ESMO Open
Cancer Horizons

Somatic cancer genetics in the UK: real-world data from phase I of the Cancer Research UK Stratified Medicine Programme

Colin R Lindsay,[1,2,3] Emily C Shaw,[4,5] Fiona Blackhall,[1,2,3] Kevin G Blyth,[6,7,8] James D Brenton,[9,10,11] Anshuman Chaturvedi,[12,13] Noel Clarke,[13] Craig Dick,[6,14] Thomas R J Evans,[6,15,16] Geoff Hall,[17,18] Andrew M Hanby,[17,19,20,21] David J Harrison,[22,23] Stephen R D Johnston,[24,25] Malcolm D Mason,[26,27,28] Dion Morton,[29,30] Julia Newton-Bishop,[17,31] Andrew G Nicholson,[32,33] Karin A Oien,[6,16] Sanjay Popat,[25,33,34] Doris Rassl,[10,35] Rowena Sharpe,[4] Phillipe Taniere,[30,36] Ian Walker,[4] William A Wallace,[23,37] Nicholas P West,[17,21] Rachel Butler,[38] David Gonzalez de Castro,[39] Mike Griffiths,[40] Peter W M Johnson[4,5]

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/esmoopen-2018-000408).

To cite: Lindsay CR, Shaw EC, Blackhall F, *et al*. Somatic cancer genetics in the UK: real-world data from phase I of the Cancer Research UK Stratified Medicine Programme. *ESMO Open* 2018;**3**:e000408. doi:10.1136/esmoopen-2018-000408

For Presented at statement see end of article.

Received 23 May 2018
Revised 30 May 2018
Accepted 31 May 2018

For numbered affiliations see end of article.

Correspondence to
Dr Colin R Lindsay; colin.lindsay@manchester.ac.uk

ABSTRACT

Introduction Phase I of the Cancer Research UK Stratified Medicine Programme (SMP1) was designed to roll out molecular pathology testing nationwide at the point of cancer diagnosis, as well as facilitate an infrastructure where surplus cancer tissue could be used for research. It offered a non-trial setting to examine common UK cancer genetics in a real-world context.

Methods A total of 26 sites in England, Wales and Scotland, recruited samples from 7814 patients for genetic examination between 2011 and 2013. Tumour types involved were breast, colorectal, lung, prostate, ovarian cancer and malignant melanoma. Centralised molecular testing of surplus material from resections or biopsies of primary/metastatic tissue was performed, with samples examined for 3–5 genetic alterations deemed to be of key interest in site-specific cancers by the National Cancer Research Institute Clinical Study groups.

Results 10 754 patients (98% of those approached) consented to participate, from which 7814 tumour samples were genetically analysed. In total, 53% had at least one genetic aberration detected. From 1885 patients with lung cancer, *KRAS* mutation was noted to be highly prevalent in adenocarcinoma (37%). In breast cancer (1873 patients), there was a striking contrast in *TP53* mutation incidence between patients with ductal cancer (27.3%) and lobular cancer (3.4%). Vast inter-tumour heterogeneity of colorectal cancer (1550 patients) was observed, including myriad double and triple combinations of genetic aberrations. Significant losses of important clinical information included smoking status in lung cancer and loss of distinction between low-grade and high-grade serous ovarian cancers.

Conclusion Nationwide molecular pathology testing in a non-trial setting is feasible. The experience with SMP1 has been used to inform ongoing CRUK flagship programmes such as the CRUK National Lung MATRIX trial and TRACERx.

Key questions

What is already known about this subject?

- The core genetics of the six cancers explored in this article are well delineated in projects such as The Cancer Genome Atlas.

What does this study add?

- Rather than being confined to one centre/city, this study looked at very high sample numbers (eg, 1885 patients for lung cancer) across eight cities, offering a 'snapshot' of real-world somatic cancer genetics in the UK.
- By taking a reductionist rather than bioinformatic approach to genetic examination/analysis, we discovered a number of striking results to validate.
- For example, the *KRAS* mutation rate was unexpectedly high in UK lung adenocarcinoma at 37%.
- In breast cancer, rates of *TP53* mutation were nearly 10-fold higher in ductal versus lobular disease.
- The core genetics of rarer pathologies such as mucinous colorectal and tubular breast adenocarcinomas were also delineated.

How might this impact on clinical practice?

- This article and programme is intended as a repository for advancing preclinical, translational and clinical cancer research.
- Establishing the prevalence of core cancer genetic alterations provides a platform to drive forward precision medicine and molecularly targeted clinical trials.
- We also convey the challenges of setting up a nationwide infrastructure for molecular testing—an experience that we hope will be informative for other countries.

INTRODUCTION

The breadth of clinical and biological genomic challenges which have become apparent in cancer over the past five years highlights how technological advance is driving expectation ever higher. Knowledge of therapeutically targetable driver mutations in genes such as *BRAF*, *EGFR* and *BRCA1/2* is now being supplemented by the discovery of a new generation of resistance mutations: these include aberrations found in the same gene, for example, *EGFR* T790M, or in genes downstream of the same cellular pathway such as MEK-mediated resistance in *BRAF*-mutated melanoma.[1–4]

The Cancer Research UK Stratified Medicine Programme (SMP1) was established in response to growing demand for prospective analysis of prognostic and predictive genetic markers in clinical tumour samples.[5 6] At inception, there were two systemic clinical deficits in the UK National Health Service (NHS) inhibiting progress towards molecular diagnostics as a component of normal cancer care[7]:

1. A nationwide infrastructure facilitating key molecular pathology tests at the point of diagnosis was necessary. This would reduce waiting times involved with requesting a test in retrospect, expediting patient access to prognostic genetic information and associated genetically targeted therapies.
2. Little infrastructure existed for the systematic genomic analysis of tumour tissue surplus to diagnostic requirements, whether obtained through biopsy or resection, in the majority of patients not recruited to clinical trials.[8]

The central aim of SMP1 was therefore to facilitate an experience and infrastructure for molecular diagnostics in solid cancers across the UK, establishing its incorporation in the normal pathway of patient care that could be further developed during subsequent phases of the programme (SMP2). We also hypothesised that, in an era where laboratory genomic studies are offering increasing degrees of complexity, 'bridging' studies such as SMP1 would be necessary to ensure the ensuing data 'storm' runs in parallel with direct translational improvements in patient care.[9] Perhaps as much progress could be gained by projects which focus on key genetic players and offer the opportunity to work backwards from identified patient responders.[10–12]

Here we offer a summary of important findings for implementation and clinical practice as well as core genetic results from patients across the UK in six disease areas: carcinomas of the breast, lung, colorectum, prostate, ovary and malignant melanoma.

