Identification of the Genetic Basis of Multicentric Tumourigenesis

James Whitworth

A thesis submitted to the University of Birmingham for the degree of Master of Science by Research

School of Clinical and Experimental Medicine

College of Medical and Dental Sciences

University of Birmingham

July 2015

Abstract

Multiple primary malignant tumours (MPMT) are an indicator of potential inherited cancer susceptibility and occur at appreciable frequency among unselected cancer patients and, particularly, among referrals to cancer genetics services. However, there is little information on the clinical genetic evaluation of cohorts of MPMT patients representing a variety of tumour types. A referral based case series was ascertained and is described. Service-based molecular genetic testing had demonstrated a pathogenic germline variant in a cancer predisposition gene in less than a quarter of referrals. To assess for evidence of unidentified variants in those who tested negative, comparisons were made with those who tested positive, revealing considerable overlap between the two groups with respect to clinical characteristics indicative of an inherited cancer syndrome. A subset of unexplained MPMT cases was subsequently analysed for mutations in *TP53*, *PTEN* and *CDKN2B* but none were detected. MPMT individuals may receive negative genetic test results for a number of reasons, which are discussed. They include referral bias and MPMT cases from a population based registry were also analysed and compared with the referral based series. The increasing application of next generation sequencing techniques in clinical services is likely to address many of these issues.

Contents

List of Figures

List of Tables

List of definitions

Variant – A change in the sequence or structure of a gene as compared to a reference dataset. Does not imply functional effect or pathogenicity.

Mutation – As per variant, though generally taken in the field of Clinical Genetics to imply that the change is deleterious and/or pathogenic. Also refers to the process of genetic changes occurring.

Missense – A genetic change resulting in the substitution of one amino acid for another in a gene's protein product.

Non-sense – A single nucleotide genetic change that results in the generation of a stop codon and a truncated protein product.

Frameshift – A genetic change resulting in loss of the translational reading frame of a gene with consequent changes to amino acids incorporated through translation downstream of the change. Frequently results in a premature stop codon and a truncated protein product.

Dominant inheritance - A condition where mono-allelic mutations are sufficient for the phenotype to manifest.

Recessively inherited condition – A condition where bi-allelic mutations are necessary for the phenotype to manifest.

Constitutional variant/mutation – A variant/mutation that exists in a significant proportion of cells in an individual.

Germline variant/mutation – A variant/mutation that is present in germ cells. Inheritance of such a variant/mutation by offspring results in a constitutional variant/mutation in that individual. Genetic testing of blood samples is often interpreted as testing a tissue that is representative of cells present at conception and identified variants/mutations may be referred to as germline.

***De novo* mutation** – A constitutional mutation that occurred in the parental gamete or post conception and is therefore not present in either parent.

Cancer predisposition gene – A gene where constitutional deleterious mutations confer predisposition to developing cancer. This may be through inheritance from a parent or a post conception event.

Penetrance – The degree to which the phenotype associated with mutations manifests in mutation carriers. A low penetrance mutation may lead to no observable phenotypic effect whereas a high penetrance mutation leads to an associated phenotype in most or all cases.

Amplicon – A genomic region amplified by an individual polymerase chain reaction primer pair.

Next generation sequencing – A group of techniques that utilise simultaneous sequencing of large numbers of nucleic acid sequences. Resulting sequence readouts are computationally aligned to a reference sequence for interpretation. The sequenced nucleic acids vary according to technique and can include effectively all genomic regions (whole genome sequencing), all protein coding regions of the genome (whole exome sequencing) or a selected group of genes (gene panel).

Much of Chapter 1 Section 4, the majority of Chapters two and three and a proportion of Chapter five have previously been published[1].

Chapter 1 – Introduction and aims

Section 1: Genetic predisposition to cancer

Cancer as a genetic disease

The concept of the cancer as a clonal expansion of cells that have undergone genomic changes conferring malignant properties is now broadly accepted. The development and testing of this hypothesis has been a process guided by the application of new technologies, in this case to analyse cellular genetic material at increasing resolution and in increasing quantity. Advances in genomic techniques are allowing the next step in this process to take place.

In the nineteenth century, microscopic analysis led to the observation that chromosome aberrations can occur in malignant cells[2]. Boveri was one of the first to put forward the idea that such aberrations might be a key causative factor in tumourigenesis. By studying abnormal mitoses in sea urchin embryos, he hypothesized that abnormal cellular properties, including malignancy, were conferred by an unbalanced chromosome complement. His work included assertions regarding “inhibiting chromosomes” i.e. those that normally act to suppress cell division and “stimulatory chromosomes,” which change a cell’s relationship with its external environment to encourage a proliferative state. These

ideas were remarkably prophetic of the proto-oncogenes and tumour suppressor genes (see below) that later became familiar[3]. It was not until the 1960's that a specific chromosomal abnormality was associated with a particular tumour when the Philadelphia chromosome (resulting from a translocation of chromosomes 9 and 22) was identified as present in all studied chronic myeloid leukemia patient samples[4].

Abnormal chromosomes may be a result of the genomic instability caused by tumourigenic processes but may also, as in the case of the Philadelphia chromosome, be important initiating events. With the development of gene sequencing techniques it became possible to study the molecular consequences of such events at the gene level and also define causative genetic abnormalities, not visible by chromosome analysis, that occur at the nucleotide level.

Oncogenes, tumour suppressor genes and cancer predisposition

A crucial step in understanding how genetic changes in tumours lead to cancer was the development of the concept of the oncogene, a gene that has undergone change rendering it as contributory to tumour development. These changes result in enhanced or altered function of a given oncogene's normal counterpart, termed a proto-oncogene. Proto-oncogenes are involved in a variety of cellular processes pertinent to cell growth/proliferation including cell cycle regulation and growth signalling. Oncogenes were initially identified through analysis of tumour cells whose malignant properties had been induced by a retrovirus. The tumorigenic potential of such viruses was found to be due to one component gene of the virus,

designated the oncogene[5]. Further research revealed that orthologues of the viral oncogenes were present in normal cells and were labelled proto-oncogenes[5].

The discovery of the other main class of gene significant in the development of cancer, the tumour suppressor gene (TSG), was to provide the key step in the understanding of inherited tumours. In its normal state, a TSG often functions to inhibit cell proliferation but loss of function mutations compromise this role and promote tumourigenesis. In some cases, inactivation of a TSG directly leads to cellular attributes that favour malignant transformation but in other cases (e.g. inactivation of a TSG involved in DNA repair mechanisms with resulting failure to repair mutations in other TSGs or proto-oncogene) the pro-tumourigenic process is indirect.

Whilst some genetic changes appear particularly important in conferring tumour defining properties to cells, a cell's transition from normal to malignant is a multi-step process. A source of much debate, often based on epidemiological evidence, has been how many changes are required for the completion of this process. Work by Nordling observed cancer mortality correlating with age and estimated that, on average, six mutational events in a given cell were required for a cancer to occur[6]. The work only studied certain cancer types and observed that many malignancies did not conform to this model.

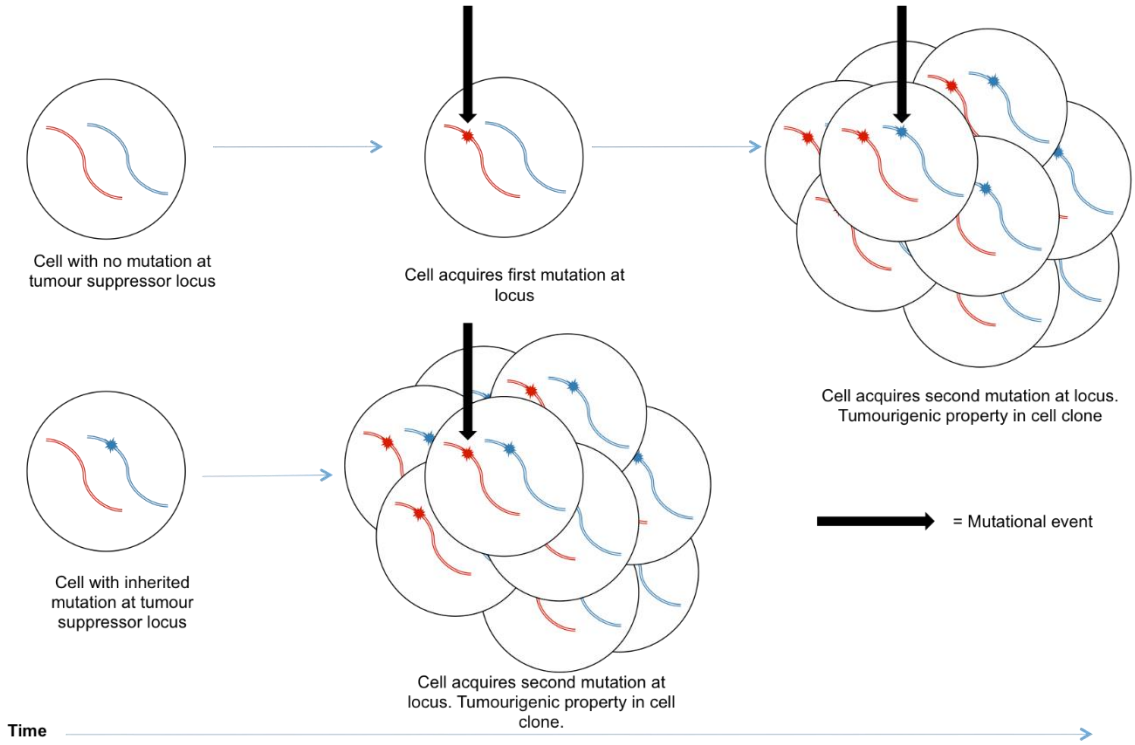
Seminal work by Knudson et al suggested that at least one cancer type (retinoblastoma, a rare childhood cancer of the eye) could be caused by two

mutational events affecting different alleles at the same locus. By observing tumour frequencies in presumed inherited and sporadic cases and estimating the mutation rate at the relevant theoretical locus, a model was proposed whereby those who inherit a mutated TSG allele from a parent require only a single mutation event ("hit") in the other allele for a tumour to be initiated. Sporadic cases, on the other hand, must acquire both hits post conception[7] (see Figure 1). This model explains the very high risk of retinoblastoma in individuals with the familial form where bilateral tumours occur, and the later onset unilateral tumours in sporadic cases.

The causative genetic factor in familial retinoblastoma is a constitutional aberration affecting the *RB1* gene. This gene was identified through analysis of retinoblastomas with Knudson's two hit hypothesis in mind. There was previous evidence that a section of chromosome 13 was the area affected by the hypothesized second hit. Some retinoblastoma patients had been shown to harbour a constitutional deletion in this area[8] and acquired loss or partial deletion of this region of chromosome 13 had been shown in retinoblastoma tumour cells[9]. Identification of the *RB1* gene within the target region was followed by demonstrating that inherited cases had an inactivating constitutional mutation and a second hit in the tumour cells, whereas tumours from non-inherited sporadic cases harboured inactivating hits in both *RB1* alleles that were not present in the patient's normal cells (so both events were somatic)[10].

Knudson's work and the subsequent identification of the *RB1* tumour suppressor gene[10] was a key step in the development of inherited cancer genetics. Apart from highlighting the role of TSGs in cancer pathogenesis, it demonstrated that inherited constitutional mutations causing cancer predisposition were identifiable through the study of affected families and that the genes affected by them can also be implicated in sporadic tumours.

Figure 1 – Knudson's two hit model



The discovery of *RB1* prompted an extensive and ongoing search for further cancer predisposition genes (CPGs) that has yielded findings of relevance to affected individuals and also tumours occurring outside of the inherited context. Identified genes may not conform to the two hit TSG model and a number of constitutionally activated proto-oncogenes have been found to be associated with

familial cancers (e.g. those in the *RET* gene leading to Multiple Endocrine Neoplasia Type 2[11,12]). The search for inherited cancer genes has focused on individuals with clinical characteristics of inherited cancer predisposition (e.g. family history of cancer, multiple tumours, phenotypic markers and young age at diagnosis) but advances in genetic technology are providing the means for large scale mutation screening in individuals without classical features of an inherited cancer syndrome.

Identifying cancer predisposition genes

A variety of approaches have been employed to identify CPGs including candidate gene analysis, positional cloning, positional-candidate gene analysis and comprehensive sequencing with next generation sequencing (NGS) technologies. Earlier studies concentrated on large families with high penetrance cancer phenotypes and used genetic linkage studies to identify regions that segregated with cancer predisposition in the family. In some cases (e.g. *RB1*) the putative localisation of the cancer predisposition gene was supported by deletion/allele loss analysis in tumour material. Having defined an interval of interest, all genes within the region were then sequenced until the gene containing germline mutations was identified. As illustrated by *RB1*, some CPGs are targets for somatic mutations in sporadic tumours. A typical candidate gene approach would be to take a gene known to be somatically mutated in a sporadic cancer and test whether inherited cases harbour constitutional mutations in that gene (e.g. *PBRM1* mutations in familial renal cell carcinoma[13]). A positional-candidate gene approach has been used widely, particularly since the advent of detailed maps of the human genome,

as following the identification of a candidate region of linkage, the genes within the target interval can be prioritised according to their relevance to cancer and analysed accordingly.

High throughput NGS techniques utilise massively parallel sequencing of short genomic sequences that are computationally aligned to a reference sequence to identify variants in the studied sample. They have enabled sequencing of whole genomes, the coding regions in their entirety (whole exomes) or a selected series of genes (through targeted amplification and capture of genomic regions of interest in the latter two techniques) in cases where genetic linkage studies are not possible, for example if DNA is only available from a single affected family member. Although the challenges of correctly interpreting the resulting large numbers of rare genetic variants should not be underestimated, these assays have greatly facilitated the identification of CPGs in research studies as well as in diagnostic laboratories. Their utility is illustrated by the discovery of *POT1* as a melanoma predisposition gene. *POT1* was identified by the application of NGS techniques to families where multiple members were affected with this cancer. Whole exome or whole genome sequencing was performed in individuals belonging to 28 families where two or more melanomas had occurred and where there were three or more members available for analysis. This latter criterion increased the ability of the study to exclude potentially causative variants that were not present in affected cases (or were present in unaffected cases). The large number of variants identified through sequencing were filtered using bioinformatic techniques to exclude those that were unlikely to be the cause of melanoma.

Variants were excluded if they were common and therefore unlikely to explain the relatively rare phenomenon of familial melanoma. They were also not retained if they were unlikely to affect the function of the corresponding protein product. 320 genes were found in the families that had variants fitting the above criteria. Crucially, only five of these were shared by more than one of the families. Further analysis focused on these and demonstrated that the variants in one of them (*POT1*) could be found in further families with familial melanoma and disrupted biological function of the gene[14].

Methodologies as described above generally begin without a hypothesis in terms of which gene may harbour the causative mutation. Initial analysis, particularly for NGS approaches, is likely to generate gene lists for further investigation and selection of putative causative genes at this stage is aided by candidature provided by existing sources of evidence. This might be expression of the gene in the organ of interest or involvement of it in a cellular process relevant to cancer (e.g. DNA repair). Observation of somatic alterations in a cancer type corresponding to that seen in the studied family may also prompt candidature. An example of this is the identification of *RB1* as illustrated above and the knowledge that *TP53* was recurrently mutated in many cancer types was helpful in the genetic characterisation of Li-Fraumeni syndrome, caused by constitutional mutations in that gene[15]. In recent years, molecular characterisation of tumours has become more extensive through cancer genome projects and is likely to further assist with identification of CPGs in this way. The cBioportal for example, contains information from over 20,000 sampled tumours from various projects[16].

Risks associated with mutations in cancer predisposition genes

While NGS technologies and improved candidature have enhanced the rate of novel CPG discovery, mutations in many of them have been estimated to cause lower cancer risks than some of the earlier discoveries such as *APC, VHL* and *TP53*. The majority of high risk cancer predisposition genes affecting large numbers of families may have been discovered, leaving more recent and future discoveries falling into a number of categories according to the risk that mutations affecting them confer.

Newly identified high risk genes are more likely to be rare and account for a smaller proportion of overall cancer burden. Current sequencing and bioinformatic techniques applied to small numbers of families are well placed to identify them. Their clinical utility will be significant for affected families and can provide insights into non-familial tumours of similar morphology.

Mutations in genes leading to moderate levels of cancer risk are unlikely to cause a characteristic cancer in all carrying members of families where they are present. Consequently, study designs relying on analysis of variants that are shared between affected family members may be less informative. Studies comparing variants seen in large series of patients with a particular cancer with a control series can reveal significant association of variant with tumour without necessarily reflecting a very high risk of that tumour developing. Examples of predisposition genes identified in this way include *BRIP1* and *PALB2*, reported to confer a relative breast cancer risk of 2 and 2.3 respectively[17,18]. Interestingly, further

observation of *PALB2* mutation carriers has revised this estimated risk to a much higher level[19].

