VEGF). Saliva samples were thawed on ice, vortexed and clarified by centrifugation at 3000 r.p.m. for 15 min. Neat saliva samples (100 µl) were applied to a high sensitivity biochip, and simultaneous quantification was performed on the Evidence Investigator, a semi-automated bench-top analyser. Cytokine and growth factor standards were used to create a standard curve, and concentrations were expressed as pg/ml. The performance of the assay in saliva was assessed by assaying known amounts of high and low concentration standards for each target added to two control saliva samples.

Nitrite analysis
Nitrite was colorimetrically determined with the Greiss reagent system (Sigma, UK) using thawed salivas that had been refrozen after a previous analysis. A nitrite standard reference curve was prepared ranging from 0 to 100 µM. All samples and standards were assayed in duplicate and the average absorbance and salivary nitrite concentration calculated. To determine whether salivary proteins interfered with the Greiss assay, deproteination was performed with alkaline zinc sulphate (5 µl of 1 M sodium hydroxide and 3 µl of 0.5 M zinc sulphate were added to 50 µl saliva) and results compared with whole saliva. Salivas were also ‘spiked’ with sodium nitrite to give a final concentration of 5 µM nitrite.

Statistical analysis
The GraphPad suite of computer programmes, StatMate, InStat and Prism, were used to perform Power calculations and analyse and present data. Power calculations used to determine the sample size were based on using paired t-tests and published data along with the authors’ pilot data for CRP and nitrite. For CRP, mean values of 2500 pg/ml were assumed with a standard deviation of 2000 and a correlation coefficient (R) of 0.8 between the paired data. Twenty-three data pairs would give 80% power to identify a difference of ~550 pg/ml at P = 0.05. For nitrite, mean values of 100 µM were assumed with a standard deviation of 90 and R = 0.8. Twenty-three data pairs would give 80% power to identify a difference of ~35 µM at P = 0.05.

Data sets were tested for Gaussian distribution using a Kolmogorov–Smirnov test, and paired Student t-test was used to compare data from salivas collected using the UKB protocol with immediately frozen salivas. Non-parametric Mann–Whitney tests were used to test for differences between samples when the variable of interest was not normally distributed. Cycle threshold (Ct) values of RT-PCR data from three groups were compared by one-way ANOVA.

Sequence analysis and identification of microbial species was performed using Parsimony and Unifrac, which are part of the Mothur suite of programs. Sequences were filtered for quality and chimeric sequences removed and identified with reference to the Silva database. Analyses included descriptive statistics of diversity and species, richness, phylogenetic identification of the bacterial taxa making up the samples and statistical comparisons of community composition.

Results
DNA and RNA expression
The total yield of DNA varied between subjects ranging from 334 to 2077 ng per 0.1 ml aliquot of saliva. The mean yield of DNA in UKB protocol saliva samples was 9.54 ± 1.61 ng for each microlitre of saliva and in frozen saliva samples was 9.30 ± 1.20 ng. The stability of host DNA in saliva samples was assessed by quantitative polymerase chain reaction (qPCR) using a primer for a chromosome 18 sequence. The Ct values for UKB protocol and frozen salivas were 23.4 ± 0.2 and 23.1 ± 0.2, respectively.

RNA concentrations of frozen and UKB protocol saliva samples were calculated and compared for all samples. On average, 3.1 ± 0.42 ng of total RNA was obtained per microlitre of saliva stored under UKB protocol conditions, 2.9 ± 0.4 ng from one microlitre of saliva frozen immediately after collection and 2.58 ± 0.41 ng from SUPERase-In™ preserved saliva. The total yield of RNA varied between volunteers ranging from 250.5 to 6966.1 ng per 0.5 ml aliquot of saliva.

The stability of salivary RNA was determined by RT-qPCR, and all saliva samples contained mRNAs
that encode for housekeeping genes glyceraldehyde 3-phosphate dehydrogenase (GAPDH), ribosomal protein S9 (RPS9) and β-Actin. The relative levels of expression (mean average cycle threshold Ct values) between UKB protocol and frozen saliva samples were similar (Figure 1a). Saliva samples treated with the ribonuclease inhibitor Superaseln before storage showed no significant differences in Ct values of all three housekeeping genes compared with freezing and UKB protocol saliva samples. In all samples, GAPDH showed the highest expression (lowest Ct value), followed by β-Actin and then RPS9 (highest Ct value). Five other genes of interest IL-1, IL-8, spermidine acetyltransferase (SAT1), S100A and S100P were also detectable in all saliva samples, and each had varying levels of expression. However, there was little difference between the Ct values of the frozen, UKB protocol or inhibitor-treated saliva samples for each gene. S100A and IL-8 were highly expressed in all three saliva groups with mean Ct values between 20 and 25. The genes IL-1, SAT1 and S100P were expressed at low levels in all saliva groups with mean Ct values exceeding 28 cycles.