METHODS

The Cancer Research UK Stratified Medicine Programme

The Cancer Research UK Stratified Medicine Programme (Research Ethics Committee reference 11/EE/0202) commenced in 2011, at clinical and laboratory sites in England, Wales and Scotland. Phase I (SMP1) took place between September 2011 and July 2013. Six different solid tumour types (breast, colorectal, lung, ovarian, prostate cancer and malignant melanoma) were chosen for study. Twenty-six participating hospitals formed the network, coordinated through Cancer Research UK and National Institute for Health Research-funded Experimental Cancer Medicine Centres (grouped to form local 'Clinical Hubs') (figure 1). At these sites, patient consent was sought for centralised molecular testing (performed at 'Technology Hubs') of surplus material from resections or biopsies of primary/metastatic tumour tissue, performed as part of routine clinical care.

Figure 1 Participating sites for phase I of the Cancer Research Stratified Medicine Programme. Yellow markers represent clinical hubs, red markers represent clinical and technology hubs. ICR , Institute for Cancer Research; RBH, Royal Brompton Hospital; RMH , Royal Marsden Hospital.

The tumour types included in SMP1 were selected for their representation of a large percentage of the UK cancer demographic: breast, colorectal, lung and prostate cancer make up over half of all incident cancer cases in the UK each year; ovarian cancer represents the fifth most common cancer in females; advanced malignant melanoma represents a clinically unmet need which has recently been successful as a driver for a new generation of genetically targeted therapies. Also, 3–5 genes of interest were prioritised for molecular testing in each tumour type, selected by relevant National Cancer Research Institute Clinical Study Groups for their prevalence and/or potential 'actionability'. Samples were examined for the following gene alterations: *PTEN*, *PIK3CA*, *BRAF* and *TP53* for breast and ovarian cancer; *KRAS*, *EGFR*, *ALK*, *BRAF* and *DDR2* (squamous only) for lung cancer; *KRAS*, *NRAS*, *BRAF*, *PIK3CA* and *TP53* for colorectal cancer (CRC); *PTEN*, *BRAF* and *TMPRSS2-ERG* for prostate cancer; *BRAF*, *NRAS*, *PIK3CA* and *KIT* for melanoma. Further information on the programme and techniques

used for molecular analysis is detailed in online Supplementary methods and supplementary methods table 1.

Patient eligibility

Eligibility criteria were designed to be broad and inclusive in order to maximise the relevance of findings to the generalUK population. Patients had to be aged 18 years or more, able to give written informed consent, and with a suspected or confirmed diagnosis of one of the following types of invasive malignancy: breast cancer (carcinoma including ductal, lobular and other subtypes); CRC (adenocarcinoma of the colon or rectum); lung cancer (carcinoma of the lung including both small cell and non-small cell subtypes but excluding carcinoid tumours and pleural malignant mesothelioma); malignant melanoma (advanced stage III or IV disease with at least regional lymph node involvement, from cutaneous primaries as well as less common mucosal sites); ovarian cancer (adenocarcinoma) and prostate cancer (adenocarcinoma).

Analysis, interpretation and reporting

Patients consenting to participate in the programme were required to have a sample submitted of formalin-fixed paraffin-embedded tumour tissue from a resection or biopsy procedure with surplus tissue, beyond that needed for making a tissue diagnosis, taken either from the primary tumour or from a site of metastasis.

The SMP1 gene sets for each tumour type comprised well-characterised hotspots in oncogenes, structural chromosomal rearrangements, as well as screening of multiple exons in tumour suppressor genes such as *PTEN* and *TP53*. More information on this is available in online supplementary methods .

Database analysis

For statistics, association of genotype with clinical characteristics and patient demographics were assessed using Fisher's exact test for dichotomous factors, and the Mann-Whitney test for continuous data, which were not adjusted for multiple testing. All p values were two-sided and considered statistically significant at <0.05. More detail on database analysis, including data curation and compilation, is available in online supplementary methods.

RESULTS

Patient population

Between August 2011 and July 2013, 10754 patients (98% of those approached) consented to participate in SMP1, with 9010 patient tumour samples examined. Of 7814 samples with data available at the time of our analyses, 53% had at least one aberration detected (table 1). Also, 44% of the samples were *wild type* for the genes and regions analysed and the remaining 3% of samples failed all gene tests. Data completeness varied between the clinical sites and between data items (table 2).

Lung cancer

Lung cancer samples were obtained from a total of 1885 patients. The baseline demographics of these patients are shown in online supplementary table 1. In total, 774 patients (41.1%) were diagnosed with adenocarcinoma, 399 patients (21.2%) with squamous cell cancer and 50 patients (2.6%) with small cell lung cancer. Also, 64 patients (3.4%) were diagnosed with non-small cell lung cancer (NSCLC) of non-specific histological subtype (online supplementary figure 1). Due to the focus on resection samples where tissue was plentiful, the majority of samples collected represented stage I–II lung cancer (1006 samples, 53.4%), followed in frequency by stage III lung cancer (21%) then stage IV (19.4%) lung cancer. A substantial percentage of baseline information was returned as 'not stated' (ie, either not tested, not documented or both), the most important of which was a deficit of smoking history.

Figure 2A offers an overview of genetic results obtained from the lung adenocarcinoma population. Of those tested, 92/774 (11.9%) were *EGFR* mutant, 287/774 (37.1%) were *KRAS* mutant, 19/774 (2.5%) were *ALK* rearranged and 18/774 were *BRAF* mutant (2.3%) (online supplementary table 2). The most common failed genetic test was *KRAS*, with no result returned in 127/774 (16.4%) samples: incidence of gene mutation/modification would have been considerably higher, particularly for *KRAS* (287/647 samples, 44.36%), had these failed samples been excluded from the total numbers in our final analysis (online supplementary table 3). A further breakdown of *EGFR* mutation results showed that 67/92 (72.8%) *EGFR*-mutant cancers harboured a solitary sensitising mutation, 10/92 (10.9%) a solitary resistance mutation, 3/92 (3.3%) had both sensitising and resistance mutations and 12/92 (13%) had mutations of unknown significance (online supplementary table 4).