Identification of the genetic basis of highly penetrant cancer predisposition phenotypes can have a major effect on management of affected families (see below) but only impact on a small minority of patients with the relevant cancer. Genome wide association studies (GWAS) of large cohorts of cancer patients have been employed to identify common genetic variants that predispose to specific cancers. The identification of such susceptibility alleles can provide pointers as to molecular pathways significant to particular tumours but generally haven't been translated into clinical practice because the increased risk associated with each variant is small. Ultimately, clinical utility might be provided by identifying individuals who harbour multiple low risk variants that combine to put the individual at a significantly increased risk. One analysis to assess this potential combined risks derived from 77 variants previously associated with breast cancer in GWAS to produce a combined risk score. The score was used to stratify over 30,000 breast cancer cases and controls into quintiles. It was found that in those without a family history, the highest scoring quintile had significantly higher lifetime breast cancer risk (16.6%) than those in the lowest scoring quintile (5.2%). This difference was more pronounced in those with a first degree relative with breast cancer (8.6% vs 24.4%)[20]. Risk estimates at the level of those assigned to the higher risk group approach those deemed sufficient for intervention in the cancer genetic clinic.

Section 2: Clinical utility of cancer predisposition gene analysis

Identification of CPGs through the study of affected individuals and families has provided the opportunity to extend genetic analysis to large numbers of individuals, often without a family history, who are considered to be at risk of an inherited cancer syndrome (e.g. a significant proportion of patients with phaeochromocytoma/paraganglioma without a relevant family history have a causative constitutional mutation[21]). This consideration is generally based on factors indicative of genetic predisposition and/or a particular syndrome such as young ages at diagnosis, multiple tumours and a particular tumour spectrum occurring in the family.

The case to perform germline genetic testing may be particularly compelling where rarer tumours cluster within the same family as alternative causes are less likely. The picture is more ambiguous where common tumours cluster. Such a scenario may represent inherited predisposition or be the result of higher population incidences of particular tumours leading to occurrence in multiple family members. Sporadic tumours that appear consistent with a constitutional genetic cause may be termed phenocopies and are more likely to be present where environmental factors frequently lead to a particular tumour type. However, there is not a simple relationship between how common a specific cancer is and whether it is genetic or environmental in origin. An assessment of the proportion of cancer cases attributable to 14 common preventable environmental exposures has shown relatively low figures for many of the most frequent tumours occurring in the

population e.g. breast (26.8%) colorectal cancers (54.4%)[22]. In many cases, clinical genetic testing is able to distinguish between clustering of common tumours caused by mutations in CPGs and due to other mechanisms.

Genetic testing is termed diagnostic in an individual who has previously been diagnosed with a cancer and where a genetic explanation is sought. Predictive testing generally interrogates the genetic status of an unaffected person for a causative variant previously identified in a family member, though advances in testing technology and knowledge of CPGs is likely to lead to much more predictive testing where such a mutation has not been found in a relative. In the ideal scenario, results from such analysis can provide accurate risk assessment and potentially insight into tumour biology in the tested individual. Clinical utility may therefore be derived in a number of ways.

Information as therapy

Risk information based on genetic test results may enhance management by providing patients with a clear explanation for cancers that are often of early onset and affect multiple family members. Negative results can provide reassurance but often leave open the possibility of unidentified pathogenic variants such that a genetic predisposition cannot be definitively excluded. A genetic diagnosis, although potentially increasing perceived future cancer risk, often comes with a defined spectrum of associated tumours. These specific risks, which can be managed, may be preferable to a generalised increased cancer risk that can be perceived by the patient prior to testing. A notable exception is Li-Fraumeni

syndrome, where the cancer risks are very broad and where there is little clinical surveillance currently demonstrated to be of benefit. Uptake of predictive testing for this condition however, is broadly similar to other commonly performed predictive tests where better management can be offered[23]. A systematic review of psychological outcomes in women with a family history of breast cancer undergoing genetic testing found a significant reduction in psychological distress in women receiving negative results and little change in those receiving positive results[24]. A study of individuals undergoing predictive testing for *BRCA1* or *BRCA2* mutations found that 92% would recommend the process to others in the same situation[25]. Assessment of benefits such as this should be seen in the context of testing uptake, which has been shown to be around half of individuals eligible for predictive testing in the most commonly seen conditions in the genetics clinic[26,27]. Those not pursuing testing may not have gone on to have an equally positive experience.

Individuals consulting clinical services for assessment for a possible predisposition syndrome frequently do so in an attempt to provide a genetic diagnosis in the family. This gives the opportunity for relatives to predict and manage their risks and may be of benefit to the consulting patient even if their own prognosis is poor. An assessment of motivations for diagnostic testing in a series of colorectal cancer patients undergoing testing showed greater importance placed on this than a desire to increase certainty as to whether or not an increased risk of colorectal cancer was present in the consultand[28].

Finally, identification of mutations in CPGs in a potential parent may facilitate reproductive decisions or lead to testing for the variant in a foetus in utero (pre-natal diagnosis (PND)) or pre-implantation embryos resulting from in vitro fertilisation (pre-implantation genetic diagnosis (PGD)). Test results can affect decisions as to whether to continue a pregnancy or which embryos to select for implantation. This is a model largely used in severe (mainly non-cancer) childhood onset disorders and is less frequently applied to most cancer predisposition syndromes due to their generally later onset and more manageable manifestations. However, a number of adult onset cancer syndromes appear on the Human Fertilisation and Embryo Authority's approved list for PGD[29] and high levels of agreement with PND for Lynch syndrome among individuals at risk of Lynch syndrome[30] suggest that this may change in future.

Clinical surveillance

The current mainstay of management following the diagnosis of a cancer predisposition syndrome are interventions designed to prevent cancers occurring or diagnose them at a more treatable stage. The latter is achieved through regular clinical surveillance of at risk tissue. This may be via a number of modalities depending on the tissue or syndrome in question and imaging, endoscopic examination and biochemical analysis are frequently used. Frequency and age at which surveillance investigations are performed are guided by observational evidence from series of affected cases. The quality of such guidance is often compromised by the rarity of a condition and/or ascertainment biases influencing which patients are included in these series. The gold standard of a prospective

study of mutation carriers to observe may be difficult to achieve but collaboration between centres is likely to lead to more accurate assessments of risks and also screening outcomes. The effectiveness of surveillance programmes is currently uncertain for most predisposition syndromes. For more common conditions, larger cohorts that can be assembled in one or a few centres can provide greater clarity on the issue. In Lynch syndrome, a number of studies comparing screened and unscreened groups have been published, allowing a systematic review. The multiple studies showing reduction in colorectal cancer incidence and related mortality in screened (with regular colonoscopy) patients provide convincing evidence for the utility of this strategy. In rarer conditions, inference may have to be made from less direct sources of information. Von Hippel Lindau disease (VHL) is one such condition that predisposes affected individuals to central nervous system haemangioblastoma, phaeochromocytoma and renal cell carcinoma. Protocols for surveillance are widely used but no prospective follow up data comparing screened with unscreened patients exists. However, life expectancy in VHL patients has been observed by identifying cases through genetics service and cancer registries. Whilst a reduced life expectancy was noted (52.5 years), the mean survival increased by 16.3 years in patients diagnosed after 1990 when the genetic service registry, and therefore increased systematic surveillance protocols, was introduced[31].

Surveillance programmes for conditions with diverse tumour risks can lead to the proposal of complex screening programmes which may have poor acceptability for patients and whose outcomes are difficult to assess. An appealing approach is

that of a single modality screening test that can address these issues. A candidate for a modality like this is a whole body MRI scan, which is currently being investigated in Li-Fraumeni syndrome[32]. Crucial outcomes will include the rate of potentially significant findings and effects of resulting interventions. This is true for any screening test but a whole body approach combined with the range of tumour risks in Li-Fraumeni syndrome make them particularly pertinent.

Clinical surveillance has more potential to do harm (e.g. through unnecessary resulting surgery) where the penetrance of a mutation in a given cancer predisposition gene is not high. In Hereditary Leiomyomatosis and Renal Cell Carcinoma (HLRCC) caused by *FH* mutations, only 15-20% of mutation carriers develop kidney cancer but of those that do, many are at an advanced stage resulting in a poor prognosis[33]. As well as assessing screening programmes in as extensive series as possible therefore, an important area of research is risk stratification within inherited cancer syndromes. Stratification might be based on the particular variant in the causative gene (see below) or through constitutional variants in other genes that may influence cancer risk (modifier genes). Alternatively, acceptability, specificity and sensitivity of screening tests might be improved for those individuals at low risk by exploiting the phenomena of circulating tumour DNA. Identification of specific genetic markers of tumour cell origin in a predisposition syndrome such as HLRCC could facilitate an effective surveillance programme based on blood sampling.

Prophylactic surgery

In some predisposition syndromes, prophylactic surgery may represent the most effective preventative strategy. Utility and uptake of this can depend on a number of factors such as level of risk reduction from tissue removal, function (and loss thereafter) of the tissue in question and likelihood of complications following the procedure. These factors need to be considered against the efficacy of surveillance strategies as an alternative. Prophylactic surgery can result in dramatic reduction in tumour risk. In hereditary breast and ovarian cancer caused by *BRCA1* and *BRCA2* mutations, bilateral mastectomy is estimated to reduce the risk of breast cancer by around 90%[34]. Similar to surveillance, utility of surgery in other syndromes may be more difficult to estimate due to rarity of the condition and/or lack of an adequate control (no surgery performed) group with which to compare survival rates. In familial adenomatous polyposis, the risk of colorectal cancer[35] has been estimated at a level sufficient to warrant colectomy in all diagnosed cases, leaving a low number of cases with an intact colon for further study.

Pharmacological management

In cancer predisposition syndromes, the benefits of chemo-preventative strategies are especially likely to outweigh disadvantages resulting from side effects or economic cost. Use of preventative agents may be based on observations relating to tumours occurring outside of the familial context. Lower colorectal cancer rates in individuals taking long term aspirin[36] led to a trial of its use in Lynch syndrome, which leads to high risks of that malignancy. A significant reduction in bowel

cancer incidence was observed in the trial leading to the recommendation that this medication should be discussed with affected individuals[37].

Recent years have brought an increasing focus on the use of a patient's constitutional mutation status to guide medical therapy to both treat and prevent tumour development. This knowledge can provide insight into the biology of the tumour sufficient to prompt clinical trials. Though such therapies are currently limited in number, they are a source of wider optimism in cancer predisposition syndromes. Strategies generally involve the inhibition of an abnormally active/upregulated gene product or cellular pathway as exemplified by the use of vismodegib in Gorlin syndrome (basal cell nevus syndrome). This is an autosomal dominant condition associated, among a wide range of other features, with multiple early onset basal cell carcinomas (BCCs). Study of affected families demonstrated that it was caused by mutations in *PTCH1*, the gene product of which performs an inhibitory function in the hedgehog signalling pathway[38,39]. This pathway is a key regulator of cellular development in early life but is abnormally active in Gorlin related BCC cells due to a "second hit" in the wild type *PTCH1* allele[40]. *PTCH1* was subsequently shown to be mutated in most sporadic BCCs and a trial of vismodegib (an agent previously known to inhibit the hedgehog pathway) was shown to be efficacious. Subsequent trials have also shown reduction in BCC occurrence in Gorlin patients[41]. The example illustrates the development of therapy based on germline mutation status and also how identification of relevant mutations in rare syndromic tumours can inform knowledge and treatment of their sporadic counterparts.

Notable exceptions to strategies inhibiting an abnormally active/upregulated gene product or cellular pathway exist such as the example of poly ADP ribose polymerase (PARP) inhibitors in *BRCA1/2* related cancers, which disrupt a DNA repair mechanism (base excision repair). Tumour cells in *BRCA1/2* mutation carriers have generally acquired an aberration affecting their *BRCA1/2* wild-type allele and are therefore deficient in a different repair process (double stranded DNA repair by homologous recombination), rendering them susceptible to PARP inhibitor induced cell death compared to other cells in the same patient[42].

More extensive definition of tumour phenotypes associated with predisposition syndromes by improved molecular analysis should continue to yield abnormalities targetable by therapeutic agents. Sequencing of breast cancers with whole genome sequencing has previously revealed a distinct mutational signature in *BRCA1* or *BRCA2* related cancers that is consistent with the known DNA repair deficit in these patients[43]. Similar work may be rewarding on tumours from patients with other syndromes, perhaps where the function of the relevant gene is less well characterised. This should provide opportunities for stratification and treatment akin to those produced by the study of sporadic cancers.

Section 3: Cancer predisposition genes

The number of genes in which constitutional mutation leads to cancer susceptibility currently stands in three figures, though defining a CPG is not without difficulties. For every gene where mutation carriers have high risks, there are others where mutation carriers are more likely to remain unaffected by the cancer/s they are at risk from. Many constitutional genetic variants have been associated with increased risk of particular cancers through GWAS but the risk conferred by these is generally not at a level prompting the management strategies described above. A comprehensive review of CPGs was published by Rahman in 2014 and included genes where rare mutations confer a doubling of relative risk of cancer and lead to 5% of carriers being affected with cancer[44]. At the lower end of these risks, however, it is doubtful whether surveillance or prophylactic surgery would be of benefit.

Phenotypic effects of mutations in cancer predisposition genes

CPGs are involved in an array of cellular processes where aberrant function can lead to cancer defining phenomena such as genomic instability, disrupted cell cycle regulation or increased proliferation. Mutations in these genes can lead to phenotypic expression consistent with various models of inheritance seen in high penetrance genetic conditions. The majority of associated phenotypes are inherited in an autosomal dominant manner, often with unaffected mutation carriers seen in the family due to incomplete penetrance. Mutations in a gene causing a dominantly inherited syndrome may be embryonically lethal in the rare

scenario that one is inherited from both parents. A number of recessive inherited cancer syndromes have also been described, most notably colonic polyposis and colorectal cancer due to bi-allelic mutations in *MUTYH*[45]. There are other intriguing examples where the inheritance of a mutation in a CPG has contrasting effects depending on whether it is inherited in the mono-allelic or bi-allelic state. This might lead to a phenotype differing qualitatively or in terms of severity. Heterozygous *SDHB* mutations cause phaeochromocytoma and paragangliomas[46] whereas the inheritance of a mutation on both chromosomes leads to a neurodevelopmental disorder[47]. The tumour risks in bi-allelic cancer predisposition mutations may still be present but often unable to manifest themselves due to reduced life expectancy. Mono-allelic *ATM* mutations are associated with an increased risk of breast cancer whereas Ataxia Telangiectasia caused by bi-allelic inheritance causes a number of features such as ataxia and haematological cancers. However, breast cancer has also been reported in longer surviving cases[48].

Some inherited cancer syndromes such as Li-Fraumeni syndrome are associated with an increased risk of a wide range of cancer types but most conditions are currently known to lead to increased risk of a small number of specific tumours. Even Li-Fraumeni related cancers are among a set of four core malignancy types in 70% of cases[49]. The reason for this specificity is yet to be elucidated in most cases, though biological explanations include the restriction of gene expression/action to particular tissues and aberrant cellular mechanisms rendering

cells susceptible to further mutation through environmental exposures only relevant to particular organs.

Some phenotypic specificity may be explained by ascertainment biases influencing the study of CPGs and their associated tumour risks. Identification of such genes has usually been through the preferential study of families where there are a number of occurrences of the same tumour or group of tumours, restricting other possible associations. Furthermore, more likely to be studied are those cases where the phenotype is more severe e.g. earlier age of tumour diagnosis. The identification of novel CPGs in these scenarios is likely to underestimate the range of tumours caused by mutations in that gene and overestimate the clinical severity of harbouring them. These effects may be exacerbated by the effect of clinical criteria to guide access to genetic testing, which appears to have occurred in Lynch syndrome.

Lynch syndrome is a cancer predisposition syndrome conferring susceptibility to a variety of cancers, primarily colorectal. It is caused by heterozygous mutations in mismatch repair genes such as *MLH1*, *MSH2* and *MSH6*. Colorectal cancer is a common condition and clinical criteria have previously been used to identify those families likely to have tumours due to Lynch syndrome as opposed to an alternative cause. The Amsterdam criteria[50] were developed to provide consistency of reporting of suspected Lynch syndrome families and require a severe family history in order to be fulfilled. The Bethesda criteria[51] are designed

to prompt the initiation of molecular investigations for Lynch syndrome and incorporate a wider range of families while still requiring relatively strong evidence of a predisposition syndrome.

Where genetic investigations were/are less available, rationing of analysis is likely to be based on such criteria on the basis that those fulfilling them are more likely to harbour a causative mutation and represent a more efficient use of resources. This may lead to an overestimate of the tumour risks associated with identified mismatch repair mutations as those families where their effect is less severe are less likely to been eligible for testing. Those who receive testing (and test positive) may also have greater risks conferred by other modifying genetic variants and not only due to the identified mutation, overestimating its effect. Indeed, earlier studies estimated a higher risk of colorectal cancer than has more recently been reported. A large 1999 analysis of registry recorded Finnish mutation carriers showed a cumulative incidence of colorectal cancer of 82% by age 70, 68 times higher than the population from which the cases were drawn[52]. However, a 2009 assessment of mutation carriers identified through genetics clinics and corrected for ascertainment bias estimated a lower cumulative incidence of 66% to age 70[53].

