

**Manuscript version: Author's Accepted Manuscript**

The version presented in WRAP is the author's accepted manuscript and may differ from the published version or Version of Record.

**Persistent WRAP URL:**

<http://wrap.warwick.ac.uk/126466>

**How to cite:**

Please refer to published version for the most recent bibliographic citation information. If a published version is known of, the repository item page linked to above, will contain details on accessing it.

**Copyright and reuse:**

The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions.

© 2019 Elsevier. Licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International <http://creativecommons.org/licenses/by-nc-nd/4.0/>.



**Publisher's statement:**

Please refer to the repository item page, publisher's statement section, for further information.

For more information, please contact the WRAP Team at: [wrap@warwick.ac.uk](mailto:wrap@warwick.ac.uk).

1 **Detection of Group B Streptococcus in pregnancy by vaginal volatile organic compound**  
2 **analysis: a prospective exploratory study**

3

4 **Authors**

5 Lauren Lacey<sup>1,2</sup>

6 Emma Daulton<sup>3</sup>

7 Alfian Wicaksono<sup>3</sup>

8 James A. Covington<sup>3</sup>

9 Siobhan Quenby<sup>1,2</sup>

10

11 1. Warwick Medical School, University of Warwick, Coventry, CV4 7AL

12 2. Department of Obstetrics & Gynaecology, University Hospitals Coventry & Warwickshire,  
13 Coventry, CV2 2DX

14 3. School of Engineering, University of Warwick, Coventry, CV4 7AL

15

16 **Corresponding author:** Lauren Lacey, Clinical Sciences Research Laboratories, University  
17 Hospitals Coventry & Warwickshire, Clifford Bridge Road, Coventry, CV2 2DX, Tel:  
18 02476964000 bleep 1351, Email: l.lacey.1@warwick.ac.uk

19

20 **Running Head:** Detection of GBS in pregnancy by VOC analysis

21 **Abbreviations**

22 EOGBS - Early-onset group B Streptococcus

23 GC - Gas chromatograph

24 GC-IMS - Gas chromatograph ion mobility spectrometry

- 25 GCMS - Gas chromatograph mass spectrometer
- 26 GBS - Group B Streptococcus
- 27 IMS - Ion mobility spectrometry
- 28 VOC - Volatile organic compound

29 **Abstract**

30 Our objective was to assess whether volatile organic compound (VOC) analysis of vaginal  
31 swabs can detect maternal Group B Streptococcus during pregnancy in a prospective  
32 exploratory study. 243 women attending a high risk antenatal clinic at one university teaching  
33 hospital in the UK consented to take part and provide vaginal swabs throughout pregnancy.  
34 VOC analysis of vaginal swabs was undertaken and compared with the reference standard of  
35 GBS detected using enrichment culture method. The chemical components that emanated  
36 from the vaginal swabs were measured by gas chromatograph ion mobility spectrometry (GC-  
37 IMS). This platform has both high sensitivity and good specificity to a range of chemical  
38 compounds. Our main outcome was to determine the sensitivity and specificity of VOC  
39 analysis for the detection of maternal GBS in vaginal swabs during pregnancy. Our study has  
40 demonstrated that the sensitivity and specificity of the VOC analysis by GC-IMS for the  
41 detection of GBS from vaginal swabs was 0.81 (95% CI, 0.71-0.89) and 0.97 (95% CI, 0.91-1)  
42 respectively. We conclude that the use of VOCs as biomarkers for the detection of maternal  
43 GBS in the vagina is a novel tool. As this test produces results within minutes and is of low  
44 unit test cost it has the potential to be used in clinical settings, where fast diagnosis is  
45 important, for example, a patient in early labour.

46 **Main research article**

47 **Introduction**

48 Group B Streptococcus (GBS) is the most frequent cause of life-threatening early onset  
49 infection in newborn infants in the UK, known as early-onset group B Streptococcus (EOGBS)  
50 disease [1]. The incidence of EOGBS in the UK in 2015 was 0.57/1000 births [2]. GBS commonly  
51 colonises the gastrointestinal and genital tract of adults, with a global mean prevalence of  
52 17%[3]. GBS only rarely causes disease in the immunocompromised adult, but it can pose a  
53 significant risk to newborn infants due to their immature immune systems.

