

BMJ Open Feasibility and acceptability of breath research in primary care: a prospective, cross-sectional, observational study

Georgia Woodfield ¹, Ilaria Belluomo,¹ Piers R Boshier,¹ Annabelle Waller,¹ Maya Fayyad,¹ Christian von Wagner ², Amanda J Cross,³ George B Hanna¹

To cite: Woodfield G, Belluomo I, Boshier PR, *et al.* Feasibility and acceptability of breath research in primary care: a prospective, cross-sectional, observational study. *BMJ Open* 2021;**11**:e044691. doi:10.1136/bmjopen-2020-044691

► Prepublication history and supplemental material for this paper is available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2020-044691>).

Received 12 September 2020
Revised 02 January 2021
Accepted 23 March 2021



© Author(s) (or their employer(s)) 2021. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹Department of Surgery and Cancer, Imperial College London, London, UK

²Institute of Epidemiology and Health Care, University College London, London, UK

³School of Public Health, Imperial College London, London, UK

Correspondence to

Professor George B Hanna; g.hanna@imperial.ac.uk

ABSTRACT

Objectives To examine the feasibility and acceptability of breath research in primary care.

Design Non-randomised, prospective, mixed-methods cross-sectional observational study.

Setting Twenty-six urban primary care practices.

Participants 1002 patients aged 18–90 years with gastrointestinal symptoms.

Main outcome measures During the first 6 months of the study (phase 1), feasibility of patient enrolment using face-to-face, telephone or SMS-messaging (Short Message Service) enrolment strategies, as well as processes for breath testing at local primary care practices, were evaluated. A mixed-method iterative study design was adopted and outcomes evaluated using weekly Plan-Do-Study-Act cycles, focus groups and general practitioner (GP) questionnaires.

During the second 6 months of the study (phase 2), patient and GP acceptability of the breath test and testing process was assessed using questionnaires. In addition a ‘single practice’ recruitment model was compared with a ‘hub and spoke’ centralised recruitment model with regards to enrolment ability and patient acceptability.

Throughout the study feasibility of the collection of a large number of breath samples by clinical staff over multiple study sites was evaluated and quantified by the analysis of these samples using mass spectrometry.

Results 1002 patients were recruited within 192 sampling days. Both ‘single practice’ and ‘hub and spoke’ recruitment models were effective with an average of 5.3 and 4.3 patients accrued per day, respectively. The ‘hub and spoke’ model with SMS messaging was the most efficient combined method of patient accrual. Acceptability of the test was high among both patients and GPs. The methodology for collection, handling and analysis of breath samples was effective, with 95% of samples meeting quality criteria.

Conclusions Large-scale breath testing in primary care was feasible and acceptable. This study provides a practical framework to guide the design of Phase III trials examining the performance of breath testing in primary care.

BACKGROUND

Late diagnosis is a common feature of patients with gastrointestinal cancers and is associated with poor survival.^{1 2} Patients

Strengths and limitations of this study

- This is the largest ever breath testing study to be conducted within a primary care setting.
- The study recruited 1002 patients from primary care for breath testing using face-to-face, telephone and SMS-messaging enrolment strategies, in patients with gastrointestinal symptoms.
- The study explored models for breath sampling including single-site sampling at local primary care practices, as well as a centralised breath sampling strategy.
- The study assessed feasibility and acceptability of breath testing in patients with gastrointestinal symptoms from both a patient and a healthcare provider perspective, using a concurrent iterative mixed methods approach.
- This study did not assess diagnostic accuracy of the breath test for diagnosis of gastrointestinal cancers or ascertain the optimum place a breath test may have in existing diagnostic pathways, where both of these factors could affect feasibility and acceptability of a future breath test.

with early oesophageal, gastric, pancreatic or colorectal cancers often have non-specific symptoms typical of many common benign conditions.^{3 4} In comparison, ‘red flag’ symptoms linked to gastrointestinal cancers often indicate advanced incurable disease.^{5–7} Currently only patients with ‘red flag’ symptoms are urgently referred for diagnostic testing.^{8 9} Opening existing diagnostic pathways to patients with non-specific symptoms can however lead to potentially harmful over-investigation that would consume National Health Service resources and cause unnecessary anxiety for the majority of patients who do not have cancer.

There remains, therefore, an unmet clinical need to establish accurate, accessible and affordable methods for early gastrointestinal cancer detection that are not reliant on traditional approaches that are invasive and expensive. The non-invasive detection of

disease markers within human breath is a promising field of research that has the opportunity to transform our ability to detect cancers of unmet need. Breath testing has the ideal characteristics of a triage test for early cancer detection, being non-invasive and acceptable to patients. A breath test could serve as a community triage test, for patients with vague symptoms that may be associated with cancer, but do not currently meet ('red flag') criteria for investigation. A breath test would support general practitioners (GPs) as well as other healthcare providers to determine which patients most warrant referral using existing gastrointestinal cancer diagnostic pathways.

The test is based on the detection of volatile organic compounds (VOCs) within exhaled breath. VOCs are produced by humans as a result of both normal and abnormal metabolism. Once released into the systemic circulation, VOCs may travel to the lungs where they are excreted in exhaled breath.¹⁰ A systematic review of breath testing for cancer identified distinctive VOC signals for different tumour sites with pooled sensitivity and specificity of 79% and 89%, respectively (including lung, breast, gastrointestinal, head and neck, prostate and gynaecological tumours).¹¹ Studies of different gastrointestinal tumour sites also showed different VOC biomarkers for oesophagogastric, pancreatic and colorectal cancers, providing the opportunity for a single breath test to diagnose different cancers based on their unique VOC signature, in a similar way to a single blood draw being used to assess for multiple diseases.¹²⁻¹⁵

Before large-scale primary care trials can occur, there is a need to evaluate different recruitment and engagement strategies to determine the feasibility and acceptability of the test. Historically, despite an ever-increasing need for high-quality research in primary care, adequate patient recruitment has been a critical barrier.^{16 17} Reasons for this include dependence on financial incentives,¹⁸ inadequate infrastructure, time constraints within busy practices, lack of buy-in and failure to show adequate recognition for those contributing to the study.^{17 19} Mitigation of these challenges is essential if GPs are to continue contributing to research and clinical trials.

The primary aim of this study was to inform the design of future large-scale studies by examining the feasibility of different recruitment and engagement strategies for breath testing in primary care. The secondary aim was to understand the acceptability of the breath test among both patients and GPs.

METHODOLOGY

Study setting and patients

The Methodological Approaches towards a Gastrointestinal Cancer Breath Test (MAGIC) study was a cross-sectional observational breath-testing study based in 26 primary care practices within Central and Northwest London (online supplemental file S1). Practices were approached based on previous research participation or expression of interest. Breath sampling was coordinated

and performed by clinical study officers (CSOs) from the National Institute of Health Research (NIHR) clinical research network North West London and local practice nurses.

The recruitment target was 1000 patients over 12 months (260 sampling days). Study eligibility criteria were patients aged 18–90 years old who were suffering from upper or lower gastrointestinal symptoms. Gastrointestinal symptoms included all 2 weeks wait (2WW) and urgent referral symptoms within National Institute for Health and Care Excellence guidelines.^{8 9} GPs and trial staff were provided with a list of all eligible gastrointestinal symptoms (online supplemental file S2–S4). Patients with chronic symptoms (lasting >2 months) were included only if they had ongoing requirement for pharmacological control. Patient eligibility was assessed by GPs at the time of a routine face-to-face appointment or from review of electronic medical records.

All subjects provided informed written consent prior to participation.

Methods of recruitment

To evaluate different methods of recruitment the study was divided into two phases. During phase 1 (29 November 2016 to 26 May 2017) 'single practice' breath sampling was conducted at 16 primary care practices. Breath sampling occurred at two practices concurrently for 2 weeks before equipment and staff were relocated to two new practices.

During phase 2 (7 November 2017 to 14 June 2018) a 'hub and spoke model' was trialled. Seven practices that were part of the Central London Healthcare GP federation recruited concurrently by referring all patients to a single central practice for breath testing (Marylebone Health Centre), regardless of the patients' registered GP practice. Local 'single practice' breath testing was also continued at three practices during phase 2 recruitment.

Methods of patient engagement

Patients who met eligibility criteria entered the study by one of four methods: face-to-face same day; face-to-face prebooking; telephoning or SMS (text) messaging. In phase 1, all four methods of patient enrolment were assessed, whereas in phase 2, SMS messaging was used exclusively.

For face-to-face enrolment, GPs identified and approached potentially eligible patients at the time of routine consultation. Those willing to participate in the study were enrolled either on the same day (face-to-face same day) or at an agreed future time and date (face-to-face prebooking).

The telephone and SMS recruitment models involved manual or automated searching of practice electronic medical records to identify potentially eligible patients (online supplemental files S5 and S6). Identified patients were contacted via either telephone or SMS message, inviting them to participate in the study. Patients who received an SMS message had previously agreed to this

form of communication with their healthcare provider and were required to respond 'yes' to request a telephone call-back. Patients were telephoned by the practice receptionist who briefly explained the purpose and requirements of the study. Patients agreeing to participate were offered an appointment in a designated breath-testing clinic. The purpose of the study was carefully explained to patients both verbally and within an approved patient information sheet prior to enrolment. All patients were told that the breath test will potentially be used in the future to detect gastrointestinal cancers, but that the current study was intended to investigate the process and feasibility of breath testing only.

Feasibility and acceptability of breath testing

Feasibility and acceptability of breath testing in primary care among staff and patients was assessed using a mixed-methods approach.

In phase 1, it was important to identify and overcome in real time, barriers to breath testing in primary care based on challenges faced by staff administering the test. Field notes were used to document weekly events and to inform Plan-Do-Study-Act (PDSA) cycles. 'Plan' involved creation of a weekly recruitment strategy accounting for surgery-specific considerations, for example, half-days and room availability. 'Do' consisted of sampling, for which investigators (GW and the lead CSO) had daily contact with CSOs and recorded verbal feedback of any recruitment, sampling or logistical problems and their solutions. 'Study' was weekly review of this process. 'Act' was achieved by planning with CSOs how to overcome barriers for the subsequent week.