Expansion of phenotype associated with mutations in a particular gene following initial discovery based on a narrow tumour spectrum is illustrated by BAP1, a recently described CPG. Its identification also illustrates the potential utility of somatic mutation databases to provide gene candidature. Subsequent association

with further tumour types demonstrates the ability of NGS techniques to aid rapid definition of a broader phenotype than that described by the original association.

Acting on observational data that suggested hereditary predisposition to uveal melanoma (UM) in a proportion of cases, Abdel Rahman et al sequenced BAP1 in a series of UM patients with a clinical indication of hereditary susceptibility. Candidature of BAP1 had been suggested by a number of lines of evidence including a previous study showing that around half of UM's had a somatic BAP1 mutation[54]. BAP1 is located on chromosome 3 and monosomy of this chromosome is frequently observed in these tumours[55]. 1/53 probands was found to have a truncating constitutional BAP1 mutation. Their UM demonstrated loss of the wild type allele and reduced protein product on immunohistochemistry staining, as did a lung adenocarcinoma diagnosed in the proband and a meningioma from a mutation carrying relative[56].

Since this discovery, mutations in this gene have been associated with a variety of other cancers, notably renal cell carcinoma (RCC). Popova et al identified a splice site mutation in a family with four individuals affected with RCC. They subsequently validated this finding by showing mutations in 11/60 families with aggregations of RCC in addition to tumours previously associated with BAP1[57]. The mutation identification was obtained through whole-exome sequencing where BAP1 was not proposed as an initial candidate, indicating that the non-hypothesis based analysis of multiple genes can lead to a greater number of gene-tumour associations on a shorter timescale than candidate gene analysis alone.

Genotype-phenotype correlation

When a variant in a CPG is detected by clinical testing, laboratory and clinical teams seek to assess that variant's pathogenicity by assessing aspects such as the likely effect on the gene product and consistency with the observed histopathological characteristics of tumours in the family. If deemed deleterious, the patient is frequently managed according to the tumour risks assigned to all deleterious variants in that gene. It is well recognised, however, that individual variants in the same gene can have contrasting clinical effects. In multiple endocrine neoplasia type 2 (MEN2), caused by activating missense mutations in the *RET* proto-oncogene, a variety of tumours are observed including medullary thyroid cancer, phaeochromocytoma and parathyroid hyperplasia/adenoma[58]. The level of risk for each of these tumours is influenced by the *RET* codon affected to the extent that mutation status is an integral part of clinical management guidelines. For example, mutations in codon 634 are associated with an increased risk of phaeochromocytoma and it is recommended to start biochemical screening for this at age eight rather than age 20 like for many other mutations[59]. They are also associated with cutaneous lichen amyloidosis, which is not reported for other variants[60]. The p.Met918Thr variant is only associated with the MEN2B clinical subtype, which includes some additional features such as gastrointestinal ganglioneuromatosis[61]. Even the finding of a non-sense or frameshift mutation may not imply that the function of the gene product is lost and will produce a risk profile similar to other patients with only one functional allele of a given gene. The BRCA2 c.9976A>T variant introduces a premature stop codon but this is towards

the 3' end of the gene and is not considered to significantly increase breast and ovarian cancer risk[62].

Section 4 – Multiple Primary Malignant Tumours

Multiple primary malignant tumours in the general population

Multiple primary malignant tumours (MPMT) describes a clinical scenario whereby two or more histologically distinct malignant tumours not due to metastasis, recurrence or local spread are diagnosed in the same individual. These may be diagnosed at the same time (synchronous) or separated by months to years (metachronous). The first description of MPMT is attributed to Billroth in 1889[63] and it initially appeared to be a rare phenomenon. However, with improved survival from many forms of cancer[64], MPMT is increasingly recognised as an important clinical problem[65]. A review of 69 European cancer registries revealed that 6.3% of registered tumours were part of an MPMT clinical picture[66] and 16% of incident cancers reported to National Cancer Institute (USA) in 2003 were in individuals with a previous diagnosis of a malignant tumour[67].

Furthermore, registry-based evidence suggests that incidence of cancer in individuals previously diagnosed with a malignant tumour is greater than the expected population incidence, with an increased risk of a wide variety of concordant (same site and morphology) and discordant (different site and morphology) tumours after an initial primary malignancy. Standardised incidence ratio (SIR) is used to measure this and expresses the ratio of observed incident cancer cases within a patient series (in this case those who had previously been diagnosed with a cancer) compared to incidence expected in a corresponding population adjusted for risk factors such as age, sex and socioeconomic status. In

a registry containing 633,964 cancer incidences, 8.5% were at least second cancers. The overall SIR of cancer in previous cancer patients was 1.3 in men and 1.6 in women, with 2.4 the highest figure for a discordant tumour in aerodigestive tract cancer patients less than one year post diagnosis. Higher SIRs may result from a number of factors discussed below. Some SIRs were lower if a cancer diagnosis had previously been made (e.g. 0.8 for prostate cancer patients). This may be due to treatment for the first tumour that reduces chances of a second occurring, especially true of further concordant cancers but also potentially applicable to discordant neoplasms (e.g. chemotherapy may also treat occult malignancies elsewhere). In addition, poor prognosis for particular cancers may lead to less active surveillance and under diagnosis of subsequent cancers compared to the general population. For example the SIR for gastric cancer in men is 0.6 after 10-38 years[68].

Aetiology of multiple primary malignant tumours

Multiple factors may contribute to the occurrence of MPMT, which may be challenging to delineate from one another.

It may occur by chance alone. A rough estimate of how frequent this is arrived at by considering the lifetime risk of developing cancer under the age of 60 years[69,70] (this study has used this age cutoff), multiplying that probability and considering it in the context of the observed age distribution in a population MPMT cohort[65]. For each age group, an expected number of cases in a given time period could be generated. Applied to the West Midlands population, the sum of these figures

suggests 16.8 cases per year, but the lifetime risk figure notably does not include non-melanoma skin cancer (NMSC). Registry based evidence from the same population (1995-1999) showed 1425 new MPMT cases per year including NMSC[64], which would be estimated to be around 200 per year without NMSC if it is considered that these cancers makes up around a 23.6% of diagnoses in the UK[71] and assumed that these had a typical age distribution and occurred evenly among first and second cancers. On this basis, observed MPMT cases appear to be more frequent than that expected by chance alone.

Increased clinical surveillance following an initial diagnosis may lead to increased detection of second malignancies through lead time bias or may identify cancers that would not present otherwise in the individual's lifetime. Identification of a tumour due to these factors may be suggested by a short intervening period between diagnoses or even a synchronous presentation. Increased diagnostic rates may be seen due to systematic imaging or examination of an organ at risk of recurrence, for example regular skin examinations following the diagnosis of cutaneous malignant melanoma[72]. They may also be seen due to surveillance imaging modalities that include other organs, as has been reported during positron emission tomography/computed tomography follow up for prostate cancer[73,74]. Similarly, surgery for one primary may expose a synchronous tumour. This can occur with the diagnosis of endometrial cancer after total abdominal hysterectomy and bilateral salpingo-oophrectomy for ovarian cancer[75], though it has been debated whether this pairing represents true independent primaries.

Radiotherapy or chemotherapy administered to treat a first cancer may predispose to second primary tumours and even non cytotoxic drug treatment may increase cancer risk as is seen for endometrial cancer after tamoxifen treatment for breast cancer[76]. Treatment related second cancers are often characterized by delayed onset and may be observed many years after the exposure to treatment occurred. Firm associations between cancer therapy and subsequent primary malignancies caused by them are challenging to arrive at for a number of reasons. Many cancers carry a poor prognosis and any carcinogenic effect of treatment is unlikely to be noted before death occurs. Additionally, treatment regimens are unlikely to be constant over time and between centres of care, preventing collation of similarly treated cases to follow up in observational studies. This difficulty is compounded in the case of tumours that occur infrequently in the population. Finally, treatment modalities for particular cancer types may become more frequently used in clinical practice with time and tumourigenic effects may be yet to manifest themselves. There is recent evidence, for example, that radiotherapy in renal cell carcinoma (a tumour that has previously been considered resistant to such treatment) may have greater utility than is widely thought[77].

Despite these challenges, a number of associations are known that demonstrate differences depending on whether chemotherapy or radiotherapy are used. Cancers may not be distinguishable from those that are not treatment related but delineation of specific features can be rewarding. Leukaemias showing high-grade microsatellite instability, for example, are common in therapy related tumours but rare where leukaemia is diagnosed without a cancer history[78].

Radiation induced second cancers generally occur ten or more years after exposure[79] and many have been documented through observational studies of survivors of events such as the Japanese atomic bomb attacks[80] and Chernobyl nuclear accident[81]. Solid cancers such as those affecting the thyroid, lung, stomach, skin and connective tissue (sarcoma) are more frequent[82] with sites reflecting tissue sensitivity and region of exposure, though haematological tumours such as leukaemias can also occur at increased rates and may occur after shorter periods of time[83]. The link between radiotherapy for Hodgkin's lymphoma and breast cancer is particularly well established and has led to revision in management strategies to reduce radiation dosage to breast tissue[84–86].

Second cancers caused by chemotherapeutic agents are more often haematological in nature and may occur after a relatively short time post treatment. Alkylating agents such as etoposide can cause acute myeloid leukaemia, usually after five to seven years and leukaemias due to epipodophyllotoxins often have a three year latency period[87]. There are also risks of solid tumours associated with chemotherapy, one example being dose responsive increased incidence of bladder cancer following cyclophosphamide treatment[88].

Carcinogenic effects of chemotherapy and radiotherapy are modified by a wide range of variables that may interact with one another to modify the risk of further

primary cancers. Higher radiation or cytotoxic agent dosage can increase risk to tissues through greater potential for mutational events but might also lead to lower risk due to induction of cell death in potentially malignant clones[78]. Tumourigenic treatment effects can also be modified by the age at which therapy is administered. A younger age at treatment results in a greater duration of time where further tumours might occur and many of the known treatment-cancer associations have been identified through follow up of children with diagnoses such as neuroblastoma[89]. Biological reasons may, in addition, account for greater risks following treatment in childhood and there is evidence that second malignancy incidence at a given follow up time point is lower in individuals where exposure occurred at a later age[79]. Theoretical reasons for this might include an increased cellular proliferation rate at an earlier age, enhancing clonal expansion of cells that have undergone tumourigenic genetic changes and increasing the probability of further cancer defining alterations occurring in their daughter cells. While chemotherapy exposes a wide range of tissues to its effects, the area covered by radiotherapy fields influences which organs are at risk (notwithstanding scattered radiation). This is seen following radiotherapy for breast cancer where an increased incidence of lung and oesophageal cancers occurs compared to patients treated without radiation[90]. Combination of modalities may complicate the picture and can itself be a risk modifying factor. For example, the administration of doxorubicin for Wilms tumour increases the risk of breast cancer following radiotherapy to treat this malignancy[91] and combined chemotherapy and radiotherapy for Hodgkin's lymphoma has been noted to confer higher incidence of gastrointestinal malignancies than would be expected if the risks from each

modality were added to one another[92]. Genetic factors may also influence likelihood of developing a second tumour due to treatment. Conditions that cause cancer susceptibility independently of treatment can enhance risks of radiotherapy or chemotherapy and manifestations of this effect include basal cell carcinomas following radiotherapy for medulloblastoma occurring in Gorlin syndrome patients[93] and radiation induced cancers in Li-Fraumeni syndrome[94]. More subtle modifying effects of genetic factors are exemplified by variation in genes encoding cytochrome p450 enzymes, effectively increasing or decreasing the administered dose of some chemotherapy agents due to altered metabolism[95].

Environmental exposures relevant to the development of two or more cancer types account for a proportion of MPMT cases and are important to consider in the assessment of such patients. These may be easily identifiable clinically such as a smoking history where, for example, an increased risk of oral and pharyngeal cancers are observed after an initial lung adenocarcinoma[96]. However, different environmental exposures in the same individual may also contribute to MPMT, particularly if those exposures are prevalent in the population. Smoking prevalence is estimated at 20% in adults in England[97] while obesity is present in ~25%[98]. A simple multiplication of probabilities would indicate that ~5% of adults have both exposures but this assumes random distribution of them in the population, which is not always the case (two carcinogenic factors that are reported to occur in association are smoking and alcohol consumption[99]).

It is widely recognised that genetic susceptibility can be a major cause of MPMT. A significant role for genetic factors in MPMT might be suggested by increased incidence of second cancers in those with a family history of the same primary tumour type (as a surrogate measure of a likely inherited genetic component). Studies based on the Swedish Family Cancer Database have shown an increased incidence of concordant and discordant second tumours in individuals previously diagnosed with breast cancer and with an affected parent. For example, the SIR (based on expected population incidence) for ovarian cancer following breast cancer was 2.0 in those with a family history of breast cancer and 1.7 in those without. The corresponding SIR's for acute lymphoid leukaemia were 12.7/1.9 and 4.6/3.0 for breast cancer[100]. Similarly, greater incidence of a second colorectal cancer has also been observed among patients who have a first degree relative with the same diagnosis, where a two-fold risk was observed compared to non-familial cases[101]. Such observations suggests an observable contribution to the burden of second cancers in the general population due to inherited genetic factors.

A proportion of multiple cancers is accounted for by cancer predisposition syndromes, which may be suggested clinically by factors such as a young age at tumour diagnosis or a family history of tumours (though not in *de novo* cases), especially if they are concordant histologically or in keeping with the phenotype associated with a particular syndrome (e.g. breast and ovarian cancer in *BRCA1*/*BRCA2* mutations and haemangioblastomas and renal cancers in VHL disease). Multiple tumours *per se* are also frequently taken as a clinical indicator.

Indeed, many predisposition syndromes are associated with a high frequency of the phenomenon[102–106]. Accordingly many patients with MPMT will be referred for evaluation by clinical cancer genetics services with a view to confirming the diagnosis of a suspected specific condition and eliciting the causative CPG mutation using genetic testing. The outcome of such evaluation has not previously been well described.

Section 5 - Study Aims

Clinical Genetics referral based series review

Although there are often large published series of individuals with a specific cancer predisposition syndrome, there are no large studies of individuals with MPMT referred for clinical genetics assessment. This information may have useful implications for clinical and laboratory practice. In particular, it is relevant to know whether individuals with MPMT who test negative for a suspected mutation/s conferring cancer predisposition are likely to represent phenocopies or whether there is evidence to indicate a need for more extensive genetic testing. To address these questions a retrospective review of referrals for MPMT to two regional genetics centres over a period of 20 years was undertaken, aiming to observe and assess:

1. Tumour types.
2. Rate of identification of pathogenic variants in CPGs.
3. Which CPGs contained pathogenic variants identified in the series.
4. The clinical evidence for undiagnosed pathogenic variants, utilising recognised clinical indicators and a newly devised scoring system.
5. Correlation between the newly devised scoring system and established scoring systems to assess likelihood of the presence of a pathogenic variant.

Candidate gene analysis

It was hypothesized that a group of patients with MPMT might harbor constitutional pathogenic variants in the CPGs *TP53*, *PTEN or CDKN2B* and analysis of these genes in a subset of individuals was consequently initiated by conventional "Sanger" sequencing.

Constitutional mutations in *TP53* cause Li Fraumeni syndrome, associated with the development of a wide variety of cancers but particularly those of the soft tissue, breast, brain and adrenal cortex[49]. Its protein product, p53 is a critical protein in a wide variety of cellular pathways and tumour suppressor functions [107,108]. Mutations in *PTEN* lead to a spectrum of disorders referred to as PTEN Hamartomatous Tumour Syndrome, the most frequent subtype of which is Cowden syndrome (CS). CS characteristically causes breast, endometrial and thyroid cancers as well as non-malignant features such as macrocephaly and developmental delay[109]. PTEN normally acts to inhibit the PI3K/Akt cellular signaling pathway and aberrant function due to *PTEN* mutation can result in loss of this inhibition and reduced apoptosis and cell cycle arrest[109]. Mutations in either gene leads to a phenotype including breast and bowel cancers, which are the most common malignancies prompting referral to clinical genetics[110]. They can also both be associated with the development of MPMT and an estimated 15% of individuals from Li-Fraumeni kindreds who are diagnosed with cancer go on to develop a second malignancy[111]. Assessments of cancer risks from the largest series of *PTEN* mutation carriers do not include an estimate of multiple primary risk *per se* although many of the cancers that such individuals are predisposed to

have individually high risks such as breast (lifetime risk 85%), colorectal (lifetime risk 34%) and endometrial (risk by age 60, 30%), suggesting an appreciable rate of second cancers[112]. Clinical testing guidelines for CPGs have been developed to target genetic analysis to those more likely to harbour a mutation based on how closely their phenotype matches that seen in the putative syndrome as previously described. The tumours associated with mutations in a particular gene however, may develop over time and as genetic testing becomes more widely available/performed. Indeed, the risk of rectal and kidney cancers and melanoma in *PTEN* mutation carriers has only recently been characterised[112]. Testing MPMT cases in this series would assess the relevance of *TP53* and *PTEN* in a less phenotypically selected group and positive results would contribute to a better definition of the associated tumour phenotype.