Table 1 Summary results of molecular analysis performed during SMP1 by tumour type

Tumour type	Breast cancer	Colorectal cancer	Lung cancer	Malignant melanoma	Ovarian cancer	Prostate cancer
Number of samples	1873	1605	1885	535	557	1359
Failed all tests (%)	5.3	0.9	2.8	3.6	3.2	4.0
Wild type for all genes (%)	45	19	64	31	40	52
Aberration in more than one gene (%)	7	33	0.5	2	4	2

Table 2 Overall SMP1 data completeness by patient disease cohort

Data item	Patient cohort						
	Breast cancer	Colorectal cancer	Lung cancer	Malignant melanoma	Ovarian cancer	Prostate cancer	Overall
Total number of patients	1873	1605	1885	535	557	1359	7814
Gender (%)	100	99	98	96	N/A	N/A	98
Year of birth* (%)	100	100	100	100	100	100	100
Year of diagnosis (%)	79	75	52	74	67	69	69
Ethnic category (%)	71	73	75	81	70	60	72
Histological subtype (SNOMED morphology) (%)	100	99	77	92	97	92	93
Histological grade† (%)	83	88	N/A	N/A	62	53	72
Pathological T classification‡ (%)	92	69	91	33	35	50	62
Pathological N classification‡ (%)	86	81	89	31	24	35	58
Pathological M classification‡ (%)	24	74	77	54	79	33	57
Integrated TNM stage‡ (%)	92	89	94	71	84	55	81

For each data item, the percentage completeness given is the percentage of patient records containing valid and informative data according to the stipulated attributes in the clinical dataset.
*Date of birth and date of diagnosis were recorded at patient level but truncated to 'year of' as an information governance measure to maintain confidentiality.
†Not mandatory where this is not a core Royal College of Pathologists (RCPath) dataset reporting item. For prostate cancer, the percentage refers to overall completeness of Gleason score components requested in separate data items.
‡Alternative staging systems used as follows with completeness given in integrated stage field: FIGO for ovarian cancer, AJCC version of TNM7 for melanoma. TNM7 has been used in all cases apart from colorectal cancer where TNM5 is currently used in the UK according to RCPath guidance.
AJCC, American Joint Committee on Cancer; FIGO, International Federation of Gynaecology and Obstetrics; N/A, not available; SNOMED, Standard Nomenclature of Medicine; TNM5/7, l'Union Internationale Contre le Cancer (UICC) Tumour/Node/Metastasis Classification of Malignant Tumours 5th/7th edition.

In squamous cell carcinoma, *DDR2* was mutated in 12/175 (6.9%), an incidence that becomes substantially higher if the large numbers of failed tests are excluded from analysis (12/119 samples, 10.1%). *EGFR* mutations were present in 3/399 squamous samples (0.8%), while *ALK* rearrangement was present in 2/399 (0.5%). *KRAS* (12/399 samples, 3%) and *BRAF* (5/399 samples, 1.3%) mutations were present in small numbers (online supplementary tables 2 and 3).

Forty-eight small cell lung cancers were analysed, with none of the samples harbouring a genetic mutation/modification (online supplementary tables 2 and 3).

Breast cancer

In total, 1873 patients with breast cancer were analysed within SMP1 (online supplementary table 5). Also, 1423 patients (76%) were diagnosed with invasive ductal carcinoma (IDC), 179 patients (9.6%) with invasive lobular carcinoma (ILC), 25 patients (1.3%) with mucinous adenocarcinoma and 35 patients (1.9%) with tubular carcinoma. Sixty-seven patients (3.6%) were diagnosed with mixed IDC/ILC (online supplementary figure 2 and table 6). Due to the focus on resection specimens, a majority of samples collected in SMP1 were taken from stage I–II breast cancer (1455 samples, 77.7%). Of patients where ER, PR and HER2 status was confirmed, ER was positive in 751/866 (86.7%) cancers, PR in 271/407 (66.6%) cancers and HER2 in 142/807 (17.6%) cancers. Only 3.2% of patients were confirmed as 'triple-negative' (60/1873 with ER- PR- HER2-), perhaps reflecting the relatively low percentage of stage III–IV patients recruited. Again, a substantial percentage of baseline information was returned as 'not stated', ranging from 35/1873 (1.8%) of patients for histological subtype to 1466/1873 (78.2%) of patients for PR status. ER and HER2 were not stated in 1007/1873 (53.8%) and 1066/1873 (56.9%) of patients, respectively.

On review of histopathological category, 420/1423 (29.5%) of IDC samples were *PIK3CA* mutant, 65/1423 (4.6%) were *PTEN* mutant, 389/1423 (27.3%) *TP53* mutant and none (0/1055 samples) were *BRAF* mutant (figure 2B and online supplementary table 6). The most common failed genetic test in IDC was *PTEN*, with no result returned in 354/1423 (24.9%) of samples tested: incidence of *PTEN* mutation would have increased to 6.1% (65/1069 samples) had these failed samples been excluded in our final analysis (online supplementary table 7). Of ILC samples, *PIK3CA* was mutated in 60/179 (33.5%), *PTEN* in 10/179 (5.6%), while *TP53* mutation was present in 6/179 samples (3.4%). Also, 25 mucinous adenocarcinomas had *PIK3CA* mutation in 3/25 samples (12%), *PTEN* mutation in 0/25 samples and *TP53* mutation in 4/25 samples (16%), while 35 tubular adenocarcinomas displayed *PIK3CA* mutation in 17/35 samples