A teleconference and subsequent focus group were held with CSOs after one and 6 months of study initiation, respectively. These events were used to explore feasibility and acceptability of the testing process, from the viewpoint of the CSOs. The teleconference was an unstructured CSO-led conversation and feedback session (six CSOs and GW). The focus group (12 CSOs: 1 male, 11 females and GW) consisted of a brief presentation summarising study progress, then a minimally structured CSO-led discussion regarding perceived feasibility, acceptability, challenges and mitigation strategies, lasting 1 hour. All CSOs working on the study were invited by email to participate, therefore representing a convenience sample, at St Mary's Hospital London. The focus group was led by GW (study lead) who was known to participants. The focus group was video recorded and later transcribed. Acquired transcripts were subject to thematic analysis to identify primary themes.²⁰ Representative quotes were selected manually to illustrate the themes identified. Finally, questionnaires were given to participating GPs to complete anonymously. Likert style questions focused on their opinions around study design and logistics, with open questions regarding the remit of breath testing in primary care (online supplemental file S7).

In phase 2, patient acceptability questionnaires were used to explore opinions about the process, equipment and concept of the breath test (online supplemental file S8). The design was influenced by other established questionnaires, using Likert scales.^{21 22}

Breath sampling and quality control

Prior to enrolling patients, staff were required to attend one of three training days at either St Mary's hospital (October or November 2016) or Marylebone Health Centre (November 2017). During these sessions staff received study-specific training regarding patient enrolment and breath sample collection and handling.

Patients were not required to follow any specific conditions, such as fasting, prior to breath sampling. Before collecting breath samples CSOs explained the breath test procedure to patients. Breath samples were collected using the ReCIVA CE-marked handheld breath sampling device (Owlstone, Medical, Cambridge, UK). The standardised method for breath sampling using this device has been previously published.²³ Breath (500 mL) was collected on to a single thermal desorption (TD) tube (Markes International, Llantrisant, UK) packed with Carbograph/Tenax sorbent. The three remaining TD tube positions within the ReCIVA device were occupied by blank tubes. Inhaled ambient air was decontaminated by passing through an activated charcoal filtration column before being entrained via a tightly fitting facemask.

To maintain breath sampling quality, CSOs were trained to monitor expiratory volume and CO₂ traces during testing. If the traces were interrupted, they optimised the mask seal, or restarted the software, documenting any problems encountered.

Sealed TD tubes were stored within an airtight container and couriered weekly between the laboratory at St Mary's hospital (Imperial College London) and the primary care practices. All samples and clinical data were anonymised with no ability to retrospectively trace patients.

TD tubes were analysed using proton transfer reaction time of flight mass spectrometry (PTR-ToF-MS; Ionicon Analytik GmbH, Innsbruck, Austria) or gas chromatography mass spectrometry (GC-MS; Agilent Technologies, Cheshire, UK) in accordance with previously developed standardised methods.^{24 25} Standard quality control procedures for instruments and equipment were implemented^{24 26} (online supplemental file S9). Breath samples within TD tubes were evaluated for quality based on detected levels of acetone and isoprene (online supplemental file S10). Acceptable thresholds for acetone and isoprene were dependent on analytical platform.

Finally, quantitative data were collected throughout the study recording TD tube transport, processing and analysis times as well as the content and quality of breath VOCs.

Patient and public involvement

Patients, nurses and GPs were engaged in the study design, recruitment methodology and running of this study on a

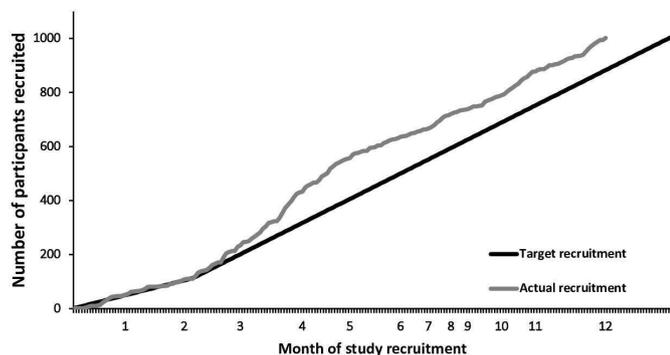


Figure 1 Total MAGIC study recruitment. Each point on x-axis represents intended sampling days only, hence uneven month distribution. MAGIC, Methodological Approaches towards a Gastrointestinal Cancer Breath Test.

daily basis. Their experiences and preferences were the material used for weekly PDSA cycles, and more formal feedback was gathered from questionnaires and the focus group, guiding changes in methodology.

RESULTS

Recruitment was successful, reaching 1002 patients within 192 of 260 allocated sampling days (figure 1). Patient demographics and reported symptoms are presented in table 1. Verification of patients against eligibility criteria found concordance in 998 (96.6%) cases. Four patients who were aged >90 year at the time of breath sampling breached eligibility criteria and were excluded.

Table 1 Demographics of eligible patients and characteristics of reported symptoms

	All patients N=998	Phase 1 N=633	Phase 2 N=365
Age, years (range)	59.7 (18–90)	59.3 (18–90)	58.8 (18–90)
Sex			
Male	409 (41%)	244 (39%)	165 (45%)
Female	578 (58%)	385 (61%)	193 (53%)
Unrecorded	11 (1%)	4 (1%)	7 (2%)
Race			
Caucasian	599 (60%)	335 (53%)	264 (72%)
Asian/Asian British	189 (19%)	161 (25%)	28 (8%)
Black/African/Caribbean/Black British	100 (10%)	73 (12%)	27 (7%)
Arab	30 (3%)	17 (3%)	13 (4%)
Other	60 (6%)	31 (5%)	29 (8%)
Unrecorded	20 (2%)	16 (3%)	4 (1%)
Current smoker	120 (12%)	72 (11%)	48 (13%)
Oral intake <5 hours	798 (80%)	458 (72%)	340 (93%)
Duration of main symptom(s)			
Today	234 (25%)	139 (22%)	95 (26%)
Recently (within 8 weeks)	351 (38%)	241 (38%)	110 (30%)
Chronic	172 (19%)	112 (18%)	60 (16%)
Unrecorded	241 (26%)	141 (22%)	100 (28%)
Patients reporting	N=921*	N=586	N=335
≥1 UGI symptom	822 (89%)	533 (91%)	289 (86%)
≥1 LGI symptom	608 (66%)	397 (68%)	211 (63%)
Single symptom reported	165 (18%)	110 (19%)	55 (16%)
UGI symptom(s) warranting urgent referral†	152 (17%)	98 (17%)	54 (16%)
UGI symptom(s) warranting non-urgent referral†	306 (33%)	199 (34%)	107 (32%)
LGI symptom(s) warranting urgent referral†	289 (31%)	178 (30%)	111 (33%)

*Symptoms were unrecorded in 77 patients, however for 44 of these patients the 'duration of symptoms' was recorded.

†Symptom(s) warranting urgent direct access endoscopic or radiological referral or urgent 2WW referral as per NICE guidelines for UGI (including pancreatic cancer) and LGI cancer.^{8,9}

LGI, lower gastrointestinal; NICE, National Institute for Health and Care Excellence; UGI, upper gastrointestinal; 2WW, 2 week wait.

Table 2 Patient engagement methods for phase 1

	Face-to-face same day	Face-to-face prebooking	Telephoning*	SMS messaging
No of GP practices using a given method†	8	2	2	8
Total no of sampling days that the given method was used‡	68	15.5	15.5	81
Eligible patients telephoned/sent SMS message*	–	–	114	2653
Patients booking an appointment	–	81	68	345
Patients attending booked appointments§	206 (100%)	69 (85%)	57 (84%)	301 (87%)
Patients recruitment per sampling day (mean)	3.0	4.5	3.7	3.7
Patients recruitment per practice per day (mean)	0.37	2.25	1.84	0.46

*Unanswered/wrong number calls are unrecorded.

†A total to 16 practices contributed to recruitment during phase 1, four practices used a combination of face-to-face enrolment and either telephoning or SMS messaging, hence they are counted twice for the purposes of this table.

‡Recruitment could occur at two GP practices at any one time, hence, for the purposes of this table only, sampling days could be counted twice, hence total is >192 days.

§Values in parenthesis represent percentage of patients agreeing to breath testing who actually completed the test.

GP, general practitioner; SMS, Short Message Service.

Methods of patient engagement

Four methods of patient engagement were assessed in phase 1: face-to-face same day, face-to-face prebooking, telephoning and SMS messaging. During phase 2, SMS messaging was used exclusively for initial patient engagement. Details of patient accrual for each of the four engagement methods are presented in [table 2](#).

The percentage of patients who completed the breath test after agreeing to be tested ranged from 84% to 100% depending of the method of initial engagement. Where patients either opted or were required to prebook a breath test, test completion rates tended to be lower reflecting a ‘drop-out’ rate of between 15% and 18%.

Methods of recruitment

During phase 1 (‘single practice’ recruitment), 633 eligible patients were recruited over a total of 119 sampling days (average 5.3 patients per day). In phase 2 (‘hub and spoke’ and ‘single practice’ recruitment), 365 eligible patients were recruited over a total of 73 sampling days (average 5.0 patients per day). For the ‘hub and spoke’ model alone, recruitment averaged 4.3 patients

per day ([table 3](#)). During phase 2, patient recruitment using the ‘single practice’ model was maintained at 5.3 patients per day.

When normalised to number of GP practices contributing to patient recruitment for each recruitment method within both phase 1 and phase 2, the average number of patients accrued per centre per day was higher for the ‘hub and spoke’ compared with ‘single practice’ method (0.61 vs 0.28) ([table 3](#)).

Feasibility and acceptability of the breath testing process

Patient recruitment

Twenty-five healthcare professionals were successfully trained to sample breath, showing feasibility of this task for a wide range of operators. Feedback obtained from field notes and the CSO-led teleconference and focus group regarding the advantages and challenges of recruitment and engagement methods are summarised in [table 4](#).