A non-sense mutation in *CDKN2B* has recently been identified as predisposing to cancer in a kindred where multiple members were diagnosed with renal cell carcinoma (RCC). The family also includes individuals with ovarian, colorectal and haematological malignancies. Further study of individuals with evidence of familial RCC demonstrated three cases with missense variants in this gene[113]. There are few other reports concerning the tumour spectrum that may be associated with constitutional mutations in this gene, though they have been shown to be not associated with familial melanoma[114,115]. *CDKN2B* is frequently disrupted by promoter methylation in a variety of haematological malignancies and encodes p15, a protein involved in cell cycle regulation[116]. Analysis was undertaken to assess whether *CDKN2B* mutations may be (given their recent discovery in cancer

predisposition) associated with tumours other than RCC or be present in those MPMT patients in this series who have previously been diagnosed with this tumour.

Registry series review

Patient series identified from Clinical Genetics referrals are subject to ascertainment biases that can influence their composition. MPMT patients are more likely to be referred for assessment if their phenotype is consistent with a recognised cancer predisposition syndrome where a genetic test can be offered but this may leave many cases un-assessed. Other possible reasons for non-referral include a poor prognosis after diagnosis preventing further consultations or non-genetics clinicians being unaware of the testing available. To assess the potential burden of constitutional genetic cancer predisposition in the general MPMT population, cases were ascertained from a regional registry and a clinical indicator based assessment was applied to predict how many cases might be due to a pathogenic variant in a CPG. Comparison with the Clinical Genetics based series was made to assess the extent and nature of tumour combinations that might not prompt referral but which appear amenable to more extensive genetic testing.

Chapter 2 - Materials and Methods

Section 1 - Clinical series review

Ascertainment of cases

In order to identify multiple primary malignant tumour (MPMT) cases referred for genetic assessment, a two stage evaluation of patients referred to two UK regional genetics services covering a combined population of >10 million[117,118] was undertaken. Firstly, the database of the West Midlands Regional Genetics Service was interrogated to identify individuals with two or more malignant tumours (of different types) diagnosed before the age of 60 years. Referrals and genetic analysis had taken place between February 1993 and February 2013. Medical and pathology records were then inspected to confirm the inclusion criteria. Those individuals with metastases, recurrence of the primary tumour or tumours of the same site and histological type were excluded from further analysis if those tumour characteristics led to non-fulfillment of the criteria. Multifocal cancers were counted as a single malignancy. The definition of MPMT was therefore made according to international guidelines[119]. Tumours known to be benign were excluded from analysis. In a further stage of ascertainment, two databases for individuals referred to the North West Regional Genetics service (in Manchester, UK) with a suspected diagnosis of hereditary breast or colorectal cancer were interrogated to identify additional cases of MPMT satisfying the same criteria specified above.

Case record review

To provide an indicator, in a wide range of clinical scenarios, of the strength of clinical evidence for an inherited cancer syndrome a “multiple tumour score” system was developed. To record this score for individuals within the series, Clinical Genetics case records were inspected to extract details of family history, age at diagnosis and site/histopathological type of cancer. The multiple tumour score (MTS) (Table 1) was designed to be analogous to the “Manchester score” for prioritising *BRCA1*/*BRCA2* testing in familial breast cancer kindreds[120]. Within a single family lineage, each cancer was assigned a score (Table 1) and the MTS calculated as the sum of these scores within a lineage, with one unaffected intervening relative permitted between individuals with a cancer diagnosis. A higher score was proposed to indicate a higher likelihood of an inherited cancer syndrome. Within the scoring system, rare or young onset cancers were weighted more heavily than common or later onset cancers (as phenocopies should be less likely in the former). Cervical cancer has been estimated to be almost entirely attributable to human papilloma virus (HPV) infection[22] but has an incidence peak at an earlier age than most other cancers in the series[121]. This may lead to overestimation of the likelihood of a predisposing mutation via an elevated MTS in families where these cancers appear. Nevertheless, these tumours were included in assessment given the incidence rate of cervical cancer when considered in the context of high HPV infection prevalence[122]. As with many cancers, the development of this malignancy may involve a degree of genetic predisposition acting as a co-factor among individuals with an existing environmental exposure, in this case HPV infection. Additional analyses, however, were also performed

excluding cases with cervical cancer as part of their MPMT clinical presentation. For individuals with a diagnosis of breast or ovarian cancer, the Manchester score[123] was also calculated and status for the revised Bethesda criteria for MSI analysis[51] was noted if colorectal or endometrial cancer had been diagnosed. Records of genetic testing were interrogated to obtain details of genes tested and mutations identified (including whether deemed causative by laboratory/clinical team). Statistical significance testing was performed using two tailed z-tests.

Table 1. Multiple tumour score

Malignant tumour	Age at diagnosis	Score
Breast, lung, colorectal, prostate, non-melanoma skin, cervical.	<30	5
	30-39	4
	40-49	3
	50-59	2
	>59	1
Any other malignant tumour	<50	5
	50-59	3
	>59	1

Section 2 - Candidate gene testing

DNA samples

Stored DNA samples extracted in the West Midlands Regional Genomics Laboratory were obtained from 62 patients identified by the West Midlands case note review. None of these individuals had been identified with a pathogenic variant in a cancer predisposition gene following assessment by Clinical Genetics services. All patients had given consent for genetic testing to be performed to identify a genetic cause of their cancer phenotype. Stock samples were stored at -20ºc.

Whole genome amplification

To ensure sufficient DNA quantity for analysis while maintaining a stock, samples were subject to whole genome amplification using the REPLI-g Mini Kit (Qiagen, Hilden, Germany). The expected yield of DNA from this kit is 10µg, which can be generated from DNA quantities as low as 0.1ng. According the manufacturers protocol, 5µl of each stock sample solution was initially denatured with denaturation buffer, a process which was terminated by addition of neutralization buffer. DNA (Phi 29) polymerase was then added to each reaction and incubated for 16 hours at 30ºc for amplification to proceed by multiple displacement amplification (MDA). MDA utilizes random hexamer primers to bind to sites across the genome. Phi 29 polymerase begins amplification at these sites and is able to generate large fragments (2 kilobases–100 kilobases) with high fidelity due to its 3'-5' proofreading activity[124]. Each reaction generated 50µl of amplified DNA

solution, which was diluted to 1:20 for use in polymerase chain reactions.

Polymerase chain reaction primer design

To design suitable primers to amplify coding regions of *PTEN*, *TP53* and *CDKN2B* by polymerase chain reaction (PCR), sequences were downloaded from the Ensembl genome browser[125]. Sequences corresponding to regions of interest were entered into Primer 3 software[126] to generate candidate primer sequences producing amplicons representing coding exons. Primer sequences were selected on the basis of a number of parameters. The melting temperature (Tm) of a primer describes the temperature at which half the primer duplex molecules will become single stranded due to denaturation. Sufficient single stranded molecules must be available for the annealing phase of the reaction and primer pairs were selected where there was a difference of 5ºc or less between the forward and reverse primers to ensure a similar ratio of the two species at that phase. Additionally, similar Tm's across different primer pairs were aimed for to ensure that different amplicons could be generated with a small number of reaction programmes (i.e. on the same 96 well plate). Quality of sequencing readout can be compromised by amplicon sizes of more than 500 base pairs and primer pairs were selected to ensure that generated fragments were under this size. An adequate number of bases 5' to the start of the region of interest (generally 50) were incorporated into amplicons as the first 5' bases of a sequencing readout are often of lower quality. Specificity of primer pairs was checked by inputting the oligonucleotide sequences into the Primer BLAST website hosted by the National Centre for Biotechnology Information[127] to avoid amplification of non-targeted genomic regions.

Assessments of potential for secondary primer structure formation provided by Primer 3 were also considered as a higher potential indicates lower availability of annealing primer oligonucleotides during the reaction process. Primer pairs, once selected, were ordered with desalted purification and synthesis scale of 25nmol (Sigma Aldrich, St Louis, MO, USA). Sequences are shown in Table 2. On receipt, primers were diluted to a 100µM solution with purified water. From these stock solutions, 1:10 dilutions were produced for use in PCR reactions. Both stock and working solutions were stored at -20ºc.

Table 2 – Primers used: *PTEN*, *TP53 and CDKN2B* coding regions

Amplicon	Forward Primer	Reverse Primer
PTEN exon 1	AGAGCCATTTCCATCCTGCAGA	ACGTTCTAAGAGAGTGACAGAA AGGT
PTEN exon 2	TCTTTCTTTTCATAACTA	ATCTTTTTCTGTGGCTTAGAAA TCTTTTC
PTEN exon 3	CCATAGAAGGGGTATTTGTTGG	ACTCTACCTCACTCTAACAAGCA
PTEN exon 4	ATTCAGGCAATGTTTGTTAGTAT	TACAGTCTATCGGGTTTAAGTT ATACAA
PTEN exon 5	GGGGAAAATAATACCTGGCTTCC	AAATTCTCAGATCCAGGAAGAGG
PTEN exon 6	ACCCAGTTACCATAGCAATTTA	AGAAAACTGTTCCAATACATG GAAGGAT
PTEN exon 7	TGACAGTTAAAGGCATTTCCTG	TAGCTTTTAATCTGTCCTTAT TTTGGATATTT
PTEN exon 8	GCAACAGATAACTCAGATTGCC	CATACATACAAGTCAACAACCCC
PTEN exon 9	GAGGGTCATTTAAAAGGCCTCT	TCATGGTGTTTTATCCCTCTTGA
TP53 exon 2	CACTGGCATGGTGTTGGG	TTTTCGCTTCCCACAGGTCT
TP53 exon 3	TCTCATGCTGGATCCCCACT	AGTCAGAGGACCAGGTCCTC
TP53 exon 4	CCATGGGACTGACTTTCTGC	GATACGGCCAGGCATTGAAG
TP53 exon 5	ACGCCAACTCTCTCTAGCTC	TCAGTGAGGAATCAGAGGCC
TP53 exon 6	TGTTCACTTGTGCCCTGACT	TTAACCCCTCCTCCCAGAGA
TP53 exon 7	CGACAGAGCGAGATTCCATC	GGGTCAGAGGCAAGCAGA
TP53 exon 8/9	TTGGGAGTAGATGGAGCCT	AGTGTTAGACTGGAAACTTT
TP53 exon 10	CAATTGTAACTTGAACCATC	GGATGAGAATGGAATCCTAT
TP53 exon 11	AGACCCTCTCACTCATGTGA	TGACGCACACCTATTGCAAG
CDKN2B exon 1	AAGAGTGTCGTTAAGTTTACG	ACATCGGCGATCTAGGTTCCA
CDKN2B exon 2	TGAGTATAACCTGAAGGTGG	GGGTGGGAAATTGGGTAAG

Polymerase chain reaction

Whole genome amplified DNA samples were incorporated into PCR reactions on a 96 well plate according to the reaction constitution outlined in Table 3. The Biomix red solution (Bioline, London, UK) contains Taq DNA polymerase, deoxyribonucleotide triphosphate molecules (dNTPs), reaction buffers and a dye to assist with transfer of reaction products. GC rich solution was incorporated if GC content of the amplicon was high and the initial reaction had failed. Control reactions without template DNA were included in all sets of reactions where a particular primer pair was used.

Table 3- Polymerase chain reaction constituents

Reaction component	Quantity per reaction
DNA solution (whole genome amplified)	4-5 µl
Water	6.5 µl
Forward primer	0.5 µl
Reverse primer	0.5 µl
Biomix red solution	12.5 µl

Reactions were run on a Bio-Rad Tetrad Thermal Cycler (Bio-Rad, Hercules, CA, USA) according to the programmes described in Tables 4 and 5. For *PTEN* and *TP53* amplicons, annealing temperature was variable according to primer pair. This was based on the Tm from each primer (2-5°c below the lowest in each primer pair) but also through optimisation of individual amplicon reactions. A high annealing temperature was aimed for as non-specific annealing of primers is less likely at higher temperatures. This is offset against a generally higher yield of PCR product at lower temperatures due to greater annealing of primers. Tm for

CDKN2B reactions was constant as there were only two amplicons. For these amplicons, annealing temperature was increased over five denaturation-annealing-polymerisation cycles in the early part of the reaction. This was intended to optimise specificity of PCR products (at the expense of yield) in the earliest phase of the reaction. A population of specific products allows a lower annealing temperature to be used in later cycles, optimising yield while minimising amplification of the now less abundant non-targeted genomic areas.

Table 4 – Thermal cycling programmes used for polymerase chain reactions: *PTEN* and *TP53*

Reaction step	Temperature	Duration	Cycles
1	95°c	10 mins	
2	95°c	60 secs	
3	Variable dependent on primer pair (55-64°c)	60 secs	
4	72°c	90 secs	Return to step 2 x39
5	72°c	10 mins	

Table 5 – Thermal cycling programmes used for polymerase chain reactions: *CDKN2B*

Reaction step	Temperature	Duration	Cycles
1	95°c	5 mins	
2	95°c	45 secs	
3	58-62°c	45 secs	
4	75°c	45 secs	Return to step 2 x4. Increase step 3 temp by 1°c each cycle
5	95°c	45 secs	
6	54°c	45 secs	
7	72°c	45 secs	Return to step 5 x36
8	72°c	5 mins	

To observe the PCR reaction's specificity and yield, agarose gels were produced by adding ethidium bromide (1µl per 100ml) to 1% agarose solution (agarose powder in Tris Acetate Ethylenediaminetetraacetic acid (EDTA) buffer). A sample of each PCR reaction was added to wells incorporated in the gel, which was bathed in Tris Acetate EDTA buffer. Following a 20 minute period run at 180v, the gel was removed and photographed under ultraviolet light to visualise DNA. A single clear band on the gel denoted a specific product with good yield suitable for sequencing. PCR products were not used for sequencing if there was evidence of non-specificity of the reaction, seen as blurred bands on the gel indicating amplicons of various sizes (i.e. not at the same genomic location) being produced. If reaction products were observed from control wells then reactions corresponding to that amplicon pair were rejected as this indicated contamination.

Sequencing reaction

2.4µl PCR reaction products were cleaned by adding an equal volume of microCLEAN (Microzone, Haywards Heath, UK) and centrifuging at 4000RPM for 40 minutes. The supernatant was then discarded by inverting the 96 well plate containing the products onto a paper towel and centrifuging at 500RPM for 30 seconds. The remaining DNA pellet was subsequently incorporated into a sequencing reaction comprising of (per reaction) 0.5µl BigDye® Terminator v3.1 Ready Reaction Mix (containing labelled chain terminating dideoxyribonucleotide triphosphate molecules (ddNTPs), deoxyribonucleotide triphosphate molecules (dNTPs) and DNA polymerase) and 2µl 5X sequencing buffer from the BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Austin TX, USA). Also

added were 2ml of 2mM primer solution (either forward or reverse primer) and 5.5µl water. The reactions were run on a Bio-Rad Tetrad Thermal Cycler (Bio-Rad, Hercules, CA, USA) with a programme described in table 6.

Table 6 – Sequencing reaction thermal cycling programme

Reaction step	Temperature	Duration	Cycles
1	95°c	1 min	
2	95°c	30 secs	
3	60°c	30 secs	
4	60°c	4 min	Return to step 1 x 34
5	12°c	10 min	

Sequencing reaction products were precipitated using a two stage ethanol based protocol. In the first stage, 2 µl 0.125M EDTA and 30 µl 100% ethanol was added to each well. The plate was centrifuged at 2000RPM for 20 mins and the supernatant removed by inverted centrifuging (on a paper towel) at 450RPM for 30 seconds. In the second stage, 90 µl 70% ethanol was added to each well and the plate centrifuged at 2000RPM for 10 mins. Supernatant was removed in the same fashion as in the first stage. 10 µl HiDi Formamide (Life Technologies, Carlsbad, CA, USA) was added to the remaining pellet in each well, which was subjected to a denaturation step by exposure to 95°c for two minutes then snap chilled on ice. Plates were subsequently loaded onto an ABI 3730 automated sequencer (Applied Biosystems, Norwalk, CT, USA).

Sequencing data analysis

Obtained chromatograms were analysed for variants visually and with the

assistance of the Mutation Surveyor software package (Softgenetics, State College PA, USA) using reference sequences generated from the GRCh37 human genome assembly[128].