54

55 The optimal screening strategy to prevent EOGBS is uncertain. Internationally there is a  
56 variation in guidelines, the 2019 ACOG recommends screening between 36+0-37+6 weeks  
57 gestation [4] but in the UK, universal screening is not currently recommended [5]. Maternal  
58 colonisation with GBS is the primary risk factor for disease (transmission to newborns is  
59 40–70% and of these 1–2% will develop infection) [6, 7]. UK guidelines advocate offering  
60 intrapartum antibiotics to women found to be colonised during pregnancy and to those with  
61 other risk factors as this reduces the risk of culture positive EOGBS disease in the neonate [1,  
62 8].

63

64 However, GBS colonisation status can be persistent but also intermittent and therefore  
65 transient during pregnancy. Up to 13% of women who are GBS positive in the before 37 weeks  
66 gestation receive unnecessary prophylactic antibiotics during labour [9, 10], which may  
67 contribute to increasing antibiotic resistance. The previous universal screening policy in the  
68 United States tested women between 35–37 weeks, studies of this had demonstrated that

69 among women who had a negative screening, 2–10% will become colonised before the onset  
70 of labour[10, 11].

71

72 The currently used diagnostic methods for colonisation with GBS utilise time-consuming  
73 enrichment culture methods (over 24 hours) [12] and therefore aren't appropriate for an  
74 intrapartum scenario. However, this method maximises GBS identification in cultures and is  
75 therefore the recommended technique in current guidance [4, 12]. Ideally, we would be able  
76 to screen for GBS colonisation at the start of labour so that only those women colonised in  
77 labour would be given antibiotics. Hence, there is a need to develop an accurate point of care  
78 test, which produces results within a few minutes, to reduce the burden of EOGBS disease.

79

80 An approach that could be applied to this medical need is to measure the volatile organic  
81 compounds (VOCs) that emanate from a vaginal sample. The concept of measuring VOCs for  
82 clinical applications is currently gaining momentum, with a broad range of biological materials  
83 and diseases being investigated. For example researchers have investigated diseases as  
84 diverse as colon cancer, pancreatic cancer, irritable bowel disease and respiratory tract  
85 infections in urine, stool, swabs and breath [13-16]. Such techniques hold considerable  
86 promise, as the test can be undertaken in a clinically relevant time period, the cost per test  
87 can be low and the instrument can be sited near or in the ward, thus ideal of point-of-care  
88 needs. The objective of this study was to determine the ability of VOC analysis to detect  
89 maternal GBS in vaginal swabs in pregnant women. Meeting this objective involved  
90 comparing GBS detected on vaginal swabs using the enrichment culture method to VOCs  
91 analysed by GC-IMS (gas chromatography ion mobility spectrometry).

92

93 **Material and Methods**

94 **Study design**

95 We conducted a prospective exploratory study at one UK hospital (University Hospitals  
96 Coventry & Warwickshire) serving a diverse population. The study protocol was approved by  
97 the NHS Research Ethics Committee West Midlands Birmingham South on 14th January 2014  
98 (13/WM/0486) and all participants gave written informed consent. The Group B Strep  
99 Support charity were consulted prior to the application for funding regarding a patient  
100 perspective about the study. Research was carried out according to The Code of Ethics of the  
101 World Medical Association (Declaration of Helsinki).

102

103 **Participants and test methods**

104 From 25th January 2017, women between 14–36 weeks gestation were consented during  
105 their attendance to a high risk antenatal clinic for women at an increased risk of spontaneous  
106 preterm birth. A speculum examination was performed as per patient routine care. No  
107 specific hygiene advice was given. As part of routine screening in the preterm prevention  
108 clinic, a vaginal swab (reference standard) for microbiology culture and sensitivity testing  
109 using the enriched culture method was taken and placed into a non-nutritive transport  
110 medium, and concurrently two cotton swabs were used to obtain index test vaginal samples.  
111 The index test swabs were then placed in a universal containers and snap frozen in liquid  
112 nitrogen and stored at  $-80^{\circ}\text{C}$ . Specimens were obtained by gently rotating the swabs across  
113 the mucosa of the vagina. Biomedical scientists independently interpreted the reference  
114 swab cultures. Demographic data including age at booking pregnancy, BMI, parity and  
115 ethnicity were collected about each woman (Table 1). Samples were taken in a consecutive

116 series from all women who consented in the clinic, some women consented to samples being  
117 taken during every attendance to the clinic.