Patient accrual rate was initially low, due to a number of recognised challenges: inconsistent referral of patients, inefficient use of CSO time, technical problems and mismatch of CSO and GP schedules. Full details of reported challenges to breath testing and mitigation strategies are provided as an online supplemental file S11. Following iterative refinement of the approach to patient accrual and breath testing there was a marked acceleration in recruitment during months two to five of the study ([figure 1](#)). This was likely due to improved CSO familiarity with equipment and study procedures over time, as well as the dynamic and adaptable study design, driven by weekly PDSA cycles, which allowed early recognition of problems and development of solutions. Dedicated breath testing clinics were set up to sample all patients who had entered the study via face-to-face prebooking, phone or SMS recruitment. This was an efficient and effective strategy that enabled testing for up to 12 patients per half day ([table 4](#)). With only one site

Table 3 Patient recruitment methods, phase 1 and phase 2 combined

	Single practice model	Hub and spoke model
Number of GP practices	19	7
Total number of sampling days	168	24
Total number of patients recruited	895	103
Patient recruitment per day (mean)	5.3	4.3
Patient recruitment per practice per day (mean)	0.28	0.61

GP, general practitioner.

**Table 4** Feedback obtained from focus group and field notes**Face to face same-day/face-to-face prebooking**

Positive	<ul style="list-style-type: none"> ▶ Method of recruitment open to all GP practices ▶ Easy to organise as reliant on direct patient interaction (without need for telephoning or SMS messaging) ▶ No requirement for administrative staff ▶ Face-to-face same-day: convenient for patients as no separate visit needed ▶ Face-to-face prebooking: allowed patients to be brought back at a time convenient for them
Negative	<ul style="list-style-type: none"> ▶ Reliant on GP engagement: CSOs having to 'remind GPs 2–3 times per morning' with some GPs admitting to 'forgetting to send in patients'. ▶ Slower recruitment in smaller (less busy) practices ▶ Inefficient: CSOs present all day for a mean yield of approximately three patients.

Telephoning

Positive	<ul style="list-style-type: none"> ▶ Method of recruitment open to all GP practices ▶ Appointments available for up to 12 patients per half day 'breath clinic'
Negative	<ul style="list-style-type: none"> ▶ Requires support of administrative staff to contact patients ▶ Administrative staff only able to give general information about study when calling patients ▶ Cost of telephoning (including staff time)

SMS messaging

Positive	<ul style="list-style-type: none"> ▶ More efficient than telephone booking as patients who were telephoned had already expressed interest in being involved in breath testing by responding to SMS message. This led to higher booking rates. ▶ Appointments available for up to 12 patients per half day 'breath clinic' ▶ Patient initial identification not reliant on attendance to GP. ▶ Potentially more convenient for patients.
Negative	<ul style="list-style-type: none"> ▶ Only open to GP practices with ability to send SMS messages to patients. ▶ Proportion of patients who were ultimately tested was lower than for other enrolment methods (approximately 10% uptake). ▶ Identification of patients based on database searching has the potential to be less reliable and could vary between different practices. ▶ Requires support of administrative staff to contact patients. ▶ Administrative staff only able to give general information about study when calling patients. ▶ Cost of telephoning (including staff time).

Single practice**Hub and spoke**

Positive	<ul style="list-style-type: none"> ▶ Patient convenience in attending own GP practice ▶ Allows for face-to-face same-day enrolment 	<ul style="list-style-type: none"> ▶ Broader recruitment cohort ▶ Fewer CSOs required (two for single hub): reduced CSO training time and potential improved consistency and quality of sampling. ▶ More flexibility for patients wanting to book an appointment and more efficient for CSOs to collect samples.
Negative	<ul style="list-style-type: none"> ▶ Narrower recruitment cohort ▶ Larger no of CSOs required, with less efficient use of their time 	<ul style="list-style-type: none"> ▶ Some patients may either not wish to or be able to travel to the central hub for testing ▶ Allocation of appointments between multiple practices, meaning that there was a requirement for a central booking system

CSO, clinical study officer; GP, general practitioner; SMS, Short Message Service.

being used for sampling, fewer staff and less equipment was required, transport and logistics were easier to coordinate, and bulk collection lowered courier costs. After recruitment acceleration, the rate then stabilised and was maintained, even after the integration of the 'hub and spoke' model. This finding reveals that testing patients at a centralised site does not negatively affect recruitment. It also indicates that a dynamic and responsive study design may be an effective strategy for primary care studies like this, as recruitment was maintained despite using 26 practices all with different environments and clinic schedules (online supplemental file S11). These findings and the lessons learnt during recruitment led to the development of a flowchart of recommendations for improving

recruitment in primary care studies (online supplemental file S12).

Acceptability of the test GP perspective

Twenty-one GPs, from 10 of the 26 participating practices, answered the GP specific questionnaire. Nine out of 10 GPs reported that asking patients to participate, sending them through to the CSO/nurse, answering patient questions and general logistics of breath testing was 'very easy' or 'easy'. Perceived barriers to participation were 'time constraints' (clinical staff and patients') and the fact that this was a research study where individual patients were not intended to directly benefit from test results. All GP

Table 5 Summary of patient acceptability questionnaire responses (n=365, phase 2)

	Very easy/very comfortable	Easy/comfortable	Difficult/uncomfortable	Very difficult/very uncomfortable	Not applicable
How easy was it to do the breath test? (%)	79	20	1	0	0
How comfortable were you while wearing the face mask? (%)	55	44	1	0	0
How did you find the experience of holding the device during the test? (%)	42	55	2	0	1
Would you be comfortable to do the breath test again, if recommended to by a doctor? (%)	64	35	0	0	1
	Too long	Acceptable amount of time	Too quick	–	Not applicable
What did you think about the time it took to give a breath sample? (%)	2	95	3	–	0
	Strongly encourage	Encourage	Discourage	Strongly discourage	Not applicable
Would you encourage family and friends who were offered a breath test to complete it? (%)	59	39	0	0	2
Free text comments regarding overall sampling experience					
'Nothing to improve because there is nothing to it. It's nice'					
'I found the breath test to be extremely satisfactory, I am happy to participate in more research'					
'I found it fine as it is. And, it was a rather nice experience. I liked it.'					

respondents reported that they had 'no concerns about the study' from their patients. GPs' opinions about the potential place of a breath test in clinical practice are detailed in online supplemental file S13.

Patient perspective

During phase 2, all 365 eligible patients completed acceptability questionnaires, providing overwhelmingly positive feedback for the breath test (table 5). Of those patients recruited using the 'hub and spoke model', only one (0.3%) commented that they found travelling to a different GP practice inconvenient. The breath test was also acceptable to patients with a wide variety of medical problems, including 197 patients with either asthma, chronic obstructive pulmonary disease, or other lung diseases. Thirteen (3.5%) patients suggested that a hands-free breath sampler would be preferable. This comment was offset by others saying they liked being 'in full control of the mask'. Despite CSOs being asked to inform patients that masks were sterile and single-use, and to open masks in front of patients, three (0.8%) patients enquired about sterility of the mask. This therefore reflected an explanation/execution issue rather than an equipment issue.

Breath sampling and quality control

Although there were minimal patient related limitations, technical issues with sampling equipment were reported. A summary of themes regarding feasibility and acceptability of the sampling process is detailed in online supplemental file S14. Problems were frequently solved by restarting or updating the computer software for the ReCIVA device. When such measures failed, CSOs

resorted to collecting breath as 'timed samples' where patients were asked to breathe into the ReCIVA for 5 min without using the device's software. This meant that the volume and flow rate of breath sampling was uncontrolled. 'Timed samples' accounted for 87 (13.7%) of the 633 eligible samples collected during phase 1 and occurred primarily at the start of study when CSOs lacked experience using the ReCIVA. In comparison during phase 2 of the study, when study logistics and methodology had been optimised, only 7 (1.9%) 'timed samples' were collected out of a total of 365 eligible samples. During the final 6 weeks of sampling there were no reported equipment failures. CSOs did not report any issues with TD tube storage or transport.

On average breath samples were analysed within 2.8 (range 0–11) days of collection. Eighty-three (13%) samples collected during phase 1 of the study were stored at -80°C for up to 13 days before analysis as a result of instrument downtime. The collection to analysis time was therefore prolonged for these samples, averaging 8.8 (range 3–14) days. There was no instrument downtime in phase 2 of this study, therefore, no storage of breath samples at -80°C was required. Twenty-six phase 2 GC-MS samples were lost due to a GC-MS instrumental error.

Breath samples were analysed by PTR-ToF-MS (n=316) and GC-MS (n=23) in Phase-2 of MAGIC. Three hundred (95%) of those analysed by PTR-ToF-MS and 21 (91%) of those analysed by GC-MS were deemed to contain adequate quantities of breath.

DISCUSSION

The analysis of VOCs within exhaled breath offers a non-invasive approach to the detection of a number of diseases including gastrointestinal cancers.¹¹ Such a test could be offered in primary care to patients presenting with non-specific symptoms that do not meet existing guidelines for referral. However, before a large phase-III clinical trial can be conducted in primary care, it is necessary to first understand the feasibility and acceptability of the breath test in this setting. The current study was designed to evaluate different recruitment and engagement strategies for breath testing in primary care. Phase 1 evaluated different engagement methods in addition to discovering optimum organisation and implementation strategies. Phase 2 was used to evaluate patient acceptability of the test, with the rationale that acceptability could only be assessed after optimisation of delivery during phase 1. The emphasis on patient recruitment during phase 1 meant a greater number of patients were accrued during this period. This was not however felt to be detrimental to the findings of phase 2.

This study showed that both sampling in a single GP practice as well as the centralised hub-and-spoke model of referral were viable and acceptable to patients and study staff. It was hypothesised that attendance and attitudes towards the breath test may be negatively affected by having to travel to a central location. However in this study it was observed that centralising breath testing reduced staffing and equipment requirements with no discernible negative impact on patient feedback. Transport and logistics were easier from one single location, and bulk collection lowered courier costs. In terms of organisation within primary care services, a breath test is comparable to a blood test. If we consider breath testing as a complete service, where the testing, results and any referrals to secondary care were managed as a streamlined pathway, we could draw comparisons to other centralised services such as diabetes care, which lowers costs.²⁷ The hub-and-spoke model evaluated in phase 2 of this study explored the concept of testing patients in a central location, in this instance a GP practice. Finding may be applicable to other centralised testing centres such as diagnostic centres and hospitals.