The sizes and distribution of RNA from both UKB protocol and immediately frozen saliva samples are shown as a gel in Figure 1b. They were compared with salivas preserved with Superaseln and RNA extracted from human TR146 cells as a reference. In general, RNA was present in all saliva samples and the TR146 cell line. There appeared to be little difference in the quality of RNA between salivas frozen immediately, stored under UKB conditions or preserved with Superaseln. A visual inspection of the electrophoretic profile for frozen, UKB protocol and Superaseln saliva samples showed similarities, and two bands were present, which appeared to correspond to 18S and 28S ribosomal RNA. However, the large number of sharp bands and high background are owing to the presence of bacterial 23S, 16S and 5S RNA along with some RNA degradation. This resulted in low RNA integrity numbers (RINs) of 2.5 and 2.8 for frozen and UKB protocol-collected saliva, respectively. As expected, RNA extracted from TR146 cells was intact, of high quality and exhibited no degradation and two distinct 18S and 28S bands and a high RIN of 9.3 (Figure 1b).

Salivary protein composition and levels of cytokines and peptide growth factors

Staining of SDS-PAGE gels of samples from 11 subjects revealed similar patterns under UKB conditions compared with immediate freezing (Figure 2a; Supplementary Figure S1A, available as Supplementary data at IJE online). Five out of 11 samples showed metachromatically stained proline-rich proteins, and these were less intensely stained in samples handled by the UKB protocol. Similar staining patterns were obtained with acetone-precipitated saliva samples after isolation of mRNA (data not shown). Western blotting detected distinct β-Actin bands in eight samples, and in five of these, the staining was visibly reduced in UKB saliva samples compared with frozen samples (Figure 2b; Supplementary Figure S1B, available as Supplementary data at IJE online). Cystatin S was detected in all samples, and there was no obvious difference between the UKB protocol and immediate freezing (Figure 2b). A prominent high molecular weight band was demonstrated by PAS staining of gels (Supplementary Figure 1C, available as Supplementary data at IJE online), and the position of this band corresponded to that of MUC5B in the western blots (Figure 2b; Supplementary Figure S1D, available as Supplementary data at IJE online). In both PAS staining and western blots, the MUC5B band appeared not be affected by the processing conditions.

Cytokines and growth factors were assayed using multiplex BAT and showed no statistically significant difference between UKB protocol and frozen saliva samples. EGF and IL-8 were present at levels above the detection limit of the high sensitivity assay used (Table 1). Components of saliva did not appear to interfere with the BAT assays, as those salivas to which known amounts of cytokine or growth factor had been added gave the expected results, i.e. a concentration of target molecule reflecting the saliva concentration and the added standard concentration (Table 2).

Levels of steroid hormones, C-reactive protein and nitrite

Concentrations of cortisol in saliva were higher in UKB protocol saliva samples compared with immediately frozen samples, whereas DHEA showed little difference with different handling/storage (Table 3). Salivary concentrations of CRP were also similar in all samples with median values of ~1734 and 1494 pg/ml under UKB conditions and immediate freezing, respectively, with some subjects showing substantially higher values (Supplementary Figure S2, available as Supplementary data at IJE online).

Overall nitrite concentrations were reduced by 30% in UKB protocol saliva samples compared with immediately frozen samples (Table 3). However, there was great individual variation in the amounts of nitrite; some subjects showing a small difference between freezing and UKB conditions, but others showing much greater differences (Figure 3). The Greiss reaction appeared to be suitable for use on whole saliva samples, as deproteination of some test saliva samples did not change the measured nitrite concentration, whereas saliva samples spiked with exogenous nitrite gave the predicted readings.