118

### 119 **Chemical analyser**

120 Chemical vapour analysis of the index swabs was undertaken in the BioMedical Sensors  
121 Laboratory, School of Engineering, University of Warwick. Here a Gas Chromatograph-Ion  
122 Mobility Spectrometer (GC-IMS) was used. Our group have previously used this instrument  
123 on a range of medical conditions including respiratory tract infections, Coeliac's disease and  
124 irritable bowel disease [16-18]. This instrument was chosen over more traditional gas  
125 chromatograph mass spectrometer (GCMS) as the basic sensitivity of the instrument is much  
126 higher than GCMS, it can use nitrogen/air as the carrier gas (so no need for expensive carrier  
127 gases such as helium), has a lower purchase/test cost than GCMS and has a much smaller  
128 form factor, making it applicable for a ward setting.

129

130 The GC-IMS instrument used was manufactured by G.A.S. (GC-IMS is also the product name,  
131 Dortmund, Germany). In use the samples, in this case formed of a mixture of VOCs that  
132 emanate from the vaginal swab, are injected into the GC-IMS. These VOCs are pre-separated  
133 by the gas chromatograph (GC) column, which takes the complex mix of chemicals and  
134 separates them based on their interaction with the long column coated with a retentive  
135 layer. Thus chemicals elude from the column at different times (known as the retention time).  
136 These pre-separated chemicals exit the GC and enter a drift tube ion mobility spectrometry  
137 (IMS) detector. Here the molecules are ionized using a radioactive source (in this case tritium)  
138 and then released into the drift tube in a controlled manner. The ions are then moved along  
139 the drift tube using an electric field. At the same time a buffer gas (nitrogen) is fed in the

140 opposite direction to the ions. The resultant impacts between the ions and the buffer gas  
141 reduce the velocity of the ions. Thus ions achieve different velocities due to its interaction  
142 with both the electric field and the buffer gas, which is inversely proportional to their size,  
143 mass and charge and then are collected on a Faraday plate, to provide a time-dependent  
144 signal corresponding with ion mobility. The device can measure substances in the low ppb  
145 range. The instrument was suitable for a clinical setting and was placed on a work surface  
146 (dimensions: 45 x 50 x 20 cm; mass: 20 kg).

147

### 148 **Chemical testing and analysis**

149 In total 607 samples from 243 women were tested using the G.A.S. GC-IMS system. Swabs  
150 were thawed and transferred to a 20ml glass vial in batches of 20. The vials were then sealed  
151 with a crimp top lid fitted with a PTFE septum. The index samples were then placed in a vial  
152 tray cooled and maintained at 4°C to reduced unwanted odour emission and sample  
153 degradation, whilst other samples were being tested. Prior to the sample measurement,  
154 samples were heated to 40°C for 10 minutes. The sample line for the GC-IMS was inserted  
155 into the septa of the vial using a needle and 2mls of sample were then extracted from the vial  
156 and injected into the analytical platform. The machine settings were as follows: E1: 150  
157 ml/min (for the drift tube IMS), E2: 20 ml/min (for the GC column) and the pump at 25%. The  
158 total run time was 10 minutes. The temperatures were set to: T1: 45°C, T2: 80°C, and T3:  
159 70°C.

160

### 161 **Statistical analysis**

162 The data was analysed using the statistical pipeline successfully used in [16-18]. In summary,  
163 the GC-IMS data was first extracted using the L.A.V. software (v2.2.1, G.A.S, Germany), which

164 converts the data from its native file format to a text file. This was followed by a pre-  
165 processing step to reduce the dimensionality of the data, making the statistical analysis less  
166 computationally intensive. A typical GC-IMS output file (of a single sample) contains typically  
167 11 million data points. Though the number of data points is high, the information content is  
168 sparse, with the all of the values containing non-background information being located  
169 around the centre of the dataset. Thus, we are able to crop the central section of the data  
170 and then apply a threshold to make the background values all be zero. These values are  
171 selected by visual inspection of the data using the LAV software and results in around a 500  
172 fold reduction in the number of non-zero data points. Once completed, the data was analysed  
173 using a 10-fold cross validation approach. In each fold, the data was split into a 90% training  
174 set and a 10% test set. Features with discriminatory power were identified from the training set  
175 using a rank-sum test and 50 features with the lowest p-value were taken forward for  
176 classification. Here, five different classifiers we used, specifically sparse logistic regression,  
177 random forest, Gaussian process classifier, support vector machine and neural network (this  
178 set is commonly used within our pipeline). Once the training models had been created, they  
179 were applied to the same features in the test set. This process is repeated ten times until all  
180 the data has a test result. This process provided test probabilities for each sample and from  
181 this, statistical values, including sensitivity and specificity were calculated.