Four methods of engagement were evaluated in phase 1. Each was deemed to be feasible and acceptable. The method of enrolment adopted in future trials and ultimately clinical practice will largely reflect the intended purpose of the test (eg, triaging symptomatic patients or screening asymptomatic populations). SMS messaging, and to a lesser extent telephoning, has the potential to reach large numbers of patients. However, as highlighted, this approach may only result in 10%–50% of patients being assessed, akin to population screening. Alternatively, opportune identification of patients by GPs may be more representative of a targeted triage test that could be used as an adjunct to the existing 2WW referral pathway. A flow chart of how to optimise patient recruitment in

primary care studies, taken from lessons learnt during this study, is detailed in online supplemental file S12.

The breath test received almost universal acceptance. The overwhelming majority of patients found the test easy to complete, with wide representation from patients of different age, gender, comorbidity and ethnicity. Selection bias may, however, have influenced findings given that enrolled patients were those who were more likely to seek medical attention and engage with medical research. Although gastrointestinal cancers are more common in men, a greater number of women participated in this study, possibly influencing results. Dutch data reported that women are 18% more likely to consult their GP than men after adjustment for gender-specific factors.^{28 29} The focus group was also predominantly female, potentially influencing results.

During the last 6 weeks of the study, after optimisation of sampling methodology and consolidation of staff training, technical failures of breath collection were eliminated. Analysis of breath samples within a central laboratory was achieved with established quality control procedures to ensure instrumental consistency.¹¹ Ninety-five per cent of all samples that were analysed were deemed to contain adequate quantities of breath. For implementation of breath testing on a wider scale, standardisation across different laboratories is required. Alternatively point of care devices could be developed to streamline the analysis and receipt of test results.

It has been previously highlighted that time and financial pressures can be a major barrier to conducting high-quality research in primary care.¹⁸ Importantly GPs and research staff were supportive of conducting breath research in primary care. Patient enrolment and sampling using SMS messaging and a central sampling hub helped to reduce the workload of GPs as they were no longer responsible for identifying and approaching potentially eligible patients. Access to research nurses from the NIHR likewise helped to minimise additional burdens to GP services during study recruitment. GP practices also received a modest financial incentive, as they were remunerated for every patient recruited by the NIHR, at a rate of £20 and £25 per patient for phases 1 and 2, respectively. This may have encouraged participation and provided some recognition for the additional workload caused by the study. However, these factors may not apply outside of the research setting, potentially influencing acceptability of breath testing to GPs, particularly where responsibility for implementing testing, interpreting and actioning results may fall to them.

No previous study has sought to define how breath testing can be successfully integrated into primary care with the engagement of both patients and clinical stakeholders. Strengths of this study were its two phased design and concurrent iterative mixed methods approach. Limitations were that the demographics and views of patients who did not respond/agree to breath testing were not recorded. Such information would have been valuable in determining barriers to patient's participation. The fact that this was a research study without

direct clinical benefit to patients may have contributed to patients declining to participate. The rate of uptake of the test within the target population, and influencing factors, while not the focus of the current study, should nevertheless be clarified in future studies. A broader assessment of the opinions of key stakeholders may have established greater consensus as to the role of the breath test in clinical practice as well as challenges to its adoption. Patients were not required to follow any specific conditions prior to the breath test, as there are currently no evidence based guidelines for sampling breath in clinical practice. This means that the study may not be fully representative of a future breath testing pathway. Finally, this study was not designed to assess the diagnostic performance of the breath test.

This study determined that it was feasible to collect and conduct high-quality analysis of large numbers of breath samples from primary care. This provides encouraging new evidence to support the use of wide-scale breath testing in this setting. In parallel to existing and ongoing diagnostic accuracy and standardisation studies, breath testing appears to be a feasible and acceptable and an accurate method of assessing patients with unexplained gastrointestinal symptoms. This study provides a practical framework to guide the design of larger phase 3 trials examining the performance of the proposed breath test in primary care. The design and methodology can also be applied to other large-scale primary care studies, particularly as it provides valuable insights as to how to optimise recruitment in this well-known challenging research sector.

Acknowledgements The authors wish to acknowledge the invaluable input and expertise of Professor Wendy Atkin, who supervised this project until she very sadly passed away in December 2018. The authors wish to thank Lucinda Hetherington (Northwest London Clinical Research Network) who was the lead clinical study officer for the MAGIC study and coordinated the study at multiple primary care practices.

Contributors All authors made substantial contributions to the work; specifically GW and GBH conceived the work with IB, PRB, CvW and AJC making substantial contributions to the design, CvW in particular with regards to the patient acceptability questionnaire. GW planned and executed the study process. GW, IB, PRB, AW and MF contributed to data acquisition and analysis, with interpretation overseen by CvW, AJC and GBH. GW, IB and PRB drafted the first manuscript version, where all authors revised it critically and then approved the final version. All authors can account for the integrity of this work.

Funding The research was supported by awards from Rosetrees and Stonegate Trusts via Imperial College Charity (grant number 171812) and by infrastructure funding from the NIHR London In Vitro Diagnostic Cooperative programme (grant number P67522). Piers R Boshier was funded by the NIHR (grant number not applicable).

Disclaimer The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval Ethical approval was obtained from the Camden & Kings Cross Research Ethics Committee (14/LO/1136).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on reasonable request. Relevant anonymised data are included in the article or uploaded as online supplemental information. Additional anonymised data are available on reasonable request.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been

peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Georgia Woodfield <http://orcid.org/0000-0003-4958-6842>

Christian von Wagner <http://orcid.org/0000-0002-7971-0691>

REFERENCES

- 1 Cancer Research UK. Bowel cancer survival statistics - by stage, 2018. Available: <https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer/survival#heading=Three> [Accessed August 8, 2020].
- 2 Cancer Research UK. Oesophageal cancer statistics. oesophageal cancer survival, 2017. Available: <https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/oesophageal-cancer#heading=Two> [Accessed August 20, 2020].
- 3 Barwick TW, Scott SB, Ambrose NS. The two week referral for colorectal cancer: a retrospective analysis. *Colorectal Dis* 2004;6:85–91.
- 4 Gibbs JF, Rajput A, Chadha KS. The changing profile of esophageal cancer presentation and its implication for diagnosis. *J Natl Med Assoc* 2007;99:620–6.
- 5 Public Health England. National cancer intelligence network: routes to diagnosis 2015. Available: http://www.ncin.org.uk/publications/routes_to_diagnosis [Accessed January 17, 2017].
- 6 Elliss-Brookes L, McPhail S, Ives A, et al. Routes to diagnosis for cancer - determining the patient journey using multiple routine data sets. *Br J Cancer* 2012;107:1220–6.
- 7 Smyth EC, Lagergren J, Fitzgerald RC, et al. Oesophageal cancer. *Nat Rev Dis Primers* 2017;3:17048.
- 8 National Institute of Clinical Excellence. Suspected cancer: recognition and referral. NICE guideline [NG12] UK 2015 [updated July 2017]. Available: <https://www.nice.org.uk/guidance/ng12/chapter/1-Recommendations-organised-by-site-of-cancer#lower-gastrointestinal-tract-cancers> [Accessed May 21, 2020].
- 9 National Institute of Clinical Excellence. Suspected cancer: recognition and referral. NICE guideline [NG12] UK 2015 [updated July 2017]. Available: <https://www.nice.org.uk/guidance/ng12/chapter/1-Recommendations-organised-by-site-of-cancer#upper-gastrointestinal-tract-cancers> [Accessed August 20, 2020].
- 10 Adam ME, Fehervari M, Boshier PR, et al. Mass-Spectrometry analysis of Mixed-Breath, Isolated-Bronchial-Breath, and Gastric-Endoluminal-Air volatile fatty acids in esophagogastric cancer. *Anal Chem* 2019;91:3740–6.
- 11 Hanna GB, Boshier PR, Markar SR. Accuracy and methodologic challenges of volatile organic Compound-Based exhaled breath tests for cancer diagnosis: a systematic review and meta-analysis. *JAMA oncology* 2018:e182815 (published Online First: 2018/08/22).
- 12 Markar SR, Chin S-T, Romano A, et al. Breath volatile organic compound profiling of colorectal cancer using selected ion Flow-tube mass spectrometry. *Ann Surg* 2019;269:903–10.
- 13 Markar SR, Wiggins T, Antonowicz S, et al. Assessment of a noninvasive exhaled breath test for the diagnosis of oesophagogastric cancer. *JAMA Oncol* 2018;4:970–6.
- 14 Markar SR, Brodie B, Chin S-T, et al. Profile of exhaled-breath volatile organic compounds to diagnose pancreatic cancer. *Br J Surg* 2018;105:1493–500.
- 15 Durán-Acevedo CM, Jaimes-Mogollón AL, Gualdrón-Guerrero OE, et al. Exhaled breath analysis for gastric cancer diagnosis in Colombian patients. *Oncotarget* 2018;9:28805–17.
- 16 Tognoni G, Alli C, Avanzini F, et al. Randomised clinical trials in general practice: lessons from a failure. *BMJ* 1991;303:969–71.
- 17 Bell-Syer SEM, Thorpe LN, Thomas K, et al. Gp participation and recruitment of patients to RCTs: lessons from trials of acupuncture



- and exercise for low back pain in primary care. *Evid Based Complement Alternat Med* 2011;2011:1–7.
- 18 Foy R, Parry J, McAvoy B. Clinical trials in primary care. *BMJ* 1998;317:1168–9.
- 19 O'Dowd T. Research in general practice: who is calling the tune? *Br J Gen Pract* 1995;45:515–6.
- 20 Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006;3:77–101.
- 21 National Institute of Health Research. Faecal immunochemical testing (fit) for adenoma surveillance. fit for follow-up: National Institute of health research, 2019. Available: <https://www.journalslibrary.nihr.ac.uk/programmes/hta/0922192#/> [Accessed February 3, 2019].
- 22 Cross AJ, Wooldrage K, Robbins EC, et al. Faecal immunochemical tests (fit) versus colonoscopy for surveillance after screening and polypectomy: a diagnostic accuracy and cost-effectiveness study. *Gut* 2019;68:gutjnl-2018-317297:1651–2.
- 23 Doran SLF, Romano A, Hanna GB. Optimisation of sampling parameters for standardised exhaled breath sampling. *J Breath Res* 2017;12:016007.
- 24 Romano A, Doran S, Belluomo I, et al. High-Throughput breath volatile organic compound analysis using thermal desorption proton transfer reaction time-of-flight mass spectrometry. *Anal Chem* 2018;90:10204–10.
- 25 Chin S-T, Romano A, Doran SLF, et al. Cross-platform mass spectrometry annotation in breathomics of oesophageal-gastric cancer. *Sci Rep* 2018;8:5139.
- 26 Romano A, Hanna GB. Identification and quantification of VOCs by proton transfer reaction time of flight mass spectrometry: an experimental workflow for the optimization of specificity, sensitivity, and accuracy. *J Mass Spectrom* 2018;53:287–95.
- 27 Honkasalo MT, Linna M, Sane T, et al. A comparative study of two various models of organising diabetes follow-up in public primary health care - the model influences the use of services, their quality and costs. *BMC Health Serv Res* 2014;14:26.
- 28 Vedsted P. [Gender differences in the use of health care system]. *Ugeskr Laeger* 2007;169:2403–8.
- 29 Jørgensen JT, Andersen JS, Tjønneland A, et al. Determinants related to gender differences in general practice utilization: Danish diet, cancer and health cohort. *Scand J Prim Health Care* 2016;34:240–9.