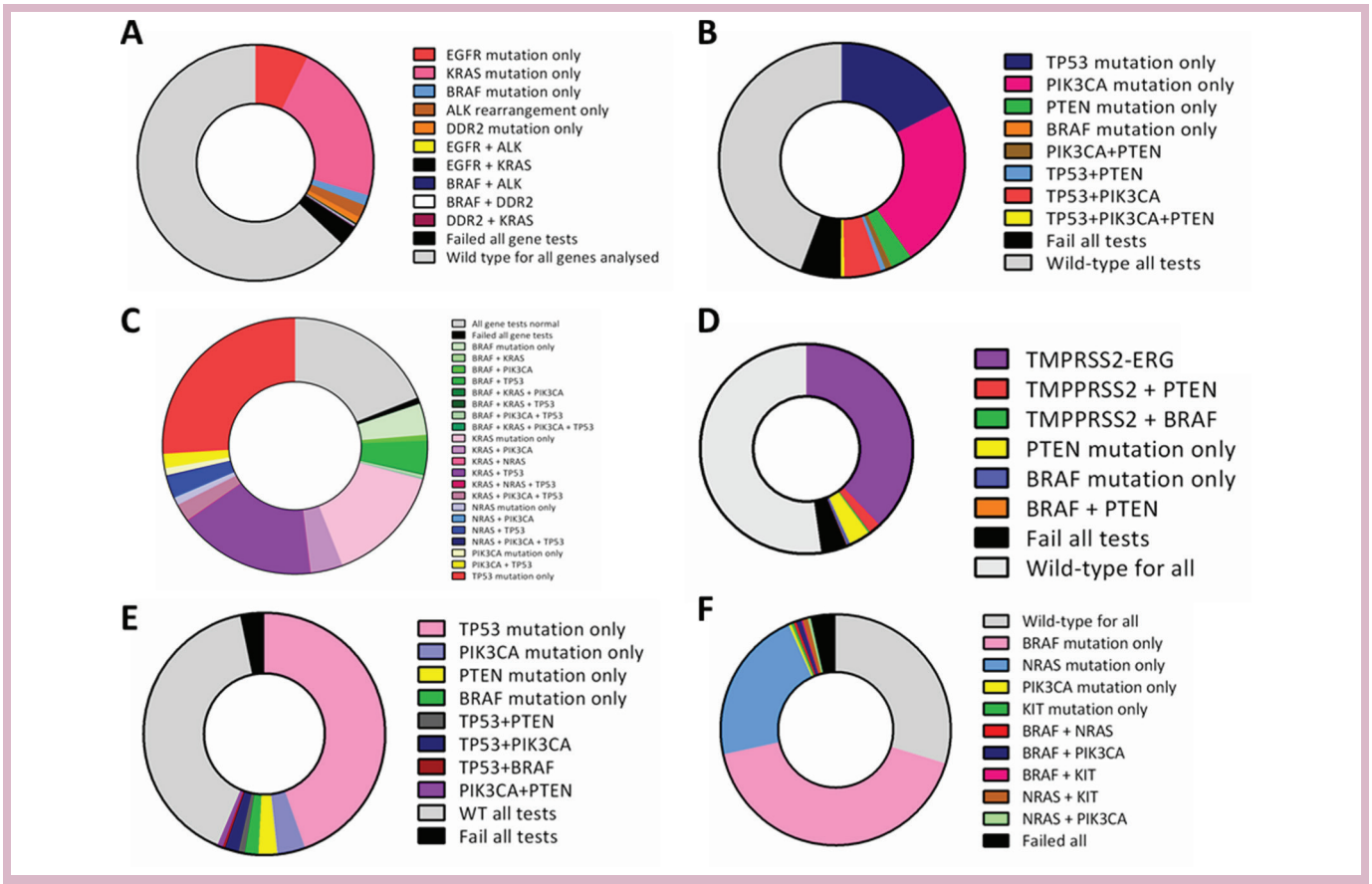

Figure 2 Overview of cancer genetics in the SMP1 cohort: lung cancer (A), breast cancer (B), colorectal adenocarcinoma (except mucinous subtype) (C), prostate cancer (D), ovarian cancer (E) and melanoma (F).

(48.6%), *PTEN* mutation in 1/35 samples (2.9%) and *TP53* mutation in 0/35 samples (online supplementary table 6 and 7).

Colorectal cancer

CRC samples were obtained from 1605 patients (online supplementary table 8). In total, 1508 patients (94%) were diagnosed with adenocarcinoma, with mucinous adenocarcinoma representing the most common histological variant (52/1605 samples, 3.2%). Also, 579 samples (36.1%) represented TNM stage I–II CRC, 602 samples (37.5%) stage III and 241 samples (15%) stage IV. The majority of tumours were graded as moderately differentiated (1126/1605, 70.2%). Again, a substantial percentage of baseline information was returned as 'not stated', ranging from 17/1605 (1.1%) of patients for histological subtype to 506/1605 (31.5%) of patients for lymphovascular invasion ('LVI', used as surrogate marker for extramural vascular invasion in core dataset).

The vast inter-tumour heterogeneity of CRC is clearly represented despite our analysis of only five genes, including myriad double and triple combinations of genetic aberrations (figure 2C). Of adenocarcinomas tested, 581/1508 (38.5%) were *KRAS* mutant, 824/1508 (54.6%) were *TP53* mutant, 144/1508 (9.5%) *BRAF* mutant, 61/1508 *NRAS* mutant (4%) and 158/1508 were *PIK3CA* mutant (10.5%) (online supplementary table 9).

The most common failed genetic test was *TP53*, with no result returned in 291/1508 (19.3%) of samples tested, the incidence of this mutation would have increased to 67.7% (824/1217 samples) had these failed samples been excluded from our final analysis (online supplementary table 10).

In CRC mucinous adenocarcinoma, *BRAF* was mutated in at least 20/52 (38.5%) of samples (online supplementary tables 9 and 10). Codon 600 *BRAF V600E* mutations accounted for 19 of the 20 samples from *BRAF*-mutated mucinous CRC. Incidence of *BRAF* mutation in mucinous CRC was significantly higher relative to its incidence in the SMP1 CRC adenocarcinoma population as a whole ($p<0.0001$). Of other genes tested, *TP53* mutations were significantly less frequent (3/52 samples, 5.8%) relative to adenocarcinoma overall ($p<0.0001$), and *PIK3CA* mutations were more common (11/52 samples, 21.2%; $p=0.022$). *NRAS* and *KRAS* mutations were not significantly different to that observed overall.