182

### 183 **Results**

184 Between January 2017 and August 2018, 243 women had vaginal swabs taken throughout  
185 pregnancy. The demographic data of these women shows the majority of women were white  
186 (79.0%) and two thirds were multiparous and one third nulliparous (Table 1). The maternal

187 GBS colonisation rates as defined by a positive enriched culture from vaginal swabs was 13.6%  
188 (corresponding to 33 women).

189

190 Figure 1 shows a typical output of the GC-IMS to a positive swab. The background is  
191 represented in blue with the non-blue areas showing that the instrument is detecting  
192 chemicals. The intensity of the peak (with red being the highest intensity) represents the  
193 amount of ions (and thus the chemical) detected. In general, each of the circular areas of  
194 higher intensity represent a different chemical. Furthermore, it can be see that the majority  
195 of the reponse is in the central section of the output. **What was found is that the number of**  
196 **chemical peaks changed significantly across the cohort (independent of them being GBS**  
197 **positive or negative), which is likely to reflect vaginal biome, but were not investigated further**  
198 **in this study.** The data from the G.A.S. GC-IMS was analysed as described and the statistical  
199 output is shown in Table 2. The high sensitivity and specificity indicates a strong signal is  
200 associated with GBS colonisation. The shape of the ROC curve illustrates the test has an  
201 excellent ability to discriminate between those with GBS from those without (Figure 2, Table  
202 2) To help visualise these differences, we have also created a box plot of the probabilities  
203 generated by the classifier for each sample, as shown in figure 3. The line in the centre of the  
204 box plot is the median and the upper/lower boundaries are the 25<sup>th</sup> and 75<sup>th</sup> percentiles and  
205 error bars defining the 10<sup>th</sup> and 90<sup>th</sup> percentiles. Data points outside this region are  
206 individually plotted as outliers. The plot shows that there are significant differences in the  
207 probabilities of GBS and non-GBS samples.

208

209 **Discussion**

210 This is the first exploratory study to our knowledge to report the use of VOC analysis to detect  
211 GBS. In our prospective cohort study we investigated the potential of VOCs in the detection  
212 of GBS on vaginal swabs taken at the same time as samples for traditional GBS testing  
213 methods. The ultimate aim is to be able to implement this technology as a point of care test  
214 for women intrapartum, reducing the incidence of EOGBS disease by appropriate  
215 administration of antibiotic prophylaxis.

216

217 VOC profiles from vaginal swabs taken from pregnant women discriminated those whose  
218 swabs grow GBS from those who did not with a high sensitivity and specificity. Our results  
219 suggest that women who are colonised with GBS have chemically different vaginal swabs to  
220 those who are not colonised. The vagina has its own varied microbiome, but our data suggests  
221 that despite this, there are differences in VOCs from vaginal swabs in those who are colonised  
222 with enough GBS to be detected by the enriched culture method. These GBS associated  
223 differences in VOCs were demonstrated and are detectable with this novel technology. The  
224 VOCs detected are believed to be the gaseous waste products produced from the metabolic  
225 pathways of the bacteria in the vaginal, which occur as a result of the complex interactions in  
226 the vagina between, the vaginal and cervical epithelial cells, the vagina flora and invading  
227 pathogens. False positive tests could be driven by alterations in either the vaginal flora or the  
228 maternal host response. This study suggests that GBS produces a unique VOC fingerprint in  
229 pregnancy. A recent meta-analysis of the VOC literature found that VOCs could differentiate  
230 11 other microbial pathogens in multiple disease states, but did not include the detection of  
231 GBS in pregnancy [19, 20].