SUPPLEMENTARY FILES

Methodological Approaches towards a Gastrointestinal Cancer (MAGIC) breath test in primary care

Authors:

Georgia Woodfield, Ilaria Belluomo, Piers R Boshier, Annabelle Waller, Maya Fayyad, Christian Von Wagner, Amanda J Cross, George B Hanna

SUPPLEMENTARY FILE 1: Names of 26 GP practices participating in the MAGIC breath test study:

Phase 1:

1. Aksyr Medical Practice, NW10 8RY
2. Oxgate Gardens Surgery NW26EA
3. Hillcrest surgery, W3 9RA
4. Dr Jefferies and Partners, Fulham, SW6 6BQ
5. The Law Medical Group practice Willesden NW10 5UY
6. The Law Medical group practice Harrow HA96QQ
7. The Gill medical practice, Feltham TW14 0AB
8. Grove Park Terrace Surgery Chiswick W4
9. The Bush Doctors W12 8PP
10. Twickenham Park Medical Centre, TW13 6HD
11. Buckingham Road Surgery NW10 4RR
12. Fulham Medical Centre, SW6 1BG
13. Acre Surgery, HA6 1TQ
14. Gladstone Medical Centre, NW2 6JH
15. Cuckoo Lane Practice, Hanwell, W7 1DR
16. Wembley Park Medical Centre, Wembley, HA9 8HD

Phase 2:

17. Pimlico Health, SW1V 3EB
18. Lonsdale Medical Centre, NW6 6RR
19. The Good practice, SW10 0LR

7 practices as part of Central London Healthcare (CLH) GP federation:

20. Woodfield Road Medical Centre, W9 3XZ
21. Covent Garden Medical Centre, WC2H 9AA
22. Cavendish Health Centre, W1G 9TG
23. Marylebone Health centre, NW1 5LT
24. Fitzrovia Medical Centre, W1T 6EU
25. Newton Medical Centre, W2 5LT
26. Crawford Street Surgery, W1H 2HJ

SUPPLEMENTARY FILE 2: Summary of National Institute for Health and Care Excellence (NICE) guidelines for gastrointestinal (GI) cancer referral 2016, available at:

<https://www.nice.org.uk/guidance/ng12/chapter/1-Recommendations-organised-by-site-of-cancer#upper-gastrointestinal-tract-cancers>

Upper GI cancers

Two week wait (2WW) direct access oesophago-gastro-duodenoscopy (OGD) for:

- 1) Dysphagia
- 2) Age >55 years with weight loss AND upper abdominal pain/reflux/dyspepsia

Non urgent direct access OGD:

- 1) Haematemesis
- 2) Age >55 years with
 - persistent dyspepsia OR
 - upper abdominal pain WITH anaemia OR
 - raised platelets with nausea/vomiting/weight loss/reflux/dyspepsia/upper abdominal pain OR
 - nausea/vomiting with weight loss/reflux/dyspepsia/upper abdominal pain

2WW computerised tomography scan/abdominal ultrasound scan for:

-Abdominal mass
(stomach and gallbladder and liver cancers)

Pancreatic cancer

2WW appointment for:

- Age >40 years with new jaundice

2WW CT scan/Ultrasound scan:

- Age >60 years AND weight loss AND diarrhoea/back pain/abdominal pain/nausea/vomiting/constipation/diabetes

Colorectal cancers

2WW appointment for:

- 1) Age >40 years with weight loss and abdominal pain
- 2) Age >50 years with rectal bleeding
- 3) Age >60 years with anaemia/change in bowel habit/positive faecal occult blood test
- 4) Abdominal mass
- 5) Age <50 years with rectal bleeding AND abdominal pain/ change in bowel habit /weight loss/anaemia

SUPPLEMENTARY FILE 3: Poster for GPs

Imperial College
London

Imperial College Healthcare 

Breath Testing for Gastrointestinal Disease 2016- 2017

- Please ask all patients with **current/recent** (within 2months) **gastrointestinal symptoms** if they would consider speaking to a research nurse about this study.
- Patients are also eligible with **chronic gastrointestinal problems**, even if controlled on medication.
- Patients must be >18 or <90 years of age, and be able to speak to a research nurse in order to consent to giving a breath sample.

We are testing the feasibility of breath testing for the diagnosis of GI disease.

- **ANY/ALL gastrointestinal symptoms are accepted in this study.** These could include abdominal pain, diarrhoea, constipation, reflux/dyspepsia, dysphagia, change in bowel habit, nausea, vomiting, weight loss, anaemia, GI bleeding, jaundice, or others.

- The study involves giving a breath sample and will take 5-10minutes.
- A research nurse will explain the breath test to the patient and consent them.

Please advise patients to ask at reception to see the clinical research nurses who will be based in the GP practice, or please email: ichc-tr.breathtest@nhs.net with the patient details.

Primary Care Poster v.3.0 16/12/2016. Dr G Woodfield, Prof G Hanna

SUPPLEMENTARY FILE 4: Letter for GPs

Imperial College
London


**National Institute for
Health Research**

Clinical Research Network
North West London

Division of Surgery
Department of Surgery and Cancer
10th Floor, QEQM Building
St Mary's Hospital
Praed Street, London, W2 1NY

Tel: +44 (0)20 3312 2125
Fax: +44 (0) 020 3312 6309

g.hanna@imperial.ac.uk
www.imperial.ac.uk

Professor George Hanna PhD FRCS
Head of Division of Surgery

14th December 2016

Dear North West London General Practitioners,

Thank you very much for agreeing to participate in the study:

Non-invasive testing for the diagnosis and assessment of gastro-intestinal disease – Primary Care feasibility Study. November 2016 to March 2017.

The study aims to trial the use of a breath test in patients with gastrointestinal symptoms, as a future device for detecting oesophago-gastric, pancreatic and colorectal cancer. This initial trial is a feasibility study of 500 patients. The breath test is not currently a diagnostic tool, as there is no validated "positive" or "negative" result to be gained currently, GPs and patients therefore will not be informed of results. It therefore should not influence patient management or referral pathways in any way.

Recommendation for GPs:

- Please look out for patients who have/have had GI symptoms. Symptoms include **dysphagia, weight loss, abdominal pain, reflux, dyspepsia, nausea, vomiting, diarrhoea, constipation, change in bowel habit, abdominal mass, GI blood loss, jaundice** or any other variation of GI symptoms. Patients are eligible if they have had **current/recent symptoms (within 2 months)**, OR if they have ever consulted the GP with GI symptoms, where the symptom is **chronic or requires medication to control it**. This includes all patients on frequent anti-reflux medications, laxatives or anti diarrhoeals for example. The symptom does not have to be present on the day of testing.
- Please ask these patients if they would mind talking to a research nurse about performing a breath test as part of our research study. If it is outwith the time that the research nurse is in attendance, please ask them if they would mind being contacted by phone.
- Research nurses will be stationed in practices for 1-2 week blocks. If a research nurse is in your practice that week, please send the patient to see the nurse straight away.
- If it is out of hours or on a week where the research nurse is not present, please email their Name, Phone number and Practice name to ichc-tr.breathtest@nhs.net
- The research nurse will then contact them in a few days time to ask them to come in to perform the breath test.
- Research nurses will give an information leaflet to patients, consent them, document basic medical history and perform the breath test. It should take about 5-10 minutes.
- Practices will be compensated financially for their time in sending patients to see the research nurse for a breath test, and for every patient they highlight via email for the study. (£5 per patient referred)
- Please be aware that neither GPs nor patients will get any feedback/results after their breath test, as both the technology and analysis are still in development. This is not currently a diagnostic tool.

Should you have any questions please don't hesitate to contact me. Thanks again for your cooperation with this study.