Sequencing of the salivary microbiome

Saliva samples generated 40705 sequences, which were initially screened by selecting sequences with an end position of >10000 in the SILVA 16S rRNA reference alignment14 and a length of >450 bp, leaving 38817 sequences. Use of the uchime algorithm to
detect chimeric sequences detected 373 chimeras, which were removed from the analysis. Sequences were assigned to operational taxonomic units at a distance of 0.015 and identified with reference to the HOMD reference database. The composition of the community found in each sample is shown as Supplementary data (Supplementary Table S2, available as Supplementary data at IJE online). Visual inspection of Supplementary Table S2 (available as Supplementary data at IJE online) shows that the composition of the libraries obtained from the same samples treated in the two different ways was highly similar, and this was confirmed in the dendrogram constructed using the thetayc calculator, which describes the similarity of the samples (Figure 4). It can be seen that the pairs of libraries obtained from each sample always clustered with each other rather than with other samples. Comparisons of the community structure obtained with the two methods of sample handling showed no difference with either the parsimony or weighted unifrac methods (ParsScore 20, \( P = 1 \); UWScore 0.39, \( P = 1 \)).

**Figure 1** Analysis of gene expression in salivas. Salivas collected using the UK Biobank protocol were compared with immediately frozen saliva with and without the RNAase inhibitor SuperaseIn (SI). (a) RT-qPCR of mRNA for GAPDH, \( \beta \)-Actin, RPS9, IL-1, IL-8, SAT1, S100A and S100P. The histogram compares expression quantified as Ct (cycle threshold) values, and there was no statistically significant difference for the salivas handled under different conditions. (b) Gel generated from Agilent 6000 Bioanalyser using RNA Picochip. Representative electrophoretic profiles are as follows: (i) Saliva frozen immediately; (ii) Saliva stored under UKB conditions; (iii) Saliva preserved with SuperaseIn and frozen; and (iv) RNA extracted from cell line TR146. An RNA ladder (L) showing sizes of the nucleotides was used as a reference and RNA integrity number (RIN) is denoted below each electrophoretic profile.
Discussion

The UKB study will follow a cohort of subjects as they age and is intended to help elucidate the factors that predispose to disease development. Significant numbers will accumulate in cohorts for many diseases. For example, 20,000 cases of type 2 diabetes and >10,000 cases of coronary heart disease are predicted to emerge within 10 years. The UKB database and sample repository, therefore, offer a great opportunity for studying genetic associations and phenotypes of disease cohorts. The baseline biofluid samples collected before the appearance of clinical disease can yield prognostic biomarkers.

The potential to assay a wide range of salivary components providing insight into disease development

Table 2 Comparison of cytokine and growth factor concentrations assayed in saliva immediately frozen at −80°C after collection or held at 4°C for 24 h and then frozen at −80°C as in the UKB protocol

<table>
<thead>
<tr>
<th>Cytokine</th>
<th>Freezing (n = 15) (95% CI)</th>
<th>UKB (n = 15) (95% CI)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>IL-2</td>
<td>1.68 ± 0.19 (1.28–2.08)</td>
<td>1.95 ± 0.27 (1.37–2.53)</td>
<td>0.49</td>
</tr>
<tr>
<td>IL-4</td>
<td>2.29 ± 0.17 (1.93–2.08)</td>
<td>2.27 ± 0.14 (1.96–2.58)</td>
<td>0.88</td>
</tr>
<tr>
<td>IL-6</td>
<td>40.4 ± 14.0 (10.3–70.5)</td>
<td>37.1 ± 12.7 (9.9–64.3)</td>
<td>0.4</td>
</tr>
<tr>
<td>IL-8</td>
<td>&gt;854</td>
<td>&gt;854</td>
<td>ND</td>
</tr>
<tr>
<td>IL-10</td>
<td>2.66 ± 0.60 (1.38–3.95)</td>
<td>2.39 ± 0.60 (1.11–3.67)</td>
<td>0.1</td>
</tr>
<tr>
<td>IFN-γ</td>
<td>2.83 ± 0.99 (0.72–4.95)</td>
<td>3.05 ± 0.86 (1.19–4.9)</td>
<td>0.63</td>
</tr>
<tr>
<td>MCP1</td>
<td>604.8 ± 93.5 (404.2–805.3)</td>
<td>552.2 ± 94.4 (349.7–754.6)</td>
<td>0.05</td>
</tr>
<tr>
<td>TNF-α</td>
<td>24.95 ± 4.72 (14.82–35.07)</td>
<td>21.81 ± 3.55 (14.19–29.44)</td>
<td>0.51</td>
</tr>
<tr>
<td>EGF</td>
<td>&gt;723</td>
<td>&gt;723</td>
<td>ND</td>
</tr>
<tr>
<td>VEGF</td>
<td>2089.7 ± 143.7 (1782–2398)</td>
<td>1990.7 ± 148.5 (1672–2309)</td>
<td>0.26</td>
</tr>
</tbody>
</table>

The multiplex assay used an array of antibodies on a biochip (BAT) and concentrations (pg/ml) are expressed as mean ± SEM. The cytokines and growth factors assayed are IL-2, −4, −6, −8, −10, IFN-γ, TNF-α, MCP1, EGF and VEGF.