232

233 Our results demonstrate that the G.A.S. GC-IMS instrument has a very high specificity and  
234 negative predictive value for the detection of GBS in vaginal swabs. This technology could  
235 now be developed as a bedside test for GBS colonisation. Previous studies have demonstrated  
236 high patient acceptability for intrapartum testing for GBS [21]. In the acute intrapartum  
237 scenario, women could have a swab taken and analysed in a hand held device in minutes.  
238 Where the results are positive, this could guide clinicians to prompt and appropriate  
239 administration of intrapartum antibiotics, reducing the risk of EOGBS. This would reduce  
240 residual GBS disease as it would allow us to treat the 10% women who may be negative at  
241 screening but convert to positive by the time of labour. The high negative predictive value of  
242 the test could be used to counsel families about the low likelihood of colonisation with GBS  
243 and bring into question whether administration of antibiotic prophylaxis is necessary,  
244 reducing unnecessary antibiotic exposure to both the mother and infant. Furthermore, a large  
245 number of women need to be tested for a screening program and the cost of this test is  
246 minimal.

247

248 Strengths of the study include the large number (n=607) of swabs analysed using the G.A.S.  
249 GC-IMS instrument and compared to the reference standard. We complied with the STARD  
250 statement and minimised bias as far as possible. However, there were a few limitations to our  
251 study. Our prevalence of colonisation with GBS is lower than expected at 13.6%, compared  
252 with the global average of 17.9% and European average 19.0% [3]. In the clinic women have  
253 a vaginal swab only (as part of their screening for risk of spontaneous preterm birth), this is  
254 not in keeping with recommendations for specimen collection for detection of colonisation  
255 of GBS. Swabbing both the lower vagina and rectum increases the culture yield when  
256 compared to sampling the vagina only [6, 22]. Previous studies sub-analysis has illustrated

257 that colonisation from low vaginal swabs only had a mean prevalence of 14.2%[3], this has  
258 more similarity to our cohort. In future studies we would aim to comply with the  
259 recommendations of a low vaginal and rectal swab to identify which women are colonised  
260 both for our reference and index test. Furthermore, we have not attempted to identify the  
261 specific biomarkers associated with GBS. This would require the use of a more sophisticated  
262 measurement platform (such as GCMS), which we were unable to undertake within this study.

263

264 The Centre for Disease Control and Prevention recommend that in order to be considered  
265 clinically useful in the intrapartum period, a point of care test should have sensitivity and  
266 specificity equal to or greater than 90% [6]. It is possible that optimising our sample collection  
267 (for both the index and reference test) as recommended may further increase our sensitivity  
268 to reach this threshold. Furthermore, due to logistical issues, samples were not analysed on  
269 site and therefore had to be frozen and then transferred for analysis at a later date. This may  
270 have caused degradation of the VOCs and influenced the fingerprints obtained as has been  
271 shown in previous studies[23].

272

273 At present an optimal screening strategy to prevent EOGBS has not been established, the high  
274 specificity and sensitivity obtained in this study suggest that with further work, it could be  
275 possible to implement this type of technology into a clinically useful screening pathway. It  
276 would allow the timely initiation of therapeutics by clinicians to prevent EOGBS and  
277 furthermore prevent women and their babies unnecessarily remaining in hospital for  
278 observation and/or antibiotic administration to the neonate when not indicated. This type of  
279 analysis and use of VOCS as biomarkers has huge potential in medical diagnostics, VOCs are  
280 thought to reflect complex changes in the vagina and therefore this technology has the

281 potential to be utilised in the assessment of a variety of diseases in obstetrics and  
282 gynaecology.

283

284 In conclusion, EOGBS disease remains the leading infectious cause of morbidity and mortality  
285 amongst neonates. Preventative efforts have reduced the burden of this disease over time  
286 but at present worldwide no universal screening tool or pathway can be agreed. This study  
287 has shown that the VOC signature present in vaginal swabs of pregnant women distinguished  
288 those swabs from which GBS was detected. Using the G.A.S. GC-IMS analytical platform with  
289 a sensitivity and specificity for GBS colonisation of 0.81 and 0.97 respectively. Development  
290 of this technology has the potential to provide clinically useful and cost-effective universal  
291 screening intrapartum for colonisation with GBS.