Yours sincerely,

Dr Georgia Woodfield, Clinical Research Fellow Imperial College. (g.woodfield@imperial.ac.uk)

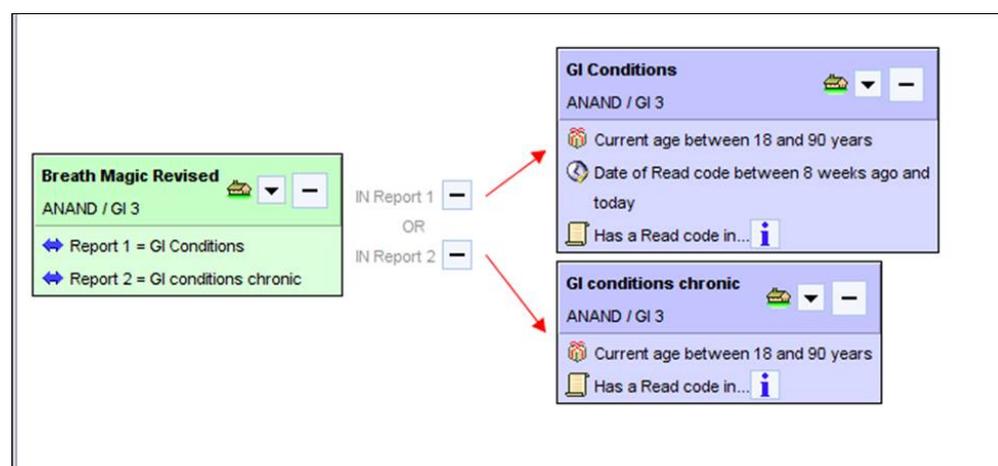
Professor George Hanna PhD FRCS, Principle Investigator

SUPPLEMENTARY FILE 5: Example of GP database searches

The following database searches were performed in order to identify patients for potential text (SMS) recruitment during phase 2 of Breath MAGIC, as part of the hub and spoke model of sampling. Of note, this search example is the most complex and thorough search done within the Breath MAGIC study. In phase 1, local GP practices did simple local database searches performed by local GP receptionists or GPs themselves based on specific symptoms or medication use. This is because only small numbers of patients were needed per practice in phase 1. The aim of this more complex search was to reach as many eligible patients as possible from 7 GP practices in the Central London Healthcare (CLH) GP federation for General Practices in Westminster, to be breath-tested at the central hub (Marylebone Health Centre). The database search was performed by central CLH administrative staff at CCG level (Ahmed Hosny and Anand Bhundia- GP Network Support Officers CLH), neither of whom worked directly in the participating practices. The large searches of CLH records for the hub and spoke sampling strategy was done as follows:

Search 1**Gastrointestinal symptoms recorded in the past eight weeks (including those coded as chronic) AND age 18-90 years inclusive**

Search 1 therefore picked up patients who fulfilled the 1st and 2nd inclusion criteria for the study (gastrointestinal symptoms today or within last 8 weeks).



gastrointestinal symptoms are represented by “i” in the diagram above. Symptoms (with database read codes) included:

Indigestion (1954.)
Abdominal pain (1969.)
Altered bowel function (19EA.)
Diarrhoea (19F2.)
Viral gastroenteritis (A07y0)
Gastro-oesophageal reflux disease with ulceration (J1020)
[D]Dysphagia (R072.)
[D]Change in bowel habit (R078.)
[D]Abdominal pain (R090.)
Gastric reflux (Ua1kQ)
Gastro-oesophageal reflux disease (X3003)
Gastritis (X301N)
Gastroenteritis (X30BN)
Nausea (X75qw)
Jaundice (X769z)
Weight loss (X76CA)
Flatulent dyspepsia (X76d5)
Campylobacter gastrointestinal tract infection (XE0QI)
Gastro-oesophageal reflux disease with oesophagitis (XE0aL)
Gastro-oesophageal reflux disease without oesophagitis (XE0aO)
Irritable bowel syndrome (XE0as)
Biliary tract disorders NOS (XE0dR)
Constipation (XE0rD)
[D]Abdominal mass (XE2nV)
Dysphagia (XM08J)
Abdominal mass (XM097)
Bacterial gastroenteritis (XM0pJ)
Nausea and vomiting (Xa1pJ)
Moderate gastric reflux (Xa7Ta)
Minimal gastric reflux (Xa7Tb)
Gastric aspirate containing blood (Xa7Tj)

Chronic conditions:

Chronic gastric ulcer (J111.)
Chronic gastrojejunal ulcer (J141.)
Chronic gastritis (J151.)
Chronic constipation with overflow (J5201)
Chronic nonspecific abdominal pain (X3062)
Chronic constipation (X30BI)
Chronic diarrhoea (X30Bn)

Search 2

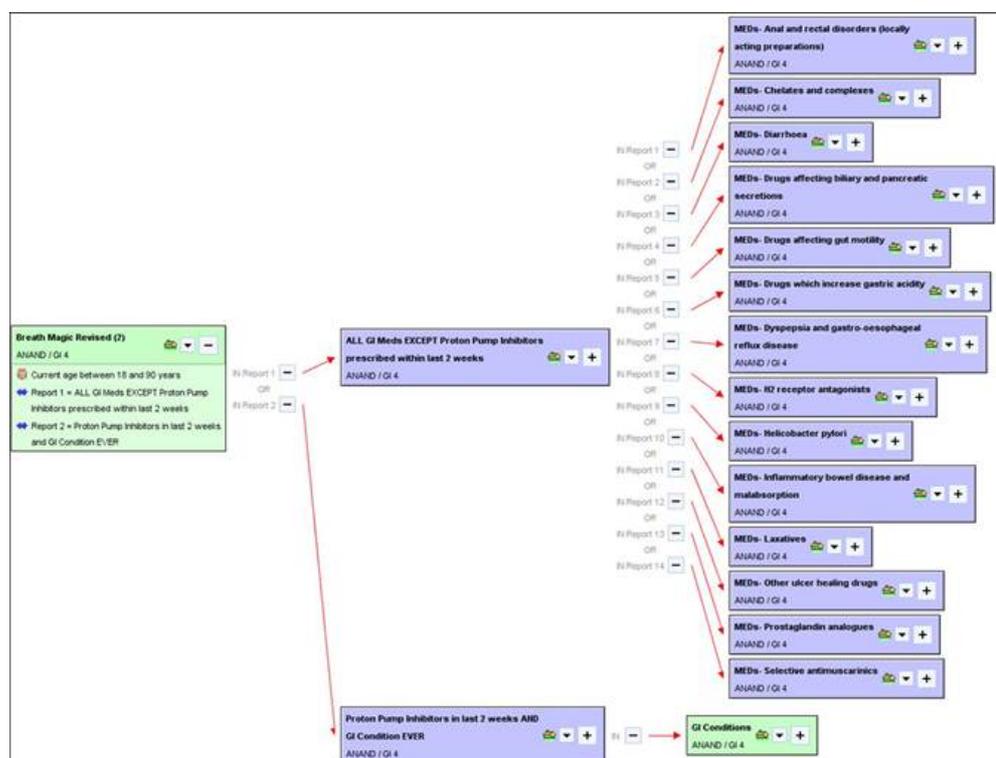
Currently on gastrointestinal medications except proton pump inhibitors (PPIs) prescribed within last two weeks AND age 18-90 years inclusive.

OR

Patients prescribed PPIs (within last two weeks) AND a recorded gastrointestinal condition at any point in their records.

Search 2 therefore picked up patients who fulfilled the 3rd inclusion criteria for the study (chronic gastrointestinal condition controlled on medication).

The caveat with PPIs was that we did not want patients who were on PPIs for non-gastrointestinal reasons; e.g. for patients who were on steroids. For this reason the search started with “All GI meds EXCEPT PPIs prescribed in last two weeks” AND “PPIs in last two weeks AND GI condition EVER”.



Gastrointestinal medications that were included as part of this search are encoded by the headings on the right of the diagram above.

Included categories and subcategories of gastrointestinal medications were:

Actions

- ↳ Gastro-intestinal
 - ↳ Dyspepsia and gastro-oesophageal reflux disease
 - ↳ Antacids and simeticone
 - ↳ Aluminium- and magnesium- containing antacids
 - ↳ Hydrotalcite
 - ↳ Antacid preparations containing simeticone or local
 - ↳ Simeticone
 - ↳ Sodium citrate
 - ↳ Raft-forming indigestion remedies
 - ↳ Alginate preparations
 - ↳ Indigestion remedies
 - ↳ Drugs affecting gut motility
 - ↳ Antimuscarinics
 - ↳ Other antispasmodics
 - ↳ Motility stimulants
 - ↳ Gastroprotection
 - ↳ H2 receptor antagonists
 - ↳ Selective antimuscarinics
 - ↳ Chelates and complexes
 - ↳ Prostaglandin analogues
 - ↳ Proton pump inhibitors
 - ↳ Other ulcer healing drugs
 - ↳ Helicobacter pylori
 - ↳ Diarrhoea
 - ↳ Adsorbent & bulk-forming drugs
 - ↳ Antimotility drugs
 - ↳ Intestinal antisecretory agents
 - ↳ Inflammatory bowel disease and malabsorption
 - ↳ Aminosallylates
 - ↳ Corticosteroids (in chronic bowel disorders)
 - ↳ Immunosuppressants (in chronic bowel disorders)
 - ↳ Food allergy
 - ↳ Other drugs to treat inflammatory bowel disease
 - ↳ Laxatives
 - ↳ Bulking agents
 - ↳ Stimulant laxatives
 - ↳ Stool softeners
 - ↳ Osmotic laxatives
 - ↳ Colonic evacuation
 - ↳ Peripheral opioid-receptor antagonists
 - ↳ Other laxatives
 - ↳ Anal and rectal disorders (locally acting preparations)
 - ↳ Soothing haemorrhoidal preparations
 - ↳ Corticosteroid haemorrhoidal preparations
 - ↳ Haemorrhoidal sclerosants
 - ↳ Anal fissures

SUPPLEMENTARY FILE 6: Text (SMS) recruitment

Text wording was as follows (one practice used personalised text messaging as this was their usual method of text communication):

“Are you available to donate your breath for cancer research? (..Name..) Surgery are asking our patients to help develop a new breath-testing device. A sample of breath will be collected at the practice during a 15-minute visit with a researcher. You will be helping to develop a new tool that could potentially be used for early cancer diagnosis in the future. Are you interested in hearing more? Text YES for a callback”

5. Would a breath test to detect/rule out cancer be a useful future tool?
(please circle)

a. Point of care device with instant results

0 1 2 3

b. Breath test done in same way as a blood test, with results electronically

0 1 2 3

0= Not useful 1= Not sure 2= Useful 3= Very useful

6. If you think breath testing could be useful, which patient groups do you think it would particularly benefit?

.....

7. What do you think is a reasonable cost for GP surgeries to pay for one patient to have a breath test?

.....

8. How many patients complaining of general gastrointestinal symptoms do you see on average per day as a GP?

.....

9. Did your request to recruit patients for the study have any impact on the quality of your consultation? (please circle)

0= negative impact

1= no impact

2= minimal impact

3= Positive impact

0 1 2 3

Please elaborate on any reasons for your answer.....

10. Did patients raise questions or concerns when you mentioned the study? If so what were their concerns?

.....

11. How could we have improved the organisation of the study?

.....