Figure 2 Sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS-PAGE) and western blot analysis of salivary proteins. Salivas were frozen immediately (F) or stored under UK Biobank (UB) conditions. (a) Coomassie Brilliant Blue-stained proteins of four volunteers. There were no consistent differences in the electrophoretic banding patterns, although metachromatically stained proline-rich proteins, when stained, tended to be decreased in the UKB protocol sample. (b) Western blots of salivas probed with antibodies against (i) MUC5B; (ii) β-Actin; and (iii) Cystatin S, showing bands at >200, 40 and 14 kDa, respectively. The levels of all proteins varied between subjects, but there appeared to be no obvious differences in levels of MUC5B and Cystatin S in samples collected under UKB conditions compared with immediate freezing. In contrast, when β-Actin was detected, it appeared to be reduced under UKB conditions.
presents both an opportunity for studying disease phenotypes and a dilemma for sample handling, as different potential biomarkers may require a variety of methods for optimal stabilization and preservation. Given that whole-mouth saliva contains large numbers of bacteria, the potential for sample degradation is high. However, as with blood and urine, the approach adopted in UKB with saliva was to avoid using specific preserving agents, e.g. protease or ribonuclease inhibitors, as such additives may interfere with the use of chewing or taste stimulation, as these add another variable, and the increased volume of saliva may dilute some important biomarker molecules. In the UKB study, a saliva volume threshold of 2–2.5 ml was requested, but, ideally, the volume of saliva collected should be measured in biomarker studies, as this aids the interpretation of the levels of components.

Assessment of genetic factors determining disease susceptibility will be achieved by gene sequencing of DNA isolated from blood samples collected in the UKB study. It is not anticipated that saliva samples will be required for this purpose, but as in other recent studies, we have found that although the quantity of DNA isolated from saliva is less than that from blood, saliva is a suitable source of host DNA of sufficient quantity and quality for epidemiological studies. Although not assessed in the present study, saliva-derived DNA is proving useful in epigenetic studies. Preserving solutions can be added to saliva to avoid the need for immediate freezing, and DNA remains stable under these conditions for long periods.

Bacterial DNA was also analysed in the present study, as it provides information on bacterial diversity. Bacterial metagenomic studies of biofluids can yield biomarkers of host lifestyle and disease, and it may be that salivary microbiomics can provide biomarkers of oral and systemic disease of relevance to UKB. Sequencing of the 16S rRNA genes by 454 sequencing showed that differences in the diversity of sequences were observed between individuals, but that samples treated by the UKB protocol provided similar data to samples immediately frozen, and therefore it can be concluded that the UKB stored samples should be suitable for studies of bacterial microbiomics.

Table 3  Levels of the steroid hormones cortisol and DHEA, CRP and nitrite in salivas immediately frozen at −80°C or held at 4°C for 24 h and then frozen at −80°C as in the UKB protocol

<table>
<thead>
<tr>
<th>Analyte</th>
<th>Stored saliva</th>
<th>n</th>
<th>Mean ± SEM</th>
<th>95% CI</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cortisol (ng/ml)</td>
<td>Freezing</td>
<td>23</td>
<td>8.33 ± 1.85</td>
<td>3.14–13.52</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>UKB</td>
<td></td>
<td>11.62 ± 2.31</td>
<td>5.8–17.44</td>
<td></td>
</tr>
<tr>
<td>C-reactive protein (ng/ml)</td>
<td>Freezing</td>
<td>23</td>
<td>2.56 ± 0.44</td>
<td>1.63–3.49</td>
<td>0.49</td>
</tr>
<tr>
<td></td>
<td>UKB</td>
<td></td>
<td>2.60 ± 0.44</td>
<td>1.69–3.51</td>
<td></td>
</tr>
<tr>
<td>DHEA (pg/ml)</td>
<td>Freezing</td>
<td>23</td>
<td>102.9 ± 40.6</td>
<td>19.2–186.5</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>UKB</td>
<td></td>
<td>108.8 ± 42.9</td>
<td>20.4–197.2</td>
<td></td>
</tr>
<tr>
<td>Nitrite (µM)</td>
<td>Freezing</td>
<td>12</td>
<td>124.8 ± 27.7</td>
<td>63.7–185.8</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>UKB</td>
<td></td>
<td>86 ± 27.1</td>
<td>26.3–145.6</td>
<td></td>
</tr>
</tbody>
</table>

Cortisol, CRP and DHEA were analysed by ELISA and nitrite by the Greiss assay.