Section 3 - Registry review

Ascertainment of cases

A series of individuals with more than one diagnosis of a malignant tumour was ascertained from the National Cancer Registration Service – Eastern Office, which covered a five-year period from 2009 to 2014. Only cases who had developed two cancers before the age of 60 years or three cancers before the age of 70 years were considered. All cancer sites were included but two or more malignancies in the same organ pair only counted as one tumour. Cases were reviewed and excluded if they did not fulfil the International Agency for Research on Cancer criteria for the classification of multiple primary cancers[119]. These criteria stipulate that cancers occurring at the same site may be classified as separate primaries if they fall into different histological groupings. For this reason, haematological tumours were subdivided into groups according to their cell lineages as described in the guidelines. Basal cell and squamous cell carcinomas of the skin counted as separate cancers but were also considered together as non-melanoma skin cancer (NMSC).

Collation and grouping of tumour combinations

Each combination of tumours that occurred in an individual patient was recorded along with the five year age band at which that combination had occurred (i.e. the age at which the later tumour of that particular combination had been recorded). Every combination of two tumours was considered. Thus, if an individual had been diagnosed with tumour A and B, then this was recorded as combination A-B. If

diagnoses had been made of tumours A, B and C, then this was recorded as 3 different combinations (A-B, A-C and B-C). To assess which combinations of cancers were more likely to be due to genetic predisposition, criteria were applied to the series based on population incidence and the relative known contribution of environmental exposure. To identify common tumours, a list of the five most frequent tumours in the UK population was used, namely breast, lung, colorectal and prostate cancer and NMSC[129]. To consider tumours less likely to be due to environmental exposure, cancers with a population attributable fraction (PAF) of greater the 50% were recorded. PAF is an estimate of the proportion of incident cancers due to a recognised environmental exposure and is based on a large scale analysis of the relative contribution of 14 such exposures undertaken in a UK population[22].

Chapter 3 - Results

Section 1 – Clinical Genetics series review

Evaluation and analysis of West Midlands MPMT series

212 individuals (with 441 tumours) referred for genetic evaluation of a multiple primary malignant tumour (MPMT) phenotype satisfied the eligibility criteria for inclusion in the study. Most individuals (179, 84.4%) were female and breast cancer was the most frequent tumour type (Table 7). The most frequent combination of tumour types was breast with ovarian followed by breast with non-melanoma skin cancer, endometrial and colorectal cancers respectively (Table 8).

Table 7 – West Midlands MPMT series subgroups

5 most common tumour types	All tumours (n=441)	Testing sent (n=234)	No testing sent (n=208)	Test+ve (n=96)	Test-ve (n=138)
Breast	128 (29%)	69 (29.5%)	59 (28.3%)	21 (21.9%)	48 (34.8%)
Ovarian	54 (12.2%)	35 (14.9%)	18 (8.6%)	14 (14.6%)	21 (15.2%)
Colorectal	71 (16.1%)	34 (14.5%)	37 (17.8%)	21 (21.9%)	13 (9.4%)
Endometrial	43 (9.7%)	28 (12%)	16 (7.7%)	12 (12.5%)	16 (11.6%)
NMSC	42 (9.5%)	22 (9.4%)	21 (10.1%)	7 (7.3%)	15 (10.9%)
Other tumours comprising < 5% total	103 (23.3%)	46 (19.6%)	57 (27.4%)	21 (21.9%)	25 (18.1%)

NMSC – Non melanoma skin cancer

111 of 212 probands (52.3%) had been tested for constitutional mutations in at least one gene relevant to inherited cancer (*BRCA1* (n=70 individuals), *BRCA2*

(n=67), *MSH2* (n=31), *MLH1* (n=30), *MSH6* (n=11), *PMS2* (n=2), *APC* (n=2), *MUTYH* (n=4), *PTEN* (n=4), *TP53* (n=3), *RB1* (n=2). Comparison of tumour types between the 111 patients tested for mutations in one or more genes and the 101 individuals who were not tested revealed broadly similar tumour frequencies with only thyroid cancer (higher, p=<0.05) and ovarian cancer (lower, p=<0.05) being significantly different in the untested group. The proportionally fewer colorectal cancers (Table 7) in the tested group is likely due to 25/101 cases who did not undergo germline genetic testing following negative microsatellite instability (MSI) and/or immunohistochemistry (IHC) testing indicating insufficient evidence of a constitutional mismatch repair gene mutation/Lynch syndrome to justify further analysis. There were no differences in gender distribution between tested and untested groups (94/111 (84.7%) female in tested group and 86/101 (85.1%) female in untested group) but the mean age at diagnosis of first (41.5 vs 43.2), second/synchronous (48.8 vs 50.7) and all (45.2 vs 47.3) tumours was lower in those who had undergone testing. Where a particular tumour occurred and a genetic test was performed, a mutation was identified in the patient diagnosed with that tumour in between 27.3% (non-melanoma skin) and 61.8% (colorectal) of cases (Table 9). The relatively high mutation rates for colorectal and endometrial tumours are again likely due to Lynch syndrome investigations instigated prior to germline testing.

Table 8 – Tumour combinations observed. West Midlands MPMT series

Quantities indicate number of combinations between x axis tumour and y axis tumour abbreviations on x axis correspond to tumour sites on y axis. NMSC – Non melanoma system, CRC – Colorectal cancer.

	Aerodig.	Anal	Bladder	Breast	Cervix	CNS	CRC	Endom.	Gastric	Haem.	Kidney	Lung	Melan.	NMSC	Oesoph.	Ovary	Pancreas	Periton.	Prostate	Retino.	Sarcoma	Sweat gl.	Small bo.	Thyroid	Ureter
Aerodigestive tract																									
Anal																									
Bladder																									
Breast	2	2	2																						
Cervix				7																					
CNS				2																					
CRC	1		9	16	2	1																			
Endometrial			1	19	1		10																		
Gastric				1			1	1																	
Haematological			1	7			1																		
Kidney				3			1																		
Lung			1	1			1	1																	
Melanoma				13				2																	
NMSC	1		4	20	1		13		1				1												
Oesophagus							1																		
Ovary			2	28	1		7	13			2	2	1	2											
Pancreas							1																		
Peritoneal				1																					
Prostate							2		1																
Retinoblastoma				1			1					1													
Sarcoma				4						1										1					
Sweat gland				1																	1				
Small bowel				1																					
Thyroid				4			1	1					1	1											
Ureter							3							1											

Table 9 – West Midlands MPMT testing sent group – Positive mutation testing results by tumour type

Tumour type	Number of occurrences where patient was mutation positive
Breast	21 (30.4%)
Ovarian	14 (40%)
Colorectal	21 (61.8%)
Endometrial	11 (39.3%)
Non melanoma skin	6 (27.3%)
Other tumours individually comprising < 5% total	22 (47.8%)

Overall, 44 of the 111 (39.6%) patients who underwent germline genetic testing in one or more genes had a pathogenic variant demonstrated (designated Test+ve group) and 67 did not (designated Test-ve group). 25 cases had negative MSI and/or IHC analysis for suspected Lynch syndrome and did not proceed to germline testing. These cases were not assigned Test-ve status due to the fact that the sensitivity of these investigations is quoted as between 55% and 83% depending on gene involved and the assay. Similarly, four cases who had positive MSI/IHC but didn't have a germline test (due to death in three cases) were not assigned Test+ve status due to high but incomplete specificity[130]. These individuals are nevertheless likely to have a mismatch repair deficit and further calculations were performed (see below) where these investigations were taken as negative or positive genetic tests respectively.

Pathogenic variants were most frequent in mismatch repair genes (n=21, *MSH2*

(n=9), *MLH1* (n=9), *MSH6* (n=3)) or *BRCA1*/*BRCA2* (n=16, *BRCA1* (n=12) and *BRCA2* (n=4)). Seven patients harboured pathogenic variants in genes relevant to rarer syndromes (*PTEN* ((n=3), *RB1* (n=2), *APC* (n=1), *MUTYH* (homozygous) (n=1)). Comparison of the Test+ve and Test-ve groups revealed that the mean multiple tumour score (MTS) was significantly higher (21 vs 14.1, p= <0.01) and the mean age at tumour diagnosis lower, though not statistically significant, (44 vs 46, p= 0.3271) (Table 10) in the Test+ve group. The proportion of cases with a concordant tumour in a first degree relative was also higher in the Test+ve group (63.6% vs 56.7% p= 0.4654) though again, this was not statistically significant.

Table 10 – Clinical parameters observed in West Midlands MPMT series

	All cases (n = 212)	**Mutation identified (Test+ve) (n = 44)**	**No mutation identified. ≥ 1 germ-line test (Test-ve) (n = 67)**	**No mutation identified (n = 168)**
Mean age at tumour diagnosis	46.2 SD 10.2	44 SD 10.9	46 (SD 9.8 p= 0.3271)	46.8 (SD 9.9 1.85 p= 0.064)
Mean age diagnosis 1st tumour	42.5 SD 11.3	39.4 SD 12.3	42.9 (SD 10 p= 0.1141)	43.1 (SD 10.7 p= 0.368)
Mean age diagnosis 2nd or synchronous tumour	49.6 SD 7.4	48.1 SD 6.8	49.2 (SD 8.4 p= 0.4473)	50.1 (SD 7.2 p= 0.089)
Mean individual age at tumour diagnosis	46.2 SD 8.4	44.2 SD 8.6	46 (SD 8.6 p= 0.2801)	46.7 (SD 8.3 p= 0.082)
% cases with individual mean age at tumour diagnosis ≤ Test+ve group median (= 46.25)	N/A	N/A	41.8 (28)	39.9 (67)
% cases with concordant tumour in 1st degree relative	52.3 (111)	63.6 (28)	56.7 (38) p= 0.4654	49.4 (83) p= 0.093
Mean multiple tumour score	14.6 SD 8.7	21 SD 14.3	14.1 (SD 5.4 p <0.01)	12.9 (SD 5.6 p <0.01)
% cases with multiple tumour score ≥ Test+ve group median (=17)	N/A	N/A	28.3 (19)	21.4 (36)

Mean individual age at diagnosis = Combined age of tumours ÷ number of tumours in individual. P values describe two tailed comparison with Test+ve group ($H_0 = \mu_1 = \mu_2$).

There was, however, significant overlap between the MTS in the Test+ve and

Test-ve groups such that ~ 28% of Test-ve patients had a value equal to or higher than the median seen in the Test+ve group (Fig 2) and 88% (59 cases) had a score equal to or higher than the lowest value in the Test+ve group. 21.4% of all cases without a mutation demonstrated (including those who had no genetic testing performed) had a score equal to or higher than the Test+ve group median. The individual mean age at tumour diagnosis (combined ages at diagnosis ÷ number of primary tumours) was also calculated and 41.8% of Test-ve cases had an individual mean diagnosis age at or below the median value observed in the Test+ve group (Fig 3). This figure was 39.9% if mutation status unknown individuals that had not undergone mutation analysis were also included in analysis. If the results of MSI/IHC analysis were given equal status to that of germline genetic testing then similar figures were observed. 28.1% of the revised Test-ve group had an MTS at or above the revised Test +ve median and 38.6% had an individual mean age of diagnosis equal to or below the revised Test +ve median. 55% had a first degree relative with a concordant tumour compared to 62.5% in the revised Test+ve group.

Figure 2 - Multiple tumour scores in Test-ve group

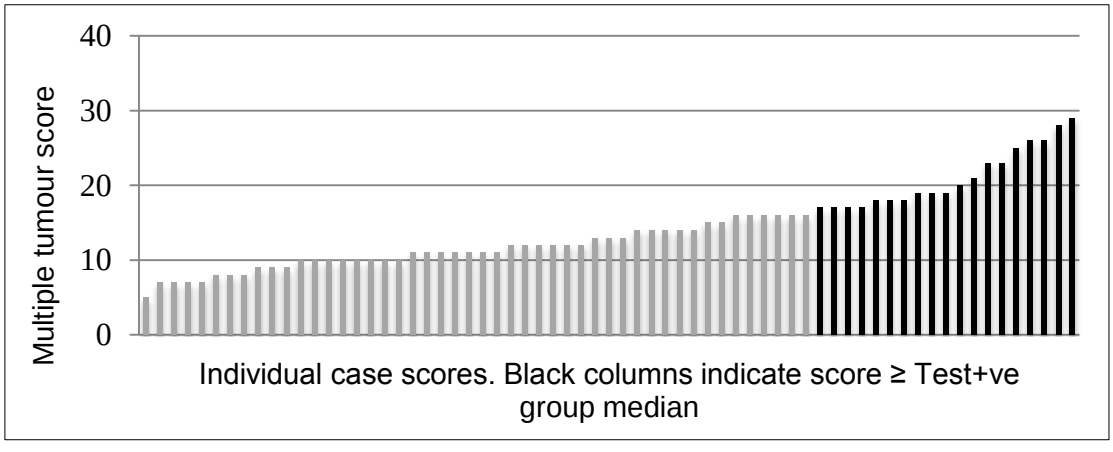
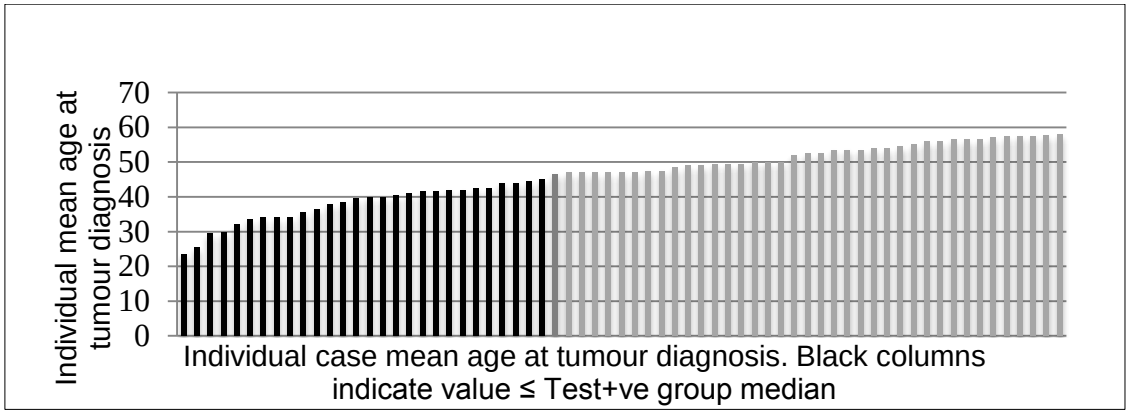


Figure 3 - Individual mean age at tumour diagnosis in Test-ve group



As previously mentioned, analysis was also performed without the cervical cancers present in the patients within the series. This led to the exclusion of nine individuals, five of whom had received genetic testing with a mutation identified in four (BRCA1, BRCA2 x2 and APC). One tumour was discounted that was diagnosed in a patient with two other primaries. Results are shown in Appendix 1 and 2 and correspond to the previous Tables 7 and 10. No significant differences in results were shown between the two analyses both in terms of tumour profile and clinical indicators of a causative mutation.

Comparison of Multiple Tumour Score with existing scoring systems

To investigate further the outcome of genetic evaluation in MPMT referrals, the characteristics of 240 further MPMT cases referred to the North West Regional Genetics Service were analysed. 230 (95.8%) cases had at least one diagnosis of breast, ovarian, colorectal or endometrial cancer. 166 (69.1%) cases had received a diagnosis of breast and/or ovarian cancer and 144 (60%) had previously had

colorectal and/or endometrial malignancy. Details of the individual tumour combinations are shown in Table 11. Breast with ovarian cancer was again the most frequent combination followed by colorectal with endometrial, colorectal with breast and endometrial with ovarian malignancy.

The data from these North West cases were combined with that from the West Midlands investigation patients who had breast and/or ovarian cancer (combined total n=320) or colorectal and/or endometrial cancer (combined total n=244). The relationships between MTS and Manchester score (a measure of indication for *BRCA1*/*BRCA2* testing) and MTS and fulfillment of revised Bethesda criteria in the two cancer phenotype subgroups were then investigated and are shown in Figure 4 and Figure 5 respectively. Manchester score showed a positive correlation with MTS. Among the colorectal/endometrial cases, 58 (23.8%) had an MTS higher than 20. Four out of the 59 (6.8%) individuals not fulfilling the Revised Bethesda criteria had scores above this level, therefore the proportion of patients with the highest MTS were less likely to fall short of criteria fulfillment.