Prostate cancer

Samples from 1359 patients with prostate cancer were analysed (online supplementary table 11). Adenocarcinoma histology was reported in 91.8% of cases diagnosed. In total, 430 cases (31.6%) were diagnosed at stages I–II, 238 (17.5%) were stage III samples and 78 (5.7%)

were stage IV samples. 613 (45.2%) of samples were of unknown stage.

Figure 2D offers an overview of genetic results obtained from the prostate cancer population. In adenocarcinoma samples, 11/937 (1.2%) were *BRAF* mutant, 67/1247 (5.4%) *PTEN* mutant and 501/1247 (40.2%) were *TMPRSS2*-rearranged. The most common failed genetic test in prostate cancer was *PTEN*, with no result returned in 303/1247 (24.3%) samples tested: incidence of gene mutation would have been higher, including for *TMPRSS2-ERG* (501/1117 samples, 44.9%), had these failed samples been excluded from our final analysis (online supplementary table 12).

Ovarian cancer

Ovarian cancer samples were obtained from 557 women (online supplementary table 13). In total, 360 patients (64.3%) were diagnosed with serous carcinoma, 33 patients (5.9%) with clear cell, 36 patients (6.5%) with endometrioid, 11 patients (2%) with mucinous and 72 patients (12.9%) with unspecified epithelial ovarian cancer (EOC) (online supplementary table 12 and supplementary figure 3). The most common stage at presentation was stage III ovarian cancer (268 samples, 48.1%), followed by stages I–II (117 samples, 21%), then stage IV (82 samples, 14.7%). A significant percentage of baseline information was returned as 'not stated', most importantly a deficit of information on low-grade versus high-grade disease.

Figure 2E gives an overview of ovarian cancer genetic results. Overall, 35/557 (6.3%) samples were *PIK3CA* mutant, 23/557 (4.1%) *PTEN* mutant, 265/557 (47.6%) *TP53* mutant and 12/516 (2.3%) were *BRAF* mutant. In serous cancer, 6/360 (1.7%) were *PIK3CA* mutant, 2/360 (0.6%) *PTEN* mutant, 181/360 (50.3%) *TP53* mutant and 7/327 (2.1%) were *BRAF* mutant (online supplementary table 14). The most common failed genetic test in serous EOC was *PTEN*, with no result returned in 114/360 (31.7%) of samples tested: percentage mutation would have been 65.3% (181/277 samples) had these failed samples been excluded from our final analysis (online supplementary table 15).

For 33 clear cell EOCs, *PIK3CA* was mutated in nine samples (27.3%), *PTEN* mutation in four samples (12.1%), while *TP53* mutation was present in seven samples (21.2%). In 36 endometrioid EOCs, *PIK3CA* was mutated in 9/36 samples (25%), *PTEN* mutation in 7/36 samples (19.4%) and *TP53* mutation in 12/36 samples (33.3%). Of 11 patients with mucinous EOC, *TP53* mutation was present in 5/11 mucinous samples (45.5%), with *PI3KCA* and *PTEN* mutations absent. *BRAF* mutation was absent in all subtypes except for serous EOC (online supplementary tables 14 and 15).

Metastatic melanoma

In total, 535 patients with metastatic melanoma were analysed (online supplementary table 16). A significant percentage of baseline information was returned as 'not stated', ranging from 23/535 (4.3%) of patients for gender to 438/535 (81.9%) for LVI. 232/535 patients (43.4%) were *BRAF* mutant, 124/535 (23.2%) *NRAS* mutant, 8/535 (1.5%) *PIK3CA* mutant and 7/535 (1.3%) *KIT* mutant (Figure 2F). For *BRAF* mutation, 219/232 samples were documented as 'V600' or 'V600E' (94.4%), with another 8/232 'V600K' (3.4%). Removing gene test failures from the total number of samples, mutation prevalence in *BRAF*, *NRAS*, *PIK3CA* and *KIT* increased to 45.7%, 29.6%, 2% and 2.1%, respectively (online supplementary table 17).

DISCUSSION

Here we have reported results from the first UK-wide study assessing molecular pathways of cancer within the UK NHS. We focused on six common cancers (lung cancer, breast cancer, CRC, prostate cancer, ovarian cancer and metastatic melanoma), finishing with a 98% consent rate for patient participation. To the best of our knowledge, this study has offered a number of novel results relevant to our future understanding of UK cancer genetics that may also have international relevance. It also highlights a number of challenges which will be important for streamlining national molecular programmes in the future.

One key result from SMP1 was a 98% rate of patient consent for participation across all six cancer types. This success was achieved by the stipulation of a blood sample and 'surplus' tissue only for eligibility, that is, tissue derived from resection or biopsy, and remaining after all necessary diagnostic tests had been performed. No additional invasive procedures were necessary, and results from clinical trials that mandate further 'research protocol' biopsies suggest that this percentage would have been lower had patients been asked to undergo this.[13] This 98% acceptance rate also suggests that concerns about genetic and clinical data privacy are not as prevalent as might be expected, despite changing data protection regulations that have caused anxiety in the research community.[14]

Table 3 offers a perspective of the advantages and challenges involved with recruitment, data collection and analysis from SMP1. We believe these data offer a unique insight, unrepresented in other genomic studies: there was no planned selection bias, large patient numbers were involved and prospective assessment of important functional mutations was implemented, demonstrating the feasibility of this programme to allow nationwide patient access to relevant novel therapies, clinical trials and other research opportunities. However, results were hypothesis-generating and should be interpreted within the context of a retrospective observational analysis requiring further clinical validation. Although selection bias may be minimised in a nationwide study such as this, the potential for unplanned bias still exists: for example, a concurrent clinical trial using SMP1 to select patients with a particular cancer genotype. Expected bias included a weighting towards specific histologies and early-stage disease, given the focus on submission of resected specimens to increase the likelihood of sufficient material