Figure 3  Nitrite concentration in saliva samples. Greiss reagent was used to assay nitrite concentrations in salivas collected from 12 volunteers, 1–26. Nitrite concentrations varied greatly between subjects, and there was a tendency for nitrite levels to be reduced in samples stored under UK Biobank (filled columns) conditions compared with salivas that were frozen immediately (light columns).
RNA is particularly susceptible to degradation, but the present study indicates that the quality of RNA extracted from saliva collected and stored under UK Biobank conditions was comparable to samples frozen immediately or preserved with an RNAase inhibitor. All three treatments generated similar Bioanalyser results and demonstrated the presence of host ribosomal RNA, in particular 18S and 28S. As found in previous studies, the calculated integrity of RNA (RIN number) was low, and this can be attributed to a number of factors. The substantial presence of RNA of oral bacterial origin reduces RIN, as the Bionalyser is unable to distinguish human and bacterial RNA in the same sample. Salivary mRNA is likely to be partially degraded owing to ribonucleases of host and bacterial origin. A very recent study has used RNA sequencing technology to demonstrate that

**Figure 4** Dendrogram showing similarities in bacterial community structure among saliva samples. Salivas were either immediately stored frozen (A) or collected and stored under UK Biobank conditions (B). Similarity was determined by means of the thetaxc calculation, indicated by horizontal distance.
many intact, small non-coding RNA molecules (microRNA and small nucleolar RNA) involved in control of gene expression are present in saliva.\textsuperscript{22} In other studies, salivary RNA preserving solutions have been used, but the quality of extracted RNA from whole-mouth saliva was similar to that in the present study (data not shown).

The host RNA isolated from whole-mouth saliva is derived not only from squamous epithelial cells and neutrophils but also from a cell-free, exosomal compartment referred to as the salivary transcriptome.\textsuperscript{23} and changes in the salivary transcriptome have been observed in both oral and systemic disease.\textsuperscript{24–26} As exosomes can be isolated after freezing and thawing of saliva, the stored UKB samples should be suitable for analysis of exosomal mRNA and microRNA, which is also present in exosomes.\textsuperscript{27} A recent publication indicates that salivary RNA can be assessed in cell-free saliva directly using RT-PCR.\textsuperscript{28}

Salivary cytokines have been examined in studies of a number of diseases, e.g. periodontal disease\textsuperscript{29} with diabetes,\textsuperscript{30} autoimmunity, mucosal ulceration and pre-malignancy.\textsuperscript{31–33} The adaptation to saliva of cytokines in UKB stored salivas, as the sample volume (100\mu{l}) required is relatively small. Salivary cytokine levels were similar after use of the UKB protocol compared with immediate freezing, although IL-10 and MCP-1 showed 10% decreases and lower P-values.\textsuperscript{34} CRP has been used as a biomarker of systemic and oral (periodontal) inflammation. It appears that salivary levels, although 1000-fold less than plasma levels, can show some correlation with the latter.\textsuperscript{35,36} The salivary concentration of CRP assayed in the present study was comparable to published data\textsuperscript{35} and showed no decrease in samples collected using the UKB protocol. Of the other proteins, which were assessed by SDS-PAGE and western blotting, \beta-actin showed some evidence of degradation under UKB processing conditions. Thus, it appears that for the purposes of biomarker development, many proteins present in saliva can withstand a period at 4°C.\textsuperscript{10,37}

Salivary concentrations of cortisol, DHEA and many other neutral steroid hormones tend to be well-correlated with serum concentrations of total hormone and of a similar quantity to serum concentrations of free hormone.\textsuperscript{38} Sequential saliva sampling has been undertaken in many behavioural studies of stress and depression. The Whitehall II epidemiological study in which saliva was sampled repeatedly from 4000 participants during a single day revealed associations between cortisol secretion patterns and cardiovascular disease.\textsuperscript{39} The results of the present study reflect those of previous studies demonstrating that salivary levels of steroid hormones are relatively stable,\textsuperscript{38} even at room temperature.\textsuperscript{40}