Figure 4 – Correlation of multiple tumour scores and Manchester scores in Clinical Genetics combined series

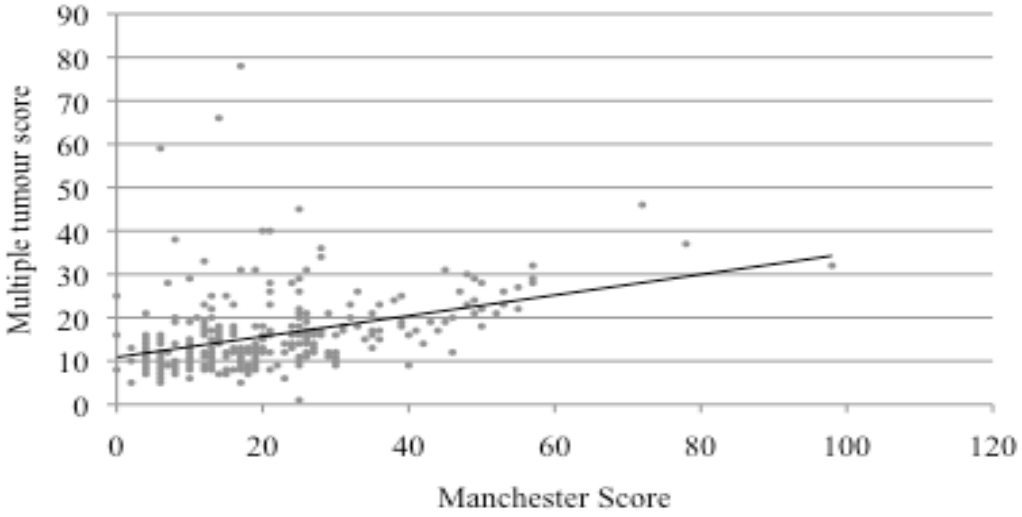


Figure 5 – Multiple tumour scores and Revised Bethesda criteria in patients with colorectal and/or endometrial cancers in Clinical Genetics combined series

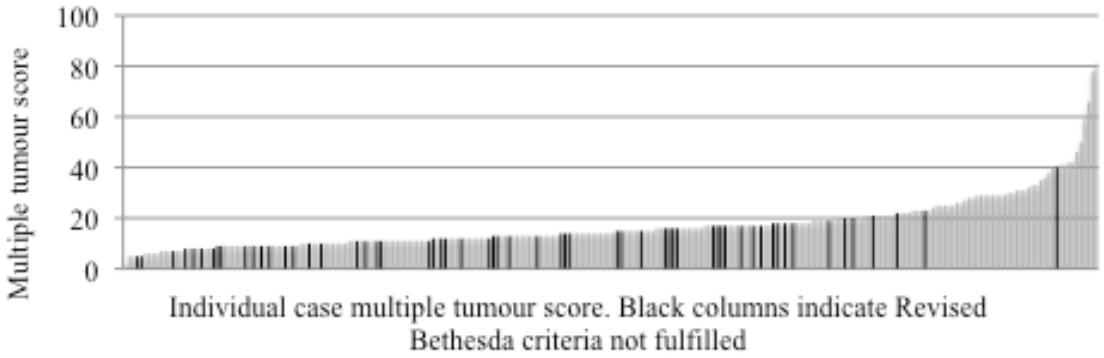


Table 11 – Tumours combinations observed. North West Regional Genetics Service series

Quantities indicate number of combinations between x axis tumour and y axis tumour observed in series. Tumour site abbreviations on x axis correspond to tumour sites on y axis. NMSC – Non melanoma skin cancer, CNS – Central nervous system, CRC – Colorectal cancer.

	Adrenal	Aerodig.	Biliary	Bladder	Breast	Cervix	CNS	CRC	Endom.	Gastric	Haem.	Kidney	Lung	Melan.	NMSC	Oesoph.	Ovary	Pancr.	Periton.	Prostate	Retinobl.	Sarcoma	Small b.	Testis	Thyro.	Ureter
Adrenal																										
Aerodigestive tract																										
Biliary																										
Bladder																										
Breast	1	1		1																						
Cervix					1																					
CNS					1																					
CRC			2	2	26	5																				
Endometrial				1	12	2	1	37																		
Gastric					1			3																		
Haematological								7	3																	
Kidney					4			5	1																	
Lung			1		2																					
Melanoma					4			1																		
NMSC					2		2	10	4																	
Oesophagus			1		2			1																		
Ovary			1	2	67			12	21					1	1	1										
Pancreas								2				1														
Peritoneal									1																	
Prostate								1																		
Retinoblastoma								1																		
Sarcoma					6		2					1					1									
Small bowel								8	1																	
Testis								2																		
Thyroid					1			2																		
Ureter				1	1			4	1								1						1			

Section 2 – Candidate gene testing

The overlap between Test+ve and Test-ve groups in MTS distributions and age at diagnosis suggested that some of the Test-ve group might harbour a mutation in a cancer predisposition gene (CPG) that had not been analysed as part of clinical service provision.

To test this hypothesis, sequencing of the coding regions of *PTEN* and *TP53* was performed in 62 patients not previously tested for mutations in these genes. The details of the tumour phenotypes of the 62 patients are shown in Table 12. 56 of the 62 (90.3%) patients had been diagnosed with a tumour previously associated with mutation in *PTEN* and/or *TP53*[49,109] (breast cancer in 42 patients, endometrial carcinoma (n=10) papillary thyroid cancer (n= 2), sarcoma (n=1) and gliomas (n=1)). 41 of the 62 (66.1%) had previously undergone testing on their *BRCA1*/*BRCA2* and/or mismatch repair genes. No germline test had been performed in 21 cases, eight of whom had negative MSI and/or IHC analysis. The 13 remaining untested patients all had been diagnosed with breast cancer in combination with another tumour but there were no colorectal cancers among this group.

No pathogenic or candidate mutations were detected in *PTEN* or *TP53*. The *TP53* Pro72Arg common variant was observed in the samples although the genotype distribution (C;C ~10%, C;G ~30%, G;G ~60%) among the series was similar to that expected in a European population[131] and no phenotypic differences were

observed between the genotype groups separated on the basis of this variant.

Of the 62 cases tested for *PTEN* and *TP53* variants, 53 were also tested for *CDKN2B*, with those not undergoing analysis indicated by an asterisk in Table 12. A heterozygous c.20G>A p.Gly7Asp (rs150973276) variant was detected in one case. This individual had been diagnosed with breast and cervical cancers at the ages of 52 and 53 respectively. Family history included a son who died from choricarcinoma at age 46 and two sisters who had breast cancer at 72 and 73. Her father had lung cancer at 50 and paternal uncle died from an unknown primary cancer aged between 30-35 (unpublished data).

This variant is rare (minor allele frequency of 0.0077% on Exome Variant server[132]), conserved (GERP score 2.92 (range -12.3 to 6.17)[133]) and predicted to be possibly damaging based on the in-silico predicted effect of amino acid substitution (Polyphen score 0.933[134]). However, it has not previously been reported clinically and does not appear on the mutation databases ClinVar[135], Human Gene Mutation Database[136] or the Leiden Open Variant Database[137].

Table 12 (i) - Tumour phenotypes in cases undergoing mutation analysis of PTEN/TP53/CDNK2B (*not tested for CDKN2B mutations)

Case	Tumour type – 1st	Age	Tumour type – 2nd or	Age	Further tumours
1	Ovarian carcinoma	22	Breast carcinoma	42	
2	Hodgkin's lymphoma	23	Bone sarcoma	27	Breast carcinoma 39
3	Cervical squamous cell	28	Breast carcinoma	49	
4*	Breast carcinoma	29	Thyroid papillary	38	
5	Anal squamous cell	29	Breast carcinoma	30	
6	Melanoma	31	Breast carcinoma	40	
7*	NMSC (BCC)	31	Endometrial carcinoma	42	Ovarian carcinoma 42
8	Cervical squamous cell	31	Breast carcinoma	51	
9	Colorectal carcinoma	31	Kidney renal cell	44	
10	Ovarian carcinoma	33	Endometrial	35	
11	Thyroid papillary carcinoma	35	Breast carcinoma	54	
12	Breast carcinoma	36	NMSC (SCC)	47	
13	NMSC (basal cell	36	Breast carcinoma	37	
14	Anal squamous cell	37	Breast carcinoma	42	
15	Polycythaemia rubra vera	37	Breast carcinoma	39	
16	Melanoma	38	Breast carcinoma	47	
17	Breast carcinoma	38	NMSC (BCC)	56	
18*	Breast carcinoma	39	Bladder TCC	42	
19	NMSC (BCC)	40	Ovarian carcinoma	55	
20	Breast carcinoma	40	Endometrial carcinoma	54	
21	Ovarian carcinoma	40	Colorectal carcinoma	40	
22	Breast carcinoma	41	NMSC (BCC)	55	
23	Melanoma	41	Breast carcinoma	47	
24	Breast carcinoma	42	Ovarian carcinoma	47	
25	Colorectal carcinoma	43	Endometrial carcinoma	51	
26	Breast carcinoma	43	Endometrial	47	
27*	Breast carcinoma	43	Ovarian carcinoma	54	
28	Bladder transitional cell	43	NMSC (BCC)	51	
29	Colorectal carcinoma	44	NMSC (BCC)	24	
30*	Melanoma	44	Breast carcinoma	50	
31	Colorectal carcinoma	44	Breast carcinoma	50	

NMSC – Non-melanoma skin cancer. BCC – Basal cell carcinoma. SCC – Squamous cell carcinoma. TCC – Transitional cell carcinoma

Table 12 (ii) - Tumour phenotypes in cases undergoing mutation analysis of PTEN/TP53/CDNK2B (*not tested for CDKN2B mutations)

Case	Tumour type – 1st	Age	Tumour type – 2nd	Age	Further tumours
32	Breast carcinoma	46	Endometrial carcinoma	59	
33	Colorectal carcinoma	46	Breast carcinoma	54	
34	Melanoma	46	Breast carcinoma	47	
35	Colorectal carcinoma	46	Endometrial carcinoma	48	
36	Non melanoma skin cancer	46	Colorectal carcinoma	55	
37	Breast carcinoma	47	Renal cell carcinoma	51	
38	Colorectal carcinoma	47	Prostate	56	Gastric adenocarcinoma 57
39*	Brain, oligodendroglioma	48	Breast carcinoma	51	
40	Cerebellar tonsil	48	Colorectal carcinoma	52	
41†	Breast carcinoma	48	Ovarian carcinoma	53	Endometrial carcinoma 53
42	Thyroid follicular	48	Breast carcinoma	57	
43*	Breast carcinoma	48	NMSC (BCC)	50	
44*	Breast carcinoma	48	Endometrial carcinoma	52	
45	Bladder TCC	48	Colorectal carcinoma	52	
46	Ovarian carcinoma	49	Breast carcinoma	50	
47	Breast carcinoma	49	Non-Hodgkin lymphoma	59	
48	Breast carcinoma	50	Melanoma	52	
49	Colorectal carcinoma	51	Breast carcinoma	53	
50	Breast carcinoma	52	Thyroid papillary	59	
51	NMSC (BCC)	53	Breast carcinoma	54	
52	Ovarian carcinoma	54	Breast carcinoma	59	
53	Pancreatic adenocarcinoma	54	Colorectal carcinoma	59	
54	NMSC (multiple BCC)	54	Breast carcinoma	54	
55	Colorectal carcinoma	54	Endometrial carcinoma	54	
56	Kidney renal cell carcinoma	55	Gastric adenocarcinoma	59	
57*	NMSC (BCC)	56	Breast carcinoma	59	
58	Thrombocythaemia	56	Breast carcinoma	56	
59	Lung carcinoma	56	Ovarian carcinoma	57	
60	Lung carcinoma	56	Ovarian carcinoma	57	
61	Ovarian carcinoma	57	Endometrial	57	
62	Lung carcinoma	57	Ovarian carcinoma	59	Bladder carcinoma 57

NMSC – Non-melanoma skin cancer. BCC – Basal cell carcinoma. SCC – Squamous cell carcinoma. TCC – Transitional cell carcinoma
† Also cervical adenocarcinoma aged 53

Section 3 – Cancer registry Review

In order to assess the frequency/nature of MPMT occurring in the UK population and compare it with a Clinical Genetics based series, information was obtained from the National Cancer Registration Service – Eastern Office, which covers a population of 5.5. million[138]. The database was interrogated to identify individuals who had been diagnosed with two or more cancers before the age of 60 years, or three or more cancers before 70 where the most recent cancer diagnosis had been made during a five year period since January 2009. In total, 530 cases were observed representing 1191 tumours. Because a constitutional genetic cancer aetiology is suggested by an earlier age of diagnosis, individuals were separated into two groups on the basis of whether their second cancer had been diagnosed at under 50 years or after this age. The older age group contained 395 individuals. 901 tumours were observed, which could be categorised into 38 subgroups on the basis of site and histology. The younger age group contained 135 individuals. 290 tumours were observed, which could be categorised into 33 subgroups. Subgroups and their frequency are listed in Appendix 3 and the most common are listed in Table 13.

Table 13 – Most frequent tumours in cancer registry series

Cancer type	Frequency	Cancer type	Frequency
NMSC	24.6% (n=222)	NMSC	20.7% (n=60)
Breast	14.4% (n=130)	Breast	18.0% (n=52)
Malignant melanoma	8.1% (n=73)	Malignant melanoma	13.1% (n=38)
Prostate	8.0% (n=72)	Haematological B-cell	5.9% (n=17)
Colorectal	6.3% (n=57)	Thyroid	3.4% (n=10)

NMSC – Non-melanoma skin cancer

Skin cancers were collectively the most common group followed by the common cancers of breast, prostate and colorectal. The three most common cancers were the same in both age groups but haematological tumours of B cell lineage and thyroid cancers were more frequent in the younger patients.

627 cancer combination occurrences (some individuals had more than one combination if diagnosed with more than two malignancies) were observed in the after age 50 group, which represented 153 combination types. The before 50 group contained 178 combination occurrences which fell into 85 combination types. Comparison of tumour combinations between the groups is illustrated in table 14. The most frequent combinations involve similar cancers though prostate and colorectal tumours were more frequent in the older age group.

Table 14 – National Cancer Registration Service – Eastern Office series tumour combinations representing >2% total in under 50 and over 50 cases

≥2 tumours under 50 (135 cases)	% of total combinations (n=178)	≥2 tumours not until after 50 (395 cases)	% of total combinations (n=627)
Melanoma - NMSC	13	Breast - NMSC	9.2
Breast - NMSC	12.4	Melanoma - NMSC	8.9
Breast - Melanoma	6.2	NMSC - Prostate	5.7
Breast - Endometrial	3.4	Haem B cell - NMSC	5
Haem B cell - NMSC	3.4	Breast - CRC	3.2
		Breast - Endometrial	3
		CRC - NMSC	2.7
		Breast - Melanoma	2.4
		Haem B cell - Prostate	2.2
		Aerodigestive - NMSC	2.1

NMSC – Non-melanoma skin cancer

To assess, from the available information, the evidence for genetic cancer predisposition in the registry series, each tumour was assigned a status based on UK population frequency and population attributable fraction (PAF). A summary of these parameters applied to the series is shown in Table 15. In the younger age group, there was a significantly higher frequency of combinations where both tumours were not common and where both tumours were not common and had a low PAF.

Table 15 - Comparison of National Cancer Registration Service – Eastern Office series tumour combinations by epidemiological characteristics of tumours under 50 vs over 50

Tumour combination characteristics	≥2 tumours under 50 (178 combinations)	≥2 tumours not until after 50 (627 combinations)	p value
1 tumour not common	142 (79.8%)	464 (74.0%)	0.11
Both tumours not common	50 (28.1%)	99 (15.8%)	0.0002
1 tumour low PAF	147 (82.6%)	499 (71.6%)	0.37
Both tumours low PAF	63 (35.4%)	223 (35.6%)	0.97
1 tumour low PAF and not common	99 (55.6%)	304 (48.5%)	0.93
Both tumours low PAF and not common	32 (18.0%)	59 (9.4%)	0.00142

PAF – Population attributable fraction
P values describe two tailed comparison between groups ($H_0 = \mu_1 = \mu_2$)

The West Midlands Clinical Genetics MPMT series (WM series) was compared to the population based series to observe differences in terms of tumours occurring. The WM series consists of individuals diagnosed with two separate cancers before the age of 60 so only those from the registry series fitting these criteria were considered.

The WM series, corresponding to a twenty year referral period, contained 212 cases and 441 tumours whereas the revised registry series contained 472 cases and 1013 tumours that occurred over a five year period. The most frequent cancers are compared in Table 16. 239 tumour combination occurrences and 66

combination types were observed in the WM series. The revised registry series contained 614 combination occurrences and 157 combination types. Combinations are shown in Table 17 and compared with the WM series in Table 18.

Table 16 – Most frequent tumours in National Cancer Registration Service – Eastern Office (≥2 tumours under 60 cases) and West Midlands series

Registry (n=1013)	Tumour number recorded	West Midlands (n = 441)	Tumour number recorded
NMSC	249 (24.6%)	Breast	128 (29%)
Breast	167 (16.5%)	Ovarian	54 (12.2%)
Melanoma	97 (9.6%)	Colorectal	71 (16.1%)
Haem B cell	72 (7.1%)	Endometrial	43 (9.7%)
Prostate	59 (5.8%)	Non-melanoma Skin	42 (9.5%)
Colorectal	53 (5.2%)		

NMSC – Non-melanoma skin cancer

Table 17 - Tumours observed in National Cancer Registration Service – Eastern office series

Quantities indicate number of combinations between x axis and y axis tumour observed in series
CNS – Central nervous system. CRC – Colorectal. NMSC – Non melanoma skin.
Yellow – Combination represents ≥2% total
Red – Population attributable fraction >50%. Blue – Among most frequent 5 UK tumours. Green – PAF ≤50% and among most frequent 5 UK tumours.