Please do not hesitate to contact me with further comments or questions:
g.woodfield@imperial.ac.uk

GP Questionnaire Version 5.0 16/1/17

SUPPLEMENTARY FILE 8: Patient acceptability questionnaireNon-invasive testing for the diagnosis and assessment of gastro-intestinal disease-
Patient Acceptability Questionnaire

General Practice name..... Study ID..... Date.....

Please tick the box corresponding to your level of agreement with the following statements:

1.	Yes, today	Yes, in the past 2 months	Yes, over 2 months ago	No, not within the past 5 years	Never
Have you seen a doctor because of stomach/ bowel/abdominal symptoms in the past 5 years?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

2.	More than 6 months	Between 2-6 months	Less than 2 months	Less than 1 week	Not applicable
How long did you have your most troubling symptom before seeing a doctor?	<input type="checkbox"/>				

3.	Extremely	Quite a bit	Moderately	Slightly	Not at all	Not applicable
How worried were you about your abdominal symptoms when you had them?	<input type="checkbox"/>					

4.	Very satisfied	Satisfied	Dissatisfied	Very Dissatisfied	Not applicable
How satisfied were you with the explanation given for how to do the breath test?	<input type="checkbox"/>				

5.	Very easy	Easy	Difficult	Very Difficult	Not applicable
How easy was it to do the breath test?	<input type="checkbox"/>				

If you found it difficult/very difficult, please explain why.....

.....

6.	Very comfortable	Comfortable	Uncomfortable	Very uncomfortable	Not applicable
How comfortable were you whilst wearing the face mask?	<input type="checkbox"/>				

Patient Questionnaire Version 9.0 5/7/17

7.		Took too long	Acceptable amount of time	Too quick	Not applicable
What did you think about the time it took to give a breath sample?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

8.	Very comfortable	Comfortable	Uncomfortable	Very uncomfortable	Not applicable
How did you find the experience of holding the device during the test?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

9.	Very comfortable	Comfortable	Uncomfortable	Very uncomfortable	Not applicable
Would you be comfortable to do the breath test again, if recommended by a doctor?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

10.	Strongly encourage	Encourage	Discourage	Strongly discourage	Not applicable
Would you encourage family and friends who were offered a breath test to complete it?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

11. How could the breath test be improved?

.....

.....

.....

.....

Please do not hesitate to contact me with further comments or questions:

Dr Georgia Woodfield
 Clinical Research Fellow, Imperial College London
 g.woodfield@imperial.ac.uk

SUPPLEMENTARY FILE 9: Quality control (QC) process for lab instruments

Two types of QC were performed daily for the proton transfer reaction time-of-flight mass spectrometer (PTR-ToF-MS) (1). A first *instrument QC* evaluated instrument stability with the three ionisation modes (H_3O^+ , NO^+ and O_2^+) against four parameters: impurities, fragmentation, mass resolution and accuracy. These were measured by using a permeation unit (a unit generating a constant flow of gaseous standard VOCs with known concentrations) which was connected directly to the PTR-ToF-MS for five minutes. Impurities levels below 10% were considered acceptable. Accuracy was evaluated through quantification of a benzene certified standard permeation tube (Kin-Tek Analytical Inc., La Marque TX). The PTR-ToF-MS quantitative measurement had to be within 20% of the certified standard in order to pass the QC. Fragmentation had to be above 60% to pass the QC. Butyric acid fragmentation was used as the check for H_3O^+ , where the ratio of diagnostic ions was used m/z 89 / (m/z 43 +71+89). For NO^+ butanal fragmentation was used, where the ratio of diagnostic ions was used m/z 71 / (m/z 43 +71+89). Resolution had to be above 1500 $m/\Delta m$. Our lab has a standard operating procedure for measuring instrument reproducibility with an action plan of what to do when parameters are not reaching the appropriate levels. No one can use the instrument for analysis until it passes the QC.

The second *standard QC* check evaluated the recovery of VOCs from TD tubes loaded from the permeation unit (2). Tube loading was performed at a flow of 0.910 (+-0.010) L/min at a temperature of 30 degrees Celsius. This was done by connecting a pocket pump to the permeation tube inlet via a TD tube. Flow across the tube was achieved by exploiting the permeation unit flow with the addition of the pocket pump. VOCs from the permeation unit were passed through the tube for 2.5 minutes. The VOC recovery was then measured by analysing the TD tubes using the PTR-ToF-MS. This test is performed daily before laboratory users can use the instrument.

For the gas chromatography mass spectrometer (GC-MS), five TD tubes loaded with a standard mixture (as explained above) were analysed daily. Retention time, peak shape and peak area were used to assess consistency and accuracy of the instrument response.

SUPPLEMENTARY FILE 10: Quality control process for TD tubes

Our lab follows a simple threshold system to identify whether there is breath present on a TD tube in sufficient quantities for it to be allowed to proceed to analysis. The full paper presenting this method is pending publication.

The QC system works by checking the VOC data from each TD tube to see that it reaches the minimum level for concentration of a particular reference compound (compound differs depending on whether the TD tube was analysed by PTR-ToF-MS H_3O^+ , NO^+ or O_2^+ ionisation or GC-MS). VOC data from TD tubes with inadequate levels of the reference compound, and therefore inadequate levels of breath within them, are discarded before data analysis.

This QC system is required because when collecting a breath sample in TD tubes it may not be immediately obvious that the full 500ml of breath has passed through the TD tube, as even where the ReCIVA software indicates the correct volume collected, breath can be lost if the caps are not tightened adequately on the tube post collection, or if the TD tube ends were not tightly sealed during breath transfer.

Our lab identified thresholds for particular reference compounds:

- **Acetone >45 ppb for PTR-ToF-MS (H_3O^+ ionisation)**
- **Isoprene >2.5 ppb for PTR-ToF-MS (NO^+ ionisation)**
- **Isoprene >5 ppb for PTR-ToF-MS (O_2^+ ionisation)**
- **Acetone >7,500,000 area counts for GC-MS**

These thresholds were identified by comparing breath samples to a control group of TD tubes with non-biological samples that consisted of (i) empty conditioned TD tubes; (ii) 500 ml of room air samples collected onto TD tubes using ReCIVA, following a procedure similar to that adopted for patient breath and (iii) TD tubes, previously conditioned and then loaded with a standard mixture of benzene (63 ppb, certified standard, Kin-Tek Analytical Inc., La Marque TX), phenol (90 ppb), butyric acid (20 ppb), pentanoic acid (5 ppb), hexanoic acid (5 ppb), decanal (4 ppb) and butanal (5 ppb), generated by a permeation unit (ES 4050P, Eco Scientific, Gloucestershire UK).

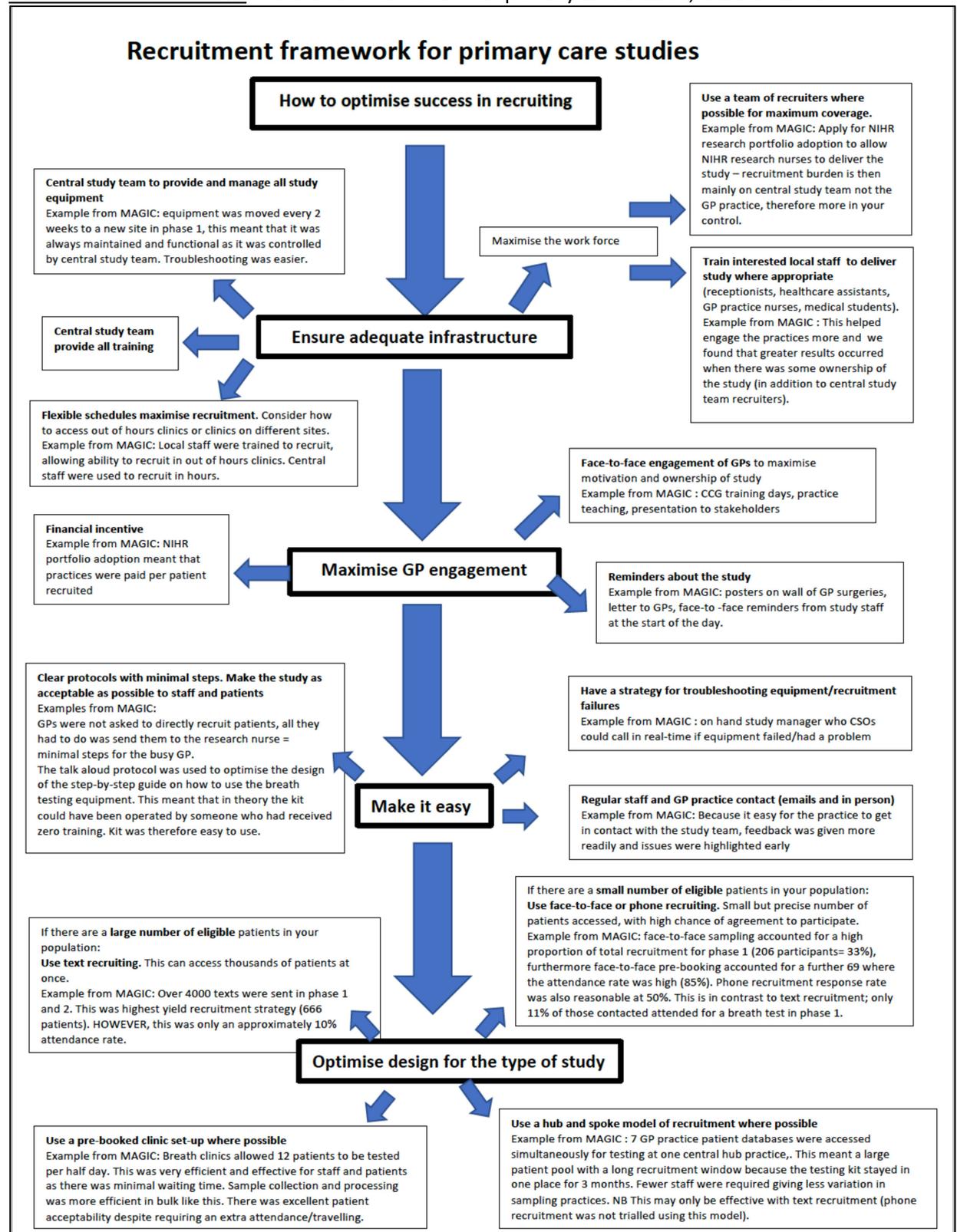
Over 100 breath samples were compared to over 100 controls for each ionisation of the PTR-ToF-MS and for the GC-MS, 1097 samples in total.