Table 3 Summary of advantages and challenges encountered in SMP1

	Advantages	Challenges
Recruitment	Broad patient eligibility, determined by histological diagnosis of one of the six cancers types, enabling inclusion of a range of patients from across the UK, all receiving care within the National Health Service (NHS) Approval granted by research ethics committee for clinical sites to use existing biobanking consent forms and information sheets, after review to confirm equivalence with CRUK SMP1 paperwork	Potential for unplanned selection biases impacting SMP1 patient recruitment due to concurrent clinical trials, for example, for patients with oestrogen receptor-positive breast cancer
Sample preparation and analysis	Insights generated into differences in tissue handling processes between different laboratories, facilitating the process of harmonising practice and understanding the impact on subsequent genetic analysis Collaborative working between technology hubs facilitating shared learning and evidence-based evolution of approach to genetic analysis and variant interpretation Move away from single gene tests using diverse techniques to multiplex panel-based next-generation sequence analysis during SMP1 Ability to adapt technology throughout SMP1 in order to incorporate additional genetic markers (eg, extended scope of *BRAF*, addition of *DDR2*) for specific add-on studies	Test failures due to variation in performance of and variations in tissue sample quality increased workload and time taken for analysis Achieving delivery of clinically relevant turnaround times proved challenging during SMP1
Data collection	Dataset drawn from existing information standards (such as the Cancer Outcomes and Services Dataset) with data item definitions according to the NHS Data Dictionary Electronic test request/report system established between clinical and laboratory sites, minimising duplication of data entry and risk of transcription errors Nationwide/cross-border network for data registration and submission established	Excessive number of data items in SMP1 dataset and focus on core rather than tumour-type-specific data items led to loss or omission of important information such as smoking history and performance status Lack of unified electronic patient record as single source of individual data items increased workload for sites
Data analysis	Large patient numbers allowing in-depth analysis for particular genetic aberrations, such as ► Relationship to clinical staging and demographics ► Relationship to other genetic modifications ► Paired samples taken from the same patient Hypothesis-generating for ongoing research	No multivariate survival analysis
Benefits to participating patients and staff	Upfront genetic diagnosis of tumour samples enabling potential access to new therapies, trials and translational research including National Lung Matrix Trial through SMP2 pre-screening Increased awareness of role of somatic mutation analysis in cancer care Creation of a collaborative, multidisciplinary knowledge network for stratified medicine	Generation of genetic data of unknown significance, for which no known treatment or trial-based approaches are available

NLMT, National Lung Matrix Trial.

being available for analysis, and the possibility that we would be more likely to recruit patients not approached for other studies due to the concern of information overload in those already participating in other research. Significant losses of important clinical information included smoking status in lung cancer, ER/PR/HER2 status in breast cancer, tumour site in CRC and loss of distinction between low-grade and high-grade serous ovarian cancers. Since implementation of this study, the collection of detailed data on each cancer diagnosed has become more established in the UK with the widespread adoption of the Cancer Outcomes Services Dataset (COSD). It is imperative that cancer treatment centres collect the key data elements within COSD which could support studies like SMP1 in the future.

SMP1 served as the basis for inception and implementation of the current second phase, SMP2, which provides patients access to molecular pre-screening of surplus diagnostic lung cancer biopsy or cytology cell block samples to inform entry to the National Lung Matrix Trial (NLMT).[11] The NLMT directly incorporates many aspects of the SMP1 infrastructure, this time facilitating multi-arm, molecularly stratified clinical trial design for patients with advanced stage non-small cell lung cancer (NSCLC). Its key aim is to demonstrate the

feasibility of combining molecular testing with clinical trial enrolment and translational progress on a national level, with its advances reported in tandem with other pioneering Cancer Research UK translational and clinical trial programmes such as TRACERx, DARWIN and PEACE.[15–17]

Author affiliations
[1]Department of Medical Oncology, The Christie NHS Foundation Trust, Manchester, UK
[2]Manchester Experimental Cancer Medicine Centre, Manchester, UK
[3]Division of Molecular and Clinical Cancer Sciences, University of Manchester, Manchester, UK
[4]Cancer Research UK, London, UK
[5]Southampton Experimental Cancer Medicine Centre, Southampton, UK
[6]Glasgow Experimental Cancer Medicine Centre, Glasgow, UK
[7]Department of Respiratory Medicine, Queen Elizabeth University Hospital, Glasgow, UK
[8]Institute ofInfection, Immunity and Inflammation, University of Glasgow, Glasgow, UK
[9]Cancer Research UK Cambridge Institute, University of Cambridge, Cambridge, UK
[10]Cancer Research UK Cambridge Centre and Cambridge Experimental Cancer Medicine Centre, Cambridge, UK
[11]Addenbrooke'sHospital, Cambridge University Hospital NHS Foundation Trust, Cambridge, UK
[12]Department of Histopathology, University Hospital of South Manchester NHS Foundation Trust, Manchester, UK
[13]Christie and Salford Royal NHS Foundation Trusts, Manchester, UK
[14]Department of Pathology, Queen Elizabeth University Hospital, Glasgow, UK
[15]Cancer Research UK Beatson Institute, Glasgow, UK
[16]Institute of Cancer Sciences, University of Glasgow, Glasgow, UK
[17]Leeds Experimental Cancer Medicine Centre, Leeds, UK
[18]St James's University Hospital, Cancer Research UK Clinical Cancer Centre, Leeds, UK
[19]Department of Cellular Pathology, Leeds Teaching Hospitals NHS Trust, Leeds, UK
[20]School of Medicine, University of Leeds, Leeds, UK
[21]Department of Pathology and Tumour Biology, Leeds Institute of Cancer and Pathology, University of Leeds, Leeds, UK
[22]School of Medicine, University of St Andrews, St Andrews, UK
[23]Edinburgh Experimental Cancer Medicine Centre, Edinburgh, UK
[24]Department of Medical Oncology, Royal Marsden Hospital, London, UK
[25]Institute of Cancer Research Experimental Cancer Medicine Centre, London, UK
[26]Velindre Hospital, Cardiff University, Cardiff, UK
[27]School of Medicine, Cardiff University, Cardiff, UK
[28]Cardiff Experimental Cancer Medicine Centre, Cardiff, UK
[29]Academic Department of Surgery, University of Birmingham, Birmingham, UK
[30]Birmingham Experimental Cancer Medicine Centre, Birmingham, UK
[31]Section of Biostatistics and Epidemiology, Leeds Institute of Cancer and Pathology, University of Leeds, Leeds, UK
[32]Royal Brompton and Harefield NHS Foundation Trust, London, UK
[33]National Heart and Lung Institute, Imperial College, London, UK
[34]Lung Unit, Royal Marsden Hospital, London, UK
[35]Department of Histopathology, Papworth Hospital, Cambridge, UK
[36]Department of Histopathology, University Hospitals Birmingham NHS Foundation Trust, Birmingham, UK
[37]Department of Pathology, Laboratory Medicine, Royal Infirmary of Edinburgh, Edinburgh, UK
[38]All Wales Genetics Laboratory, Cardiff, UK
[39]Genomic Medicine, School of Medicine, Dentistry and Biomedical Sciences, Queen's University, Belfast, UK
[40]West Midlands Regional Genetics Laboratory, Birmingham Women's NHS Foundation Trust, Birmingham, UK