	Adrenal	Aerodigestive	Anal	Biliary	Bladder	Breast	Cervix	CNS	CRC	Endometrial	Eye	Gastric	Haem (Lymphoblastic)	Haem (B-cell)	Haem (Mast cell)	Haem (Myeloid)	Haem (T cell)	Kidney	Liver	Lung	Mediastinum	Melanoma	Mesothelioma	Neuroendocrine	NMSC	Odontogenic	Oesophagus	Ovary	Pancreas	Parathyroid	Penis	Prostate	Salivary gland	Sarcoma	Skin other	Small bowel	Testis	Thymus	Thyroid	Ureter	Uterus	Vulva
Adrenal																																									
Aerodigestive						5			1	2				3						2		3		1	13																		
Anal						2														1					1																	
Biliary						1			1																																
Bladder						1			1					1						4		1			4																
Breast		5	2	1	1		7		16	24	1			13						3		21			75			11												
Cervix						7			1	3										2		1			2																
CNS						2																1			1																
CRC		1		1	1	16	1			2				6				2		1	1	3			15			4				2								
Endometrial		2		1		24	3		2					2						1					9			10												
Eye						1																			1																
Gastric						1																			2																
Haem																									1																
Haem (B-cell)		3			1	13			6	2															30																
Haem (Mast cell)																																										
Haem (Myeloid)	1												1	1											1																
Haem (T cell)		1				1							1	3											3																
Kidney		1				1			2					4											4																
Liver														1																											
Lung		2				3	2		1	1				2				2				2			1																
Mediastinum									1																																
Melanoma		3			1	21	1	1	3					4	2					2					68																
Mesothelioma																																										
Neuroendocrine		1			1	2			1			1								1																					
NMSC	1	13	1	1	4	75	2	1	15	9	1	2	1	30		1	3	4		2		68					2													
Odontogenic																																									
Oesophagus																									2																
Ovary						11			4	10				2						1		1		1	6																
Pancreas																								1																	
Parathyroid																		1																							
Penis																				1						1															
Prostate		4			10				2					10		1		2		1		5		2	26																
Salivary gland									1					1		1		1		1					2																
Sarcoma	2	1				2		1		1				1						1		1		1	5			1				1	1							
Skin other						3																																			
Small bowel																									1																
Testis		1		1					3					1		1				1		4			4							3									
Thymus																									2																
Thyroid	1	2	1			1	1	1	1	1			1	4						1		2			9							2	1	2							
Ureter										1				1																												
Uterus																																									
Vulva						2																			2																

Table 18 - Tumour combinations representing >2% total in revised National Cancer Registration Service – Eastern Office (≥2 under 60 cases) and West Midlands series (≥2 tumours under 60 cases)

Registry	% of total combinations (n=614)	West Midlands	% of total combinations (n=239)
Breast-NMSC	12.2	Breast-Ovary	11.7
Melanoma-NMSC	11.1	Breast-NMSC	8.4
Haem B cell-NMSC	4.9	Breast-Endometrial	8
NMSC-Prostate	4.2	Breast-CRC	6.7
Breast-Endometrial	3.9	Breast melanoma	5.4
Breast-Melanoma	3.4	CRC-NMSC	5.4
Breast-CRC	2.6	Endometrial-Ovary	5.4
CRC-NMSC	2.4	CRC-Endometrial	4.2
Aerodigestive-NMSC	2.1	Bladder-CRC	3.8
Breast-Haem B cell	2.1	CRC-Ovary	3

NMSC – Non-melanoma skin cancer

Chapter 4 - Discussion

Section 1 - Clinical Genetics series and population MPMT figures

The clinical features of a series of multiple primary malignant tumour (MPMT) patients referred to a regional (West Midlands) clinical genetics service between the years of 1992 and 2012 were investigated. Half of the referrals that satisfied the inclusion criteria were referred in the last seven years of the twenty-year period. This may suggest an increasing awareness among non-genetic health care professionals of the relevance of inherited cancer syndromes for MPMT patients. It was discussed above that 16.8 cases per year might be expected to occur in the West Midlands population by chance alone if non-melanoma skin cancer (NMSC) is not considered. Discounting NMSC, the last 7 years of referral in our series produced 12.6 cases per year but not all MPMT patients will be referred for genetic assessment.

These investigations complement previous epidemiological studies that have made clinical observations of MPMT patients at a population level. A series based on cancer registry data from the USA based Surveillance Epidemiology and End Results (SEER) programme collected between 1975 and 2001 identified 742,959 (54% female) individuals who had been diagnosed with two cancers. 557,685 (75.1%) patients had developed tumours at different sites and only 10.1% of these cases had both cancers diagnosed before the age of 60 years. The relative frequency of first cancer types is quoted in the study and among this group, the

most common first cancers were breast (19.7%), lymphoma (11.6%), female genital tract (10.9%), melanoma (9.5%) and colorectal (6.8%)[65]. The high proportion of lymphoma cases may, in part, be accounted for by carcinogenic treatments for this condition, in particular the risk of breast cancer post mantle radiotherapy in lymphoma cases diagnosed <30 years, which is known to be high[84]. Analysis of the most frequent first tumours observed in the West Midlands Clinical Genetics series (WM series) reveals a similar picture, albeit with higher proportions of *BRCA1*/*BRCA2* or Lynch syndrome associated cancers with breast tumours making up 32.5% of first diagnoses, female genital tract 18.9%, colorectal 17% and melanoma 5.2%. 9% of first cancers were non-melanoma skin cancers but incidence of these was not quoted in the SEER data. There were no cases of lymphoma as a first diagnosis. Further evidence for a preferential pattern of MPMT cases referred to clinical cancer genetics could be the relatively low numbers of cases with lung (5 cases) or prostate cancers (3 cases), both common tumours.

The most frequent tumours observed overall in the WM series, therefore, reflect both incidence in the general population and the Clinical Genetics referral based nature of the patients. Breast, ovarian and colorectal cancer referrals make up over 90% of cancer genetic referrals to regional genetics services in the UK[110] with a figure of 82.2% in a similar analysis of French services[139]. In both reports, breast and bowel cancer accounted for around 60% and 20% of referrals respectively. Breast, colorectal and NMSC were among the five most frequent tumours in the WM series, an observation that is also seen in the UK population[129]. Genetic cancer predisposition is often suggested by rarer tumours but assessment of

population attributable fraction (PAF) shows relatively low figures for many of the most frequent tumours in the series including breast (26.8%), colorectal (54.4%), ovarian (20.7%) and uterine (36.9%)[22]. The remaining proportion would include known cancer predisposition syndromes, other environmental exposures and as yet unidentified genetic factors. Chemo/radiotherapy related cancers are also likely to account for a number of cases though information regarding treatment of cancers in the series was generally not recorded in the clinical records analysed.

Section 2 – Assessing the evidence for unidentified cancer predisposing mutations in undiagnosed patients referred to Clinical Genetics services

Comparison of most frequent tumours between those MPMT cases that were, and were not, found to harbour a constitutional mutation (Test+ve and Test-ve groups respectively) revealed broadly similar patterns for cancer subtypes (the five most common tumours in each group were similar). However, a higher proportion of breast cancers was observed in the Test-ve group, likely reflecting the fact that colorectal cases that would have gone on to test negative are likely to be excluded from germline genetic testing following microsatellite instability (MSI)/ immunohistochemistry (IHC) tumour analysis.

The clinical evidence for underlying mutations in cancer predisposition genes (CPGs) was assessed in those individuals not demonstrated to have such a mutation by service based investigation. While, on average, clinical comparison between Test+ve and Test-ve groups revealed a lower age at diagnosis, higher multiple tumour score (MTS) and more familial (1st degree relative affected) cases in the Test+ve group, the analysis also revealed a significant number of Test-ve individuals with clinical indicators equally or more indicative of a constitutional mutation than many of the Test+ve group cases. This was also the case if all individuals without a demonstrable mutation were analysed, including those who had not had any genetic testing performed. This suggests that there may be an appreciable, as yet unidentified mutation burden in these individuals.

A significant limitation of the MTS was that it did not consider the contribution of environmental exposures to tumourigenesis. Some strongly associated cancers have comparatively high incidence rates at younger ages (e.g. cervical cancers) and could be scored relatively highly, as could rare cancers caused by environmental factors. Although the exposure information was generally not available from the clinical records analysed, PAF figures can give a guide as to probability of a cancer being due to environmental causes. PAF is imperfect as it only considers a certain number of known exposures but is still likely to be of assistance in identifying MPMT individuals more likely to be due to genetic predisposition. With this in mind this assessment was incorporated in later analysis of MPMT cases ascertained from a regional cancer registry.

Additional analysis revealed that individuals with a higher MTS were more likely to have an elevated Manchester score or fulfill Revised Bethesda criteria. These are validated tools to predict mutation likelihood implying that MTS is also likely to be higher in patients harbouring a mutation predisposing to tumour development. This in turn suggests that the development of validated scores along similar lines incorporating all tumours may be useful in the assessment of individuals referred to cancer genetics and that further genotyping might be particularly likely to reveal mutations in cases within our series with an MTS at the higher end of the range.

Section 3 - Candidate gene analysis

Although many individuals in the Test-ve group had cancers that are associated with *PTEN* or *TP53* constitutional mutations, none were detected in the tested cases. Previous studies have assessed the frequency of mutations in these genes within an MPMT cohort. Shiseki et al found a *TP53* mutation in one out of five patients diagnosed with three primary cancers, though that individual did conform to the Li-Fraumeni syndrome (LFS) phenotype according to classical diagnostic criteria[140]. Analysis of *TP53* in 59 cases diagnosed with a second primary cancer revealed four mutation carriers[141]. None of those individuals fulfilled diagnostic criteria for LFS though all met the Chompret criteria for *TP53* testing, which incorporates many of the core features of the classical criteria but is less stringent[142]. Only one case of 21 (5%) breast and sarcoma double primaries not fulfilling classical criteria for LFS was found to have a mutation in another study[143]. Deleterious *TP53* mutations were also not identified in a series of 88 breast cancer cases with a personal or family history of MPMT[144].

De Vivo et al identified five individuals harbouring constitutional *PTEN* variants (V119L x 3 and V158L x 2) from a series of 103 MPMT cases drawn from a cohort of 32,826 nurses. There was no evidence of a diagnosis of Cowden syndrome among the variant carriers but all had been diagnosed with a tumour associated with that condition (breast and/or endometrial cancer). *In-vitro* functional studies of the identified variants via *PTEN* null cell line transfection assays suggested an increase in cell size and number[145]. The V119L variant is uncommon, involves a

conserved residue and is predicted as “probably pathogenic” on the basis of likely structural effect on the protein product by Polyphen2[132] but no additional clinical information regarding carriers of either of these variants is available. It is therefore difficult to interpret these findings as evidence of *PTEN* mutations being significant within the context of an MPMT cohort.

Testing of *CDKN2B* also revealed no firmly deleterious variants, though one case harboured a rare missense change in the gene where *in-silico* predictions were consistent with pathogenicity. However, the patient’s personal and family history did not contain occurrence of any tumours so far reported as associated with mutations in this gene and the variant is not found in publicly available databases relating genotype to disease phenotype. There is therefore insufficient evidence to assign this variant as significant in the aetiology of this patient’s cancers, particularly given that their tumours are common in the general population and could have occurred by chance. As *CDKN2B* becomes more established as a CPG both in research and clinical practice, it is likely that more robust information regarding associated tumour spectrum beyond renal cell carcinoma will become available along with a more comprehensive list of firmly pathogenic variants.

Section 4 –Possible reasons for non-detection of causative genetic changes in the MPMT series

Although the findings of this study do not suggest that *TP53*, *PTEN* or *CDKN2B* should be widely analysed in MPMT cases referred to Clinical Genetics services as single gene tests, evidence was identified in favour of the hypothesis that many may harbour constitutional mutations in CPGs. Possible reasons for non-detection both in this study and by Clinical Genetics services fall into a number of categories.

Novel genes and new phenotypic associations

TP53 and *PTEN* are associated with multiple cancers and a broad tumour spectrum but are only two CPGs from a canon now numbering over 100. Recently identified genes associated with common tumours in the WM series such as *POLD1*, *POLE* (colorectal and endometrial cancers[146]) and *POT1* (cutaneous melanoma[14]) were not available as a clinical test at the time of referral or included in our analysis and might explain some cases. It is also highly likely that a battery of novel CPGs is yet to be discovered. Many existing CPGs have well established associated phenotypes but the cancers associated with them may be amenable to revision due to previous ascertainment biases or inadequate time for the phenotype to be fully established. This study did not find any evidence to support the revision of the well-known phenotype associated with mutations in *TP53*, *PTEN* or expansion of the range of tumours caused by *CDKN2B* mutations. Other untested genes, however, that would not be expected to be causative based on

clinical assessment may account for further proportion of cases.

Mosaicism

Negative test results may result from somatic mosaicism for a mutation in a CPG, which is more likely in the absence of a family history of cancer. This phenomenon is well recognized as a cause of tumour predisposition that may evade detection by conventional genetic testing. Neurofibromatosis type 2 is a condition associated with various central nervous system tumours, particularly vestibular schwannomas. It is caused by mutations in the NF2 gene and mosaicism for a cell population containing them is estimated to account for around a third of cases[147]. Mosaicism has significant implications aside from influencing mutation detection in the laboratory. It can lead to attenuated phenotypes that may prevent further investigation for the condition in question and, when detected, is of reassurance to other family members as mosaic mutations occur post conception rather than being inherited (notwithstanding the small possibility of germline mosaicism where the cell population with the mutation is present in ovaries or testes and can be passed to offspring).

The detection of mosaicism by blood sampling depends on cells carrying mutations making up at least a proportion of circulating nucleated cells. If this is the case, the probability of detecting them will be enhanced by a greater number of distinguishable molecular enquires in the analysed DNA sample for a given genomic coordinate of interest. Chromatogram peaks from Sanger sequencing (as performed by this study) present the relative proportions of base “calls” at a

particular position from the sequencing reaction products. They may reveal mosaicism but this is not reliable and is easy to interpret as an imperfect sequence readout. Next generation sequencing (NGS) techniques are often more sensitive in detecting low levels of mosaicism due to their ability to call a particular base hundreds of times in a sample, revealing mutations that are present in only a small proportion of cells from which DNA was extracted.

Cell populations containing mutations in CPGs may not be represented in blood and present obvious difficulties with detection, even with next generation sequencing (NGS) techniques. More examples of this situation are emerging such as the finding of identical *HIF2A* mutations in a patient's paraganglioma and somatostatinoma that explained both tumour's formation. The mutation was not detected, however, in blood or other samples including urine, buccal cells and nails [148]. In the not uncommon scenario where multiple tumours occur in the same patient[66], it may be advantageous to perform genetic analysis on both tumours. The finding of a mutation common to both that is not present in blood would be reassuring for family members (as a hereditary condition becomes much less likely) and may guide treatment. Such analysis may become more widespread as NGS technologies are applied in surgical and oncological settings.

Chromosome aberrations and exonic deletions/duplications

More extensive constitutional genetic aberrations such as chromosomal abnormalities or whole exon deletions/duplications can predispose to tumour development. In the case of chromosome changes, a deletion can lead to an

individual possessing only one functional tumour suppressor allele. A duplication can affect the function of a tumour suppressor gene or increase the gene dosage of a proto-oncogene and translocation can also disrupt key genes. This study did not include karyotyping, which in any case can only detect larger rearrangements. Additionally, Sanger sequencing is not well suited to the detection of exonic deletions as the other corresponding allele still present in the sample is likely to generate a sequence readout indistinguishable from that generated by two alleles. Even NGS techniques such as multiple gene panel assays do not reliably detect chromosome abnormalities or whole exon deletion/duplications. Additional laboratory techniques such as multiplex ligation-dependent probe amplification (MLPA) are often required by diagnostic laboratories for detection of exonic deletions/duplications with karyotyping only performed in certain clinical scenarios where chromosome aberrations are more frequent (e.g. renal cell carcinoma). Tools exist for deletion/duplication detection from whole exome sequencing data though this is more difficult than where whole genome sequencing is used[149]. The latter technique also has the advantage of being able to detect and define translocation breakpoints as sequence readouts incorporating two genomic regions normally at distant chromosomal locations can be identified in the data.