292

#### 293 **Disclosure of interests**

294 None declared

295

#### 296 **Contribution of authorship**

297 Conception: JC & SQ. Planning: LL, JC and SQ. Carrying out: LL, ED, AW, JC and SQ. Writing: LL,  
298 JC and SQ. Supervision: JC and SQ

299

#### 300 **Details of ethics approval**

301 The study protocol was approved by the NHS Research Ethics Committee West Midlands  
302 Birmingham South on 14th January 2014 (13/WM/0486)

303

#### 304 **Acknowledgments**

305 We would firstly like to thank the women who participated in the study. We would also like  
306 to thank the team in the Biomedical Research Unit based at University Hospitals Coventry &  
307 Warwickshire and the Arden Tissue Bank for their support during the study.

308

309 The study was funded by the Rosetrees Trust. They did not have a role in study design; in the  
310 collection, analysis and interpretation of data; in the writing of the report; or in the decision  
311 to submit the article for publication.

312

313 **References**

- 314 [1] Prevention of Early-onset Neonatal Group B Streptococcal Disease. *BJOG: An*  
315 *International Journal of Obstetrics & Gynaecology* 2017;124(12):e280-e305.
- 316 [2] O'Sullivan C LT, Efstratiou A, Patel D, Cunney R, Meehan M, et al. P3 Group B  
317 Streptococcal (GBS) disease in UK and Irish infants younger than 90 days, 2014-2015.  
318 *Archives of Disease in Childhood* ; 2016;101:A2.
- 319 [3] Kwatra G, Cunnington MC, Merrall E, Adrian PV, Ip M, Klugman KP, et al. Prevalence  
320 of maternal colonisation with group B streptococcus: a systematic review and meta-  
321 analysis. *The Lancet Infectious Diseases* 2016;16(9):1076-84.
- 322 [4] The American College of Obstetricians & Gynecologists. ACOG Committee Opinion  
323 Number 782: Prevention of Group B Streptococcal Early-Onset Disease in Newborns.  
324 The American College of Obstetricians & Gynecologists,; 2019.
- 325 [5] Royal College of Obstetricians & Gynaecologists. Prevention of Early-onset Neonatal  
326 Group B Streptococcal Disease: Green-top Guideline No. 36. *BJOG*  
327 2017;124(12):e280-e305.
- 328 [6] Centers for Disease Control & Prevention. Prevention of Perinatal Group B  
329 Streptococcal Disease. Revised Guidelines from CDC 2010. *MMWR* 2010;59(RR10).
- 330 [7] Van Dyke MK, Phares CR, Lynfield R, Thomas AR, Arnold KE, Craig AS, et al.  
331 Evaluation of Universal Antenatal Screening for Group B Streptococcus. *New England*  
332 *Journal of Medicine* 2009;360(25):2626-36.
- 333 [8] Schrag SJ, Zywicki S, Farley MM, Reingold AL, Harrison LH, Lefkowitz LB, et al. Group  
334 B Streptococcal Disease in the Era of Intrapartum Antibiotic Prophylaxis. *New*  
335 *England Journal of Medicine* 2000;342(1):15-20.
- 336 [9] Hansen SM, Uldbjerg N, Kilian M, Sørensen UBS. Dynamics of Streptococcus  
337 agalactiae colonization in women during and after pregnancy and in their infants.  
338 *Journal of clinical microbiology* 2004;42(1):83-9.
- 339 [10] Asrat T, Rumney P, Towers C, Preslicka C. The accuracy of late third trimester  
340 antenatal screening for group B streptococcus in predicting GBS colonization at  
341 delivery. *American Journal of Obstetrics and Gynecology* 2006;195(6,  
342 Supplement):S40.
- 343 [11] Feuerschuette OHM, Silveira SK, Cancelier ACL, da Silva RM, Trevisol DJ, Pereira JR.  
344 Diagnostic yield of real-time polymerase chain reaction in the diagnosis of  
345 intrapartum maternal rectovaginal colonization by group B Streptococcus: a  
346 systematic review with meta-analysis. *Diagnostic Microbiology and Infectious*  
347 *Disease* 2018;91(2):99-104.
- 348 [12] Church DL, Baxter H, Lloyd T, Miller B, Elsayed S.
- 349 [13] Arasaradnam RP, McFarlane MJ, Ryan-Fisher C, Westenbrink E, Hodges P, Thomas  
350 MG, et al. Detection of colorectal cancer (CRC) by urinary volatile organic compound  
351 analysis. *PLoS One* 2014;9(9):e108750.
- 352 [14] Arasaradnam RP, Wicaksono A, O'Brien H, Kocher HM, Covington JA, Crnogorac-  
353 Jurcevic T. Noninvasive Diagnosis of Pancreatic Cancer Through Detection of Volatile  
354 Organic Compounds in Urine. *Gastroenterology* 2018;154(3):485-7.e1.
- 355 [15] Arasaradnam RP, Westenbrink E, McFarlane MJ, Harbord R, Chambers S, O'Connell  
356 N, et al. Differentiating coeliac disease from irritable bowel syndrome by urinary  
357 volatile organic compound analysis--a pilot study. *PLoS One* 2014;9(10):e107312.