Presented at
Oral presentations: National Cancer Research Institute annual meeting, Liverpool 2015. British Thoracic Oncology Group annual meeting, Dublin 2015. The Farr Institute annual conference, St Andrews 2015. National Cancer Research Institute annual meeting, Liverpool 2014. 6 th European Congress of Pathology, London 2014. 7th Joint Meeting of the British Division of the International Academy of Pathology and the Pathological Society of Great Britain and Ireland, Edinburgh 2013. Poster presentations: ASCO general meeting, 2014. 26 th EORTC-NCI-AACR Symposium on Molecular Targets and Cancer Therapeutics, Barcelona 2014.

Acknowledgements CRUK would like to thank all the patients who participated in phase 1 of the Stratified Medicine Programme as well as everyone at the clinical and laboratory sites that worked so hard to make the programme a success and the members of our numerous advisory groups and governance board.

Collaborators We have attempted to make the following list of collaborators at the clinical and technology hub sites as exhaustive as possible and apologise for any unintentional omissions: Pauline Rehal, Samantha Butler, Matthew Smith, Rachel Doak, Anna Tanska, Graham Halford, Lisa James, Chris Kotara, Gareth Masson, Sam Clokie, Jennie Bell, Fiona Macdonald, Mike Griffiths, David Gonzalez de Castro, Lisa Thompson, Debbie Mair, Suzanne Lillis, Dorte Wren, Robert Hollifield, Keeda Dover, Manisha Maurya, Damian Brooks, Belen Gomez, Lisa Grady, Thomas Jones, Chantal Hooper, Daphne Webster, Jolyon Travis, Stephanie Ogwuru, Jana Gazdova, Denise Collins, Elaine Chapman, Lisa Leavey, Paula Proszek, Sanna Hulkki, V Peter Collins, Ash Ibrahim, Kat Brown, Jo Burge, Karen Burnett, Ginny Devonshire, Ellen Moseley, Bev Haynes, Charlotte Hodgkin, Merche Jimenez-Linan, Linda Jones, Gilly Kenyon, Betania Mahler-Araujo, Karen Payne, Jo Piper, Doris Rassl, Sue Richardson, Ed Rytina, Anne Warren, Liz Coker, Gemma Godsall, Mark Arends, Amanda O'Neill, Katy Rintoul, Donna Goymer, Julie Taylor, Claire Matthews, Harshil Bhayani, Tina Osalador, Zakiya Niwaz, Anna Higgins, Olivia Bamsey, the Thoracic Surgical Team at the Royal Brompton Hospital, Janine Salter, Louise Renouf, Glenn Noel-Storr, Helen Roberts, Kasia Gierejko, Paola Knapman, Andrew Wotherspoon, Gordon Stamp, Ayoma Attygailye, Steve Hazell, Peter Osin, Ash Nerurkar, Steven Francis, Marion Runde, Jo Arch, Xavier Chitnis, Bernard Siu, Debra Townsend, Laura Hennelly, Natalie Taylor, Bernadette Johnson, Susie Banerjee, Lynda Pyle, Monica Hamill, Jenny Gyertson, Angela George, Krishna Patel, Karla Pearce, Kim Edmonds, Sarah Sarker, Rosalind Eeles, Liz Bancroft, Natalie Taylor, Sarah Thomas, Yukie Kano, Lisa Rowland, Karen Brooks, Sanjay Popat, Mary O'Brien, Jaishree Bhosle, Kathy Priest, Bee Ayite, Jo Severn, Helen Beedham, Nicky Lucas, Kim Tye, Alison Lorentzos, Janine Webb, Sarah Kerr, Lisa Corestav, Diego Bottero, Laura Jell, Janet Thomas, Cheryl Marriott, Neil Rajah, Andy Cole, Dieu Ly, Philippe Taniere, Brendan O'Sullivan, Clare Swift, Frances Hughes, Desley Neil, Andrew Hanby, Roz Banks, Dolapo Ajayi, Alison Barclay, Bexley Wing Phlebotomy Team, Leeds Julia Newton Bishop, Debbie Beirne, Andrew Bernard, Maxine Berry, Jo Bentley, Tim Bishop, Amy Chambers, Jude Clarke, Anne Crossley, Narinder Gahir, Debbie Gibson, Rona Good, Konstantina Grosios, Geoff Hall, Pat Harnden, Kate Hasler, Damien Hindmarch, Sharon Jackson, Colin Johnstone, Anne-Marie Jones, Gil Lambert, Sally Lane, Nicola McNicholas, Rebecca Millican-Slater, Cath Moriaty, Alex Newsham, Kara O'Connell, Leeds Oncology Outpatient Clinic Teams, Leeds Oncology and Surgery Consultants, Lisa Ripley, David Sebag-Montefiore, Mary Simpson, Val Speirs, Joh Sugden, Lauren Tate, Emma Tidswell, Chris Twelves, Christy Walker, Barry Waterhouse, Martin Waugh, Louise White, Elizabeth Wright, Jane Rogan, Garry Ashton, Caron Abbey, Michelle Greenhalgh, Daisuke Nonaka, Elwyn Shing, Carmen Gibbard, Georgina Burton, Naomi Fawkes, Angela Marsden, Rachael Waddington, Phil Harrison, Shahrzad Moghadam, Kate Murray, Sarah Brown, Christy Mitchinson, Richard Booton, Rajesh Shah, Fiona Blackhall, Noel Clarke, David Harrison, Anca Oniscu, William Wallace, Frances Rae, Craig Marshall, Linda McLeod, Morag Charles, Sarah Jane Sutherland, Carol Dawson, Paul Mitchell, Alex MacLellan, Sandra Muir, Lynne Johnstone, John O'Connor, Shirley Johnstone, Jim McPherson, Jane Hair, Massimo Pignatelli, Roma Armstrong, Karin Oien, Jeff Evans, Margaret Burgoyne, Karen Blessing, Fraser Duthie, Colin Moyes, Elizabeth Mallon, David Millan, Fiona Roberts, Morag Seywright, Siobhan Fraser, Craig Dick, Ian Ford, Sharon Kean, Marion Flood, David Grant, Claire McDonald, Tom Moffat, Hugh McLelland, Alistair Kyle, Graham Cameron, Martin Wright, Stephen Kenny, Karen McAuslan, Andrew Jones, Ted Fitzsimons, Fiona Graham, Alexandra Bell, Phil Duffy, Alec Fisher, Alexis Smith, Elaine Shannon, Bryan Woods, Colin Hutchison, Angela Booth, Lyndsay Duffy, Gillian McCulloch, Hudda Sadiq, Susan Deakin, Glasgow Pre-op Assessment Staff, Glasgow Cardiothoracic department Staff, Steven Haywood, Glasgow Pathology Staff, Malcolm Mason, John Chester, Alison Parry-Jones, Abby MacArthur, Suzanne Williams, Cardiff/ Swansea/ Gwent local Wales Cancer Bank nurses, clinicians and histology lab staff, David Griffiths, Fiona Morgan, Hazel Bailey and the CANISC team.