Epimutations

The techniques used in this study analysed coding regions of the candidate genes only but CPGs that do not contain nucleotide variants in their coding sequence or regulatory regions may still be disrupted by epigenetic phenomena such as methylation that influences gene expression. This manner of disruption, termed

epimutation, is not detected by any of the methods discussed above and requires additional assays not routinely used in clinical laboratories. An example of a condition where epimutations are known to be relevant is Lynch syndrome, where deletions of *EPCAM* are causative in an estimated 1-3% of affected families[150]. A functional *EPCAM* gene product is not necessary for DNA mismatch repair but deletion of a 3' section of this gene that normally acts to terminate transcription can lead to disruption of the adjacent promoter of *MSH2* via continued transcription and consequent hypermethylation[151]. Assessment of the extent of epimutations in undiagnosed cancer predisposition syndromes will require further work that considers other factors in addition to a particular gene's DNA sequence itself.

Referral biases

Many MPMT cases with genetic cancer predisposition may not be referred for assessment by clinical cancer genetics services. Such individuals would therefore not be considered for genetic testing or have been entered into this study. Patients with *de novo* mutations, for example, are less likely to have pronounced family histories of cancer and fulfill referral criteria. These criteria are also influenced by cancer types most firmly associated with recognised familial cancer syndromes and may prevent assessment of individuals with suggestive clinical parameters (e.g. young age of onset) but malignancies that do not correspond to a well-established condition.

To begin to estimate what the MPMT burden due to constitutional genetic predisposition might be in the general population, data from a UK regional cancer registry covering a similar sized population to the West Midlands Regional Genetics Service was analysed. If cases occurring before the age of 60 are considered (as per the WM series), there were around 94 MPMT occurrences in the registry compared with 15 per year among West Midlands referrals.

The most frequent tumours of each series were compared and it was found that cancers recorded in the registry generally reflected population incidence. Lung cancer, however, only made up 3% of cancers in older group, possibly explained by a poorer prognosis for this condition as a first cancer and consequent increased likelihood of dying before a second cancer is diagnosed. Haematological tumours of B-cell lineage were more frequent than their population incidence might suggest, perhaps reflecting the effects of previous cancer treatment in these individuals. The WM series reflected both population incidence and referral bias, containing an over representation of cancer combinations indicative of commonly investigated suspected inherited cancer syndromes such as Hereditary Breast and Ovarian Cancer and Lynch syndrome.

To assess for evidence of constitutional genetic cancer predisposition in registry ascertained cases, tumour combinations in the series were observed. Around 22% of combinations occurred under the age of 50 years, an age at which such susceptibility is more likely. Designation of each tumour's status based on population attributable fraction (PAF) and whether it was a common cancer was

used to further stratify this likelihood. It was shown that in ~35% of combinations in the under 50 age group, both tumours had a PAF <50% and in ~18% both were uncommon and had a PAF <50%. The latter figure was significantly higher in this age group compared to the older group.

The registry data does not contain information relating to cancer therapies administered so it was not possible to assess how many combinations may have been due to treatment for the first cancer in a pairing. It also does not reveal how many cases were indeed referred for genetic assessment but numbers of cases and the nature of the tumour combinations observed suggest there may be a significant number of un-referred cases who harbour mutations in CPGs.

Section 5: Next generation sequencing in cancer genetics clinical practice

Individuals with suspected cancer predisposition syndromes are generally referred to a clinical cancer genetics service for assessment. Such services have traditionally acted as gatekeepers to testing of the single gene that is most likely to explain the phenotype of the patient consulting their service (following extensive pre-test counselling relevant to that gene).

This model has provided extensive information to patients and opportunities for risk management but has a number of limitations. Firstly, access to testing is restricted to those patients and families who conform to a phenotypic definition known to be characteristic of mutations in a given gene. This risks non-detection of significant mutations in patients whose cancers may not fit with the presumed typical phenotype and serves to maintain potentially inaccurate tumour estimates established due to ascertainment biases (see above). Secondly, sequential testing of candidate genes may be costly in terms of laboratory resources per patient and lead to a lengthy wait for results. Thirdly, the benefits of genetic testing rely on referral from a general practitioner or specialist, which may not always occur (even when referral criteria are fulfilled).

In recent years, the development of NGS technologies have begun to challenge the standard model due to the possibility of lower cost, higher throughput analysis. Using these innovations to bring the initiation of testing more into the realm of the clinician dealing with the presenting problem (e.g. oncologist) is likely to help

address the above issues. NGS tests generate a far greater amount of data than traditional single gene tests and require extensive reference to clinical, bioinformatic and biological information for interpretation. An emerging role for the clinical cancer geneticist might be to translate those results into a clinical decision and to provide post-test counselling to guide a patient through their results.

A methodology already in widespread use is that of the gene panel. In these assays, multiple genes potentially relevant to the patient's phenotype can be sequenced simultaneously. The number of genes on a panel may vary according to the number of genes associated with the relevant phenotype and the purposes of the laboratory using it.

With analysis of increasing numbers of genes comes increasing probability of identifying variants of uncertain significance, which present difficulties to the clinician in advising patients and making management decisions. This is particularly true for most gene panels as variants in most or all of the genes analysed potentially explain the phenotype of the tested patient. Pre-test counselling for this possibility is an important consideration in the use of NGS applications for diagnostic purposes.

Results generated from greater access to gene panel testing have the potential to reduce the biases resulting from testing criteria as more mutations are found in patients with phenotypes previously considered uncharacteristic for the gene in question. They increase the knowledge base pertaining to the effects of mutations

in CPGs, particularly true if clinical and variant information from these often rare cases is shared among clinicians and researchers through public databases.

However, panel testing still restricts the number of genes analysed based on those most likely to be mutated in the phenotype that the panel was designed for. Mutations in genes hitherto thought to be unrelated will therefore not be detected through this method. The likelihood of this reduces as the number of tested genes increases and some panels aim to comprehensively cover nearly all known CPGs. A yet more agnostic approach is that of whole exome sequencing (WES), which aims to provide coverage of all coding regions of the genome through selective amplification and capture of those areas prior to sequencing. Data relating to candidate genes can be selectively analysed in a “virtual panel” technique with the remainder stored for future interrogation should new candidates come to light. Alternatively, all genes can be analysed using bioinformatic and experimental techniques to identify likely causative variants. The latter approach is largely restricted to research studies. Whole genome sequencing (WGS), also largely used in an academic setting, does not rely on selective amplification and capture and generates sequence from all coding and non-coding regions. This can produce better coverage of coding regions than WES and has the additional advantages of detecting chromosomal translocations and non-coding variants that may be significant. A further result of WES and WGS is the potential to produce incidental findings i.e. variants potentially affecting health in ways that did not prompt the test e.g. carrier status for a recessive condition. These may be

desirable for patients and families to receive but should receive attention in counselling procedures prior to these assays.

Appendices

Appendix 1 – West Midlands MPMT series subgroups (cervical cancers excluded)

5 most common tumour types	All tumours (n=423)	Testing sent (n=224)	No testing sent (n=199)	Test+ve group (n=88)	Test-ve group (n=136)
Breast	122 (28.8%)	65 (27.7%)	57 (28.6%)	18 (20.4%)	47 (34.5%)
Ovarian	54 (12.8%)	35 (15.6%)	18 (9%)	14 (15.9%)	21 (15.4%)
Colorectal	69 (16.3%)	33 (14.7%)	36 (18.1%)	20 (22.7%)	13 (9.5%)
Endometrial	43 (10.2%)	28 (12.5%)	16 (8%)	12 (13.6%)	16 (11.8%)
Non Melanoma Skin	41 (9.7%)	22 (9.8%)	21 (10.5%)	7 (7.9%)	15 (11%)
Other tumours comprising < 5% total	94 (22.2%)	41 (18.3%)	51 (25.6%)	17 (19.3%)	24 (17.6%)

NMSC – Non melanoma skin cancer

Appendix 2 – West Midland MPMT series clinical analysis (cervical cancers excluded)

	All cases (n=203)	Mutation identified (Test+ve) (n=40)	No mutation identified. ≥ 1 germ-line test (Test-ve) (n=66)	No mutation identified (n=163)
Mean age at tumour diagnosis	46.4 SD 10.1	44.2 SD 11	46.1 (SD 9.8 p=0.3692)	47 (SD 9.8 p=0.1408)
Mean age diagnosis 1st tumour	42.7 SD 11.4	39.3 SD 13	42.8 (SD 10.2 p=0.1462)	43.5 (SD 10.8 p=0.0588)
Mean age diagnosis 2nd or synchronous tumour	49.7 SD 7.2	48.1 SD 6.7	49.2(SD 8.2 p=0.4522)	50.1 (SD 7.3 p=0.0966)
Mean individual age at tumour diagnosis	46.4 SD 8.4	44.4 SD 8.8	46.1 (SD 8.6 p=0.3309)	47 (SD 8.2 p=0.0898)
% cases with individual mean age at tumour diagnosis ≤ Test+ve group median (= 46.6)	N/A	N/A	42.4% (38)	39.2% (64)
% cases with concordant tumour in 1st degree relative	52.2% (106)	65% (26)	57.6% (38) p=0.4473	49% (80) p=0.0703
Mean multiple tumour score	14.7 SD 8.9	21.6 SD 14.6	14.2 (SD 5.4 p =0.0021)	13 (SD 5.6 p=0.0003)
% cases with multiple tumour score ≥ Test+ve group median (=16.5)	N/A	N/A	28.8% (19)	21.5% (35)

Mean individual age at diagnosis = Combined age of tumours ÷ number of tumours in individual. P values describe two tailed comparison with Test+ve group ($H_0 = \mu_1 = \mu_2$).

Appendix 3 – Tumour subgroups observed in National Cancer Registration Service – Eastern Office series

≥2 tumours under 50		≥2 tumours not until after 50	
Cancer type	**No. cases. % total in brackets**	**Cancer type**	**No. cases. % total in brackets**
Non-melanoma skin	60 (20.7)	Non-melanoma skin	222 (24.6)
Breast	52 (18.0)	Breast	130 (14.4)
Melanoma	38 (13.1)	Haematological (B-cell)	75 (8.3)
Haematological (B-cell)	17 (5.9)	Melanoma	73 (8.1)
Colorectal	12 (4.1)	Prostate	72 (8)
Endometrium	10 (3.4)	Colorectal	57 (6.3)
Thyroid	10 (3.4)	Endometrium	39 (4.3)
Aerodigestive tract	8 (2.8)	Aerodigestive tract	34 (3.8)
Sarcoma	8 (2.8)	Lung	27 (3)
Cervix	7 (2.4)	Ovary	25 (2.8)
Central nervous system	7 (2.4)	Kidney	18 (2)
Bladder	6 (2.1)	Thyroid	18 (2)
Ovary	6 (2.1)	Bladder	16 (1.8)
Prostate	6 (2.1)	Testis	13 (1.4)
Testis	6 (2.1)	Neuroendocrine	12 (1.3)
Kidney	5 (1.7)	Haematological (T-cell)	8 (0.9)
Haematological (Lymphoblastic)	4 (1.4)	Cervix	7 (0.8)
Haematological (Myeloid)	4 (1.4)	Sarcoma	7 (0.8)
Salivary gland	4 (1.4)	Gastric	6 (0.7)
Haematological (T-cell)	3 (1.0)	Salivary gland	6 (0.7)
Neuroendocrine	3 (1.0)	Anus	5 (0.5)
Adrenal	2 (0.7)	Haematological (Myeloid)	4 (0.4)
Haematological (Mast cell)	2 (0.7)	Eye	3 (0.3)
Lung	2 (0.7)	Skin (other)	3 (0.3)
Biliary	1 (0.3)	Biliary	2 (0.2)
Gastric	1 (0.3)	Mesothelioma	2 (0.2)
Mediastinum	1 (0.3)	Oesophagus	2 (0.2)
Oesophagus	1 (0.3)	Penis	2 (0.2)
Skin (other)	1 (0.3)	Small bowel	2 (0.2)
Ureter	1 (0.3)	Thymus	2 (0.2)
Uterus	1 (0.3)	Ureter	2 (0.2)
Vulva	1 (0.3)	Vulva	2 (0.2)
		Central nervous system	1 (0.1)
		Liver	1 (0.1)
		Odontogenic	1 (0.1)
		Pancreas	1 (0.1)
		Parathyroid	1 (0.1)

References

1. Whitworth, J. *et al.* A clinical and genetic analysis of multiple primary cancer referrals to genetics services. *Eur. J. Hum. Genet.* 1–7 (2014). doi:10.1038/ejhg.2014.157

2. Wunderlich, V. JMM — Past and Present. *J. Mol. Med.* **80,** 545–548 (2002).

3. Boveri, T. Concerning the origin of malignant tumours by Theodor Boveri. Translated and annotated by Henry Harris. *J. Cell Sci.* **121 Suppl ,** 1–84 (2008).

4. Nowell, P. C. Review series personal perspective Discovery of the Philadelphia chromosome : a personal perspective. **117,** 2033–2035 (2007).

5. Bishop, M. Cellular Oncogenes. (1983).

6. Nordling, C. O. A new theory on cancer-inducing mechanism. *Br. J. Cancer* **7,** 68–72 (1953).

7. Knudson, A. G. Mutation and Cancer : Statistical Study of Retinoblastoma. **68,** 820–823 (1971).

8. Yunis, J. J. & Ramsay, N. Retinoblastoma and subband deletion of chromosome 13. American journal of diseases of children (1911) **132,** 161–163 (1978).

9. Balaban, G., Gilbert, F., Nichols, W., Meadows, A. T. & Shields, J. Abnormalities of chromosome #13 in retinoblastomas from individuals with normal constitutional karyotypes. *Cancer Genet. Cytogenet.* **6,** 213–221 (1982).

10. Fung, Y. T. *et al.* Structural Evidence for the Authenticity of the Human Retinoblastoma Gene. *Science (80-.).* **236,** 1657–1661 (1986).

11. Mulligan, L. M. *et al.* Germ-line mutations of the RET proto-oncogene in multiple endocrine neoplasia type 2A. *Nature* **363,** 458–460 (1993).

12. Donis-Keller, H. *et al.* Mutations in the RET proto-oncogene are associated with MEN 2A and FMTC. *Hum. Mol. Genet.* **2,** 851–856 (1993).

13. Benusiglio, P. R. *et al.* A germline mutation in PBRM1 predisposes to renal cell carcinoma. *J. Med. Genet.* (2015). doi:10.1136/jmedgenet-2014-102912

14. Robles-espinoza, C. D. *et al.* POT1 loss-of-function variants predispose to familial melanoma. *Nat. Genet.* (2014). doi:10.1038/ng.2947

15. Malkin, D. *et al.* Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. *Science* **250,** 1233–1238 (1990).

16. cBioPortal for Cancer Genomics. at <http://www.cbioportal.org/>

17. Seal, S. *et al.* Truncating mutations in the Fanconi anemia J gene BRIP1 are low-penetrance breast cancer susceptibility alleles. *Nat. Genet.* **38,** 1239–1241 (2006).

18. Rahman, N. *et al.* PALB2, which encodes a BRCA2-interacting protein, is a breast cancer susceptibility gene. *Nat. Genet.* **39,** 165–167 (2007).

19. Antoniou, A. C. *et al.* Breast-Cancer Risk in Families with Mutations in PALB2. *N. Engl. J. Med.* **371,** 497–506 (2014).

20. Mavaddat, N. *et al.* Prediction of Breast Cancer Risk Based on Profiling With Common Genetic Variants. *JNCI J. Natl. Cancer Inst.* **107,** djv036–djv036 (2015).

21. Gimenez-Roqueplo, A. P. *et al.* Phaeochromocytoma, new genes and screening strategies. *Clinical Endocrinology* **65,** 699–705 (2006).

22. Boyd, L. & Parkin, D. M. The fraction of cancer attributable to lifestyle and environmental factors in the UK in 2010. *Journal of Epidemiology & Community Health* **65,** A143–A143 (2011).

23. Lammens, C. R. M. *et al.* Genetic testing in Li-Fraumeni syndrome: Uptake and psychosocial consequences. *J. Clin. Oncol.* **28,** 3008–3014 (2010).

24. Butow, P. N., Lobb, E. A., Meiser, B., Barratt, A. & Tucker, K. M. Psychological outcomes and risk perception after genetic testing and counselling in breast cancer: a systematic review. *Med. J. Aust.* **178,** 77–81 (2003).

25. Metcalfe, K. A. *et al.* An evaluation of needs of female BRCA1 and BRCA2 carriers undergoing genetic counselling. *J. Med. Genet.* **37,** 866–74 (2000).

26. Barrow, P. *et al.* Improving the uptake of predictive testing and colorectal screening in Lynch syndrome: a regional primary care survey. *Clin. Genet.* (2015). doi:10.1111/cge.12559

27. Meijers-Heijboer, E. J. *et al.* Presymptomatic DNA testing and prophylactic surgery in families with a BRCA1 or BRCA2 mutation. *Lancet* **355,** 2015–2020 (2000).

28. Esplen, M. J. *et al.* Motivations and psychosocial impact of genetic testing for HNPCC. *Am. J. Med. Genet.* **103,** 9–15 (2001).

29. Human Fertilisation and Embryology Authority, S. and I. D. W. team,. PGD conditions licensed by the HFEA - testing and screening. at <http