Nitrate is actively transported into saliva from blood by membrane transporters in salivary gland epithelial cells, making saliva a potentially useful biofluid for monitoring dietary nitrate.\textsuperscript{41} There is great interest in the role of oral bacteria in the reduction of nitrate to nitrite and the benefit to systemic blood pressure and haemodynamics provided by increased dietary nitrates.\textsuperscript{42–45} Measurement of salivary nitrite is potentially an important parameter to assess in epidemiological studies and, in the context of UKB, might be considered alongside urinary nitrate/nitrite levels and the dietary questionnaire, which was also part of the enhanced UKB baseline study. In the present study, there was great variability between individuals in nitrite concentrations in saliva. The measured concentration of nitrite is higher than in stimulated whole-mouth saliva, which was assayed in previous publications,\textsuperscript{46,47} but this can be accounted for by the diluting effect of a higher salivary flow rate. Overall nitrite levels in the UKB protocol-collected salivas were less than in immediately frozen saliva, and the reduction may be owing to bacterial metabolism in some samples even when chilled at 4°C. The extent of the reduction did not appear to be dependent on high or low levels of nitrite in the −80°C stored samples. Previously, salivary nitrite levels assayed by capillary electrophoresis were found to be stable at up to 48h at 3°C, and this may have been because of a lower bacterial presence in the saliva samples, as they were collected using an absorbent pad.\textsuperscript{48} The salivas in the present study had been subjected to re-freezing and thawing, but we have found in studies on salivas from other volunteers that this does not impact significantly on the measured amount of nitrite (data not shown).

The influence of long-term storage at −80°C or in liquid nitrogen, as anticipated with UKB samples, has not been evaluated in the present study, which focused on what might be considered the more damaging effects of the UKB protocol and a freeze–thaw event. A previous study of salivas, collected using a protocol similar to the UKB study, showed that during a 5-year period of storage at −80°C, most analytes including protein, DNA and RNA concentrations remained stable.\textsuperscript{49} It has also been reported that certain hormones and growth factors are able to withstand long-term freezing.\textsuperscript{50,51}

In conclusion, the present study indicates that saliva collected and stored in the UKB is likely to be of use in different studies that examine a variety of potential biomarkers.

Supplementary Data

Supplementary Data are available at IJE online.

Funding

This work was supported by UKBiobank and Department of Health via the National Institute for Health...
Research (NIHR) comprehensive Biomedical Research Centre award to Guy’s & St Thomas’ NHS Foundation Trust in partnership with King’s College London.

Acknowledgements

We thank Roche Diagnostics Corporation and Randox Laboratories Ltd. for providing the consumables for 454 sequencing and the multiplex cytokine analyses, respectively. Salimetrics Europe Ltd and Sigma-Aldrich Company provided some consumables for ELISAs and nitrite assay, respectively. Professor Dallas Swallow, UCL, kindly provided the EU-MUC5B antibody.

Conflict of interest: None declared.

KEY MESSAGES

- In UKB, salivas were collected to enhance the phenotypic characterization of subjects.

- No sample modifiers or preservations were added after saliva collection, and samples were held at 4°C for up to 24h before being placed in frozen storage.

- A paired comparison of the UKB collection protocol with immediate freezing of salivas revealed that sample degradation did not impact on assayed levels of a variety of salivary components of potential use as biomarkers.

Recommendations on saliva sample collection

- In studies where a variety of non-cellular and cellular components are to be assayed, unstimulated whole-mouth saliva samples should be collected into weighed tubes on ice for a timed period of 10 min. The weight of saliva should be recorded before dividing into aliquots and freezing. Commercially available RNA preserving and protein preserving solutions can be added to two aliquots.

- If sample processing cannot be undertaken at the collection point, then salivas should be chilled at 4°C during transportation and then aliquoted and frozen.

- For some specific applications, saliva ‘home collection kits’ might be suitable. For example absorbent pads can be used to collect saliva for steroid hormone assays whilst tubes with built-in preservatives can be used for DNA or RNA.

References

1 Galloway J. Putting the teeth into the UK Biobank. Prim Dent Care 2011;18:6–12.
4 Liening SH, Stanton SJ, Saini EK, Schultheiss OC. Salivary testosterone, cortisol, and progesterone: two-week stability, interhemorine correlations, and effects of time of day, menstrual cycle, and oral contraceptive use on steroid hormone levels. Physiol Behav 2010;99:8–16.
14 Pruesse E, Quast C, Knittel K et al. SILVA: a comprehensive online resource for quality checked and aligned