Contributors All coauthors have made substantial contributions to the conception or design of the work, or the acquisition, analysis or interpretation of data. Drafted the work or revising it critically for important intellectual content. Final approval of the version published. Offered agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. All authors and coauthors

have contributed to the planning, conducting and reporting of the work described in the article.

Funding This study was supported by Cancer Research UK, AstraZeneca and Pfizer UK.

Competing interests Funding for the Stratified Medicine Programme is acknowledged from Cancer Research UK and programme founding partners AstraZeneca and Pfizer. For hosting the Stratified Medicine Programme data, thanks to the National Cancer Registration Service Eastern Office, Jim Davies and Steve Harris at the University of Oxford Department of Computer Science. AGN was supported by the National Institute of Health Research Respiratory Disease Biomedical Research Unit at the Royal Brompton and Harefield NHS Foundation Trust and Imperial College London. PJ, ES and CL have all been employed by Cancer Research UK in the past. FB was supported by Cancer Research UK Lung Cancer Centre of Excellence Funding. WAW was supported by Lothian NRS BioResource. SP acknowledges NHS funding to the NIHR Biomedical Research Centre at The Royal Marsden and the ICR. NPW was supported by Yorkshire Cancer Research. KGB was supported by a NHS Research Scotland Senior Fellowship.

Patient consent Not required.

Ethics approval UK Research Ethics Committee reference 11/EE/0202.

Provenance and peer review Not commissioned; internally peer reviewed.

Data sharing statement CRUK retains access to unpublished data from this study. Requests can be made to interrogate this further for individual cancers assessed: breast, lung, colorectal, prostate, ovarian and melanoma.

REFERENCES

1. Jänne PA, Yang JC, Kim DW, *et al*. AZD9291 in EGFR inhibitor-resistant non-small-cell lung cancer. *N Engl J Med* 2015;372:1689–99.
2. Robert C, Karaszewska B, Schachter J, *et al*. Improved overall survival in melanoma with combined dabrafenib and trametinib. *N Engl J Med* 2015;372:30–9.
3. Sequist LV, Soria JC, Goldman JW, *et al*. Rociletinib in EGFR-mutated non-small-cell lung cancer. *N Engl J Med* 2015;372:1700–9.
4. Thress KS, Paweletz CP, Felip E, *et al*. Acquired EGFR C797S mutation mediates resistance to AZD9291 in non-small cell lung cancer harboring EGFR T790M. *Nat Med* 2015;21:560–2.
5. Nowak F, Soria JC, Calvo F. Tumour molecular profiling for deciding therapy - the French initiative. *Nat Rev Clin Oncol* 2012;9:479–86.
6. Tuff-Lacey A, Shaw E, Cummings R, *et al*. A collaborative approach to enabling stratified cancer medicine in the UK. *Drug Discov Today* 2015;20:1414–8.
7. Lindsay CR, Shaw E, Walker I, *et al*. Lessons for molecular diagnostics in oncology from the cancer research UK stratified medicine programme. *Expert Rev Mol Diagn* 2015;15:287–9.
8. Cameron D, Stead M, Lester N, *et al*. Research-intensive cancer care in the NHS in the UK. *Ann Oncol* 2011;22:vii29–35.
9. Ding L, Getz G, Wheeler DA, *et al*. Somatic mutations affect key pathways in lung adenocarcinoma. *Nature* 2008;455:1069–75.
10. Chang DK, Grimmond SM, Evans TR, *et al*. Mining the genomes of exceptional responders. *Nat Rev Cancer* 2014;14:291–2.
11. Middleton G, Crack LR, Popat S, *et al*. The National lung matrix trial: translating the biology of stratification in advanced non-small-cell lung cancer. *Ann Oncol* 2015;26:2464–9.
12. Decatris MP, Farrugia D, O'Byrne KJ. Clinician perspective on molecular profiling of non-small-cell lung cancer. *J Clin Oncol* 2016;34:884–6.
13. Lim C, Sung M, Shepherd FA, *et al*. Patients with advanced non-small cell lung cancer: are research biopsies a barrier to participation in clinical trials? *J Thorac Oncol* 2016;11:79–84.
14. Mostert M, Bredenoord AL, Biesaart MC, *et al*. Big Data in medical research and EU data protection law: challenges to the consent or anonymise approach. *Eur J Hum Genet* 2016;24:1096.
15. Hiley CT, Le Quesne J, Santis G, *et al*. Challenges in molecular testing in non-small-cell lung cancer patients with advanced disease. *Lancet* 2016;388:1002–11.
16. Abbosh C, Birkbak NJ, Wilson GA, *et al*. Phylogenetic ctDNA analysis depicts early-stage lung cancer evolution. *Nature* 2017;545:446–51.
17. Jamal-Hanjani M, Wilson GA, McGranahan N, *et al*. Tracking the evolution of non-small-cell lung cancer. *N Engl J Med* 2017;376:2109–21.