Based on this work, TD tubes which contain high levels of the appropriate reference compound are assumed to have high enough concentrations of breath collected onto the TD tube. The sample is then deemed to be adequate for inclusion in our group's study data.

SUPPLEMENTARY FILE 11: Summary of barriers to recruitment and subsequent implementations, during Phase 1 of Breath MAGIC study (from Field notes, teleconference and focus group of CSOs)

Recruitment Barrier	Cause of this problem (CSO view point)	Implementation made	Effect of implementation	Resolved?
1. Slow referral of patients, leading to slow recruitment and inefficient use of time for CSOs (6 of the first 10 days of sampling had 0 or 1 referrals per day only)	-Busy GP practices: GPs didn't have time - GPs varied in their interest levels in research -Locums were less likely to refer -Some had not heard about the study <i>"Nurses and HCAs would refer patients but GPs had to be prompted every session."</i> <i>"6 patients were found in the system to have GI symptoms, but none were referred. When asked directly, GPs and nurses said they hadn't remembered."</i>	-Emphasis on daily interaction of CSOs with GPs, reminding them of study and inclusion criteria - New GP poster and information leaflet was made	-Improved experience of CSOs meant they engaged more with GPs - Poster and leaflet helped guide discussion and acted as a reminder for GPs	Partially... - Still a marked variation between engagement levels of different individuals, often heightened by very busy practices
2. Persistence of above problem, Face-to-face sampling therefore inconsistent between practices	- As above, problem not fully solved	-Addition of phone and text bookings as well as face-to-face pre-booking. -Inclusion of referrals from other healthcare staff e.g specialist/practice nurses and HCAs doing clinics alongside GPs	-Tailored to practice resources; played to strengths of each practice -Engaged receptionists and HCAs in the recruitment task- drastically increased numbers - Efficient use of CSO time with dedicated breath clinics -However relied on staff time to pre-identify and call/text patients	Recruitment increased exponentially Target reached 6 weeks ahead of schedule
3.Smaller practices had lower recruitment	Some small practices had fewer GPs and fewer sessions. This meant that on days that there were baby clinics/other specialist clinics, no sampling could occur	-Combine practices that are close together, to boost recruitment and make CSOs time more efficient.	We combined Carepoint into Acre surgery in Northwood, so that patients could be referred from either site to see the same nurse. This helped recruitment.	Good solution for these practices, but not possible in every location.
4.Labour intensive sampling: requires a dedicated CSO present all day for sampling. Costly from central research team staff perspective	-The study design relied on central research team doing all the recruitment and breath testing.	-We trained practice staff to deliver the breath test instead of central study team. HCAs, nurse practitioners, local research nurses and a medical student (after GCP training) were trained in 4 practices, which covered the sampling over 10 study weeks	This was a huge success. It saved a lot of central study team time/resources. Local teams found it easy to recruit as they knew their colleagues and often knew the patients, familiar with the computer systems and had their own clinic rooms. They also recruited during evening clinics as this was their expected working hours at the GP practice.	Yes – This also showed how a breath test could be used in future by multiple different staff members. There was also the added bonus of great collaboration with sometimes new GP practices with great engagement from many different staff members.
5.Study timings of 9-5pm not always matching GP clinic hours	Some GPs have evening clinics and have admin time in the afternoons-not good for patient sampling	-CSO staff hours were adjusted where possible -Addition of face-to-face pre-booking enrolment so	This strategy allowed sampling outside our planned sampling hours.	Yes- recruitment was excellent during the weeks where we used local staff, and it had other positive effects

		that patients could return within hours. -Local practice staff were trained (see above) which meant they could sample in late clinics if this was their usual working pattern	Other out of hours patients could also still be recruited at a later date using face-to-face pre-booking	such as cost saving and increased engagement with local practices and different allied health professionals.
6. Equipment shortages and delays	-Mask supply ran out at one point in the study -GC Mass spectrometer malfunctioned at one point, leading to backlog of TD tubes and therefore a supply issue	-For problematic weeks, sampling was limited to 20 patients per week. This affected 8 of the sampling weeks (6 practices affected). - Samples were processed on the PTR-ToF-MS where possible, allowing clearance of back log	-20 per week cap was not a problem for most GP practices as long as they knew in advance -Only one GP practice was negatively affected where 3 days of sampling had to be cancelled -3 weeks of recruitment was significantly affected by this issue because phone bookings were not made in practices where this would have been the recruitment method.	Resolved at the time with back-up alternative instrument -Instrument malfunctions are unavoidable but this could also be mitigated by a larger back-up supply of tubes - Masks will now be ordered a year in advance for next time.
7. Problems with rooms/space	Room availability- <i>"There was a pressure on rooms today, and I had to move rooms mid-course during recruitment and find a suitable computer to work from. The recruitment was fairly low as a result".</i> Upstairs rooms – <i>"being based upstairs as opposed to downstairs where most patients were being seen impacted on recruitment"</i>	Requesting downstairs rooms where possible/rooms near where the patients were being seen.	Not clear what the effect of the intervention was, as room availability could often not be controlled, but anecdotally the nurses felt that more patients came when the sampling room was easy/accessible.	Partially resolved by enquiring and requesting certain rooms, but room availability was often largely not within our control.

SUPPLEMENTARY FILE 12: Recruitment framework for primary care studies; lessons learnt

SUPPLEMENTARY FILE 13: GP opinions about place of the breath test in future care.

Twenty-one GPs, from 10 of the 26 participating practices, answered the GP specific questionnaire. GPs felt the breath test would be best placed as a point of care test with instant results, considering this *“very useful”* (88%) or *“useful”* (8%). 10 GPs (48%) felt testing would still be useful if results were available electronically at a later date. There were varied views about which groups of patients would benefit most from a breath test, ranging from *“all age, any group”* to *“low risk cancer patients without red flag symptoms”*, to *“at risk groups”*. Of the eleven GPs who said *“at risk groups”*, six of them qualified this with specific symptoms (*“dyspepsia/weight loss”, “chronic reflux”, “chronic dyspepsia”, “elderly and frail”, “elderly with weight loss”, “lower abdominal or upper GI symptoms”*) and three others gave age cut-offs (*“>30”, “>45”* and *“45-74”*). Only two GPs gave an opinion about cost per test (*“£10”* and *“£14”*) where the others indicated *“don’t know”* or said that the CCG should decide. From the GPs’ perspective it appeared the breath test was feasible, but that its cost and place in a future referral pathway was yet to be determined.

SUPPLEMENTARY FILE 14: Summary of themes regarding feasibility and acceptability of the sampling process (from Field notes, teleconference and focus group of CSOs)

Theme	Examples of representative comments
Patient based limiting factors	<p><i>"Small print of the information sheet was an issue for one patient"</i></p> <p><i>"One patient had a bad cold and felt like they couldn't exhale properly into the mask"</i></p> <p><i>"One patient was anxious as said she was really claustrophobic, but she managed to do it in the end without a problem"</i></p> <p><i>"There were really no issues with patients, all were happy to help"</i></p>
Equipment (computer) based limiting factors	<p><i>"Computer very haphazard. Flowometer not working. About half of the patients were timed samples in the end. Cutting out (going blue) which responded to ctrl alt delete."</i></p> <p><i>"Flowometer responding to turning off and on but not always. I think it is a mask connection problem,"</i></p> <p><i>"Large number of timed samples because of flowometer issues"</i></p> <p><i>"Screen went blue"</i></p> <p><i>"Screen froze and wouldn't respond"</i></p> <p><i>"The machine displayed an error message and the breath monitor did not increase. Solved when system was rebooted but patient had gone by then."</i></p> <p><i>"Computer fault again, not recognising mask and not reading, had to do timed sample. This corrected itself the next day."</i></p>
Equipment (ReCIVA device) based limiting factors	<p>CSOs felt that the equipment was "fiddly" but "ok once they got the hang of it", particularly the spanners used to tighten screws.</p> <p>They commented on the time it took to set up the equipment at the start of the day and said they found it easier if there was a desk/workspace to lay out equipment and documents.</p> <p>When commenting on their perceptions of patient acceptability, CSOs said that they observed that patients didn't always breathe "normally" when wearing the masks, and some held their breath. This was not reported by patients in the acceptability questionnaires.</p>
Equipment (TD tube) based limiting factors	<p>1 tube in phase 1 arrived with black soot (sorbent) coming out of one end. This was sent to the company for repacking. Breath data was discarded.</p>
Training	<p>CSOs commented in the focus group that training was vital to performing the breath test because "preparation is key", "if you prepare before then it runs like clockwork and you can sample patients back to back".</p> <p>The troubleshooting manual was "useful" but the training allowed "hands on practice".</p> <p><i>"The fiddliness and multiple steps required made it not very obvious what to do next unless you had had the training."</i></p>
Human errors in sampling	<p>Early in phase 1 of the study 13 tubes in a batch arrived with no caps on, meaning that the samples would have been very contaminated. This was solved by contacting the CSO individually, (who had mistakenly forgotten to do this). This was an easily solved problem that did not recur. Breath data was discarded.</p> <p>Tubes occasionally arrived with loose caps on. However, these were likely tight enough to have held in the sample, but were easily removed by hand. This was solved by sending a reminder to CSOs and including this as a point in all subsequent training. Breath data was not necessarily discarded, but quality was checked as per all samples.</p> <p>Some tubes were overly tightened with the spanners, which could potentially damage the tubes. This was solved by supplying handheld spanners only, rather than conventional long spanners.</p>

References

1. Romano A, Hanna GB. Identification and quantification of VOCs by proton transfer reaction time of flight mass spectrometry: An experimental workflow for the optimization of specificity, sensitivity, and accuracy. . *J Mass Spectrom*. 2018;53(4):287-95.
2. Romano A, Doran S, Belluomo I, Hanna GB. High-Throughput Breath Volatile Organic Compound Analysis Using Thermal Desorption Proton Transfer Reaction Time-of-Flight Mass Spectrometry. *Analytical chemistry*. 2018;90(17):10204-10.