358 [16] Lewis JM, Savage RS, Beeching NJ, Beadsworth MBJ, Feasey N, Covington JA.  
359 Identifying volatile metabolite signatures for the diagnosis of bacterial respiratory  
360 tract infection using electronic nose technology: A pilot study. *PLoS One*  
361 2017;12(12):e0188879.

362 [17] Rouvroye DM, Wicaksono A, Bosch S, Savelkoul E, Covington AJ, Beaumont H, et al.  
363 Faecal Scent as a Novel Non-Invasive Biomarker to Discriminate between Coeliac  
364 Disease and Refractory Coeliac Disease: A Proof of Principle Study. *Biosensors*  
365 2019;9(2).

366 [18] Tiele A, Wicaksono A, Kansara J, Arasaradnam PR, Covington AJ. Breath Analysis  
367 Using eNose and Ion Mobility Technology to Diagnose Inflammatory Bowel  
368 Disease—A Pilot Study. *Biosensors* 2019;9(2).

369 [19] Bos LDJ, Sterk PJ, Schultz MJ. Volatile Metabolites of Pathogens: A Systematic  
370 Review. *PLOS Pathogens* 2013;9(5):e1003311.

371 [20] Palma SICJ, Tragedo AP, Porteira AR, Frias MJ, Gamboa H, Roque ACA. Machine  
372 learning for the meta-analyses of microbial pathogens' volatile signatures. *Scientific*  
373 *reports* 2018;8(1):3360-.

374 [21] Daniels JP, Gray J, Pattison HM, Gray R, Hills RK, Khan KS, et al. Intrapartum tests for  
375 group B streptococcus: accuracy and acceptability of screening. *BJOG: An*  
376 *International Journal of Obstetrics & Gynaecology* 2011;118(2):257-65.

377 [22] Platt MW, McLaughlin JC, Gilson GJ, Wellhoner MF, Nims LJ. Increased recovery of  
378 group B streptococcus by the inclusion of rectal culturing and enrichment. *Diagnostic*  
379 *Microbiology and Infectious Disease* 1995;21(2):65-8.

380 [23] Esfahani S, Wicaksono A, Mozdiak E, Arasaradnam PR, Covington AJ. Non-Invasive  
381 Diagnosis of Diabetes by Volatile Organic Compounds in Urine Using FAIMS and  
382 Fox4000 Electronic Nose. *Biosensors* 2018;8(4).

383

384 **Table/Figure caption list**

385 Table 1: Demographic characteristics of the women taking part in the study, the booking BMI  
386 of six patients is not known.

387

388 Table 2: Statistical output from G.A.S. GC-IMS analytical platform using a sparse logistic  
389 regression classifier.

390

391 Figure 1: Typical output of the GC-IMS to a swab positive for GBS. The x axis represents the  
392 drift time of the IMS and the y axis the retention time of the same eluding out of the GC. The  
393 non-blue areas are chemical signals be detected by the instrument.

394

395 Figure 2: ROC output for G.A.S. GC-IMS instrument for women colonised with GBS in  
396 pregnancy versus those who are not colonised

397

398 Figure 3: The boxplot representing the distribution of level of probabilities of assigning VOC'S  
399 outputs, of patients with and without GBS using the classifiers described. The closer the  
400 probability is to 1, the more certainty the classification model has to define the sample group.  
401 The probability of assigning swabs to GBS positive or negative are clearly separated indicating  
402 high assurance of the classification model to classify the patient to the correct group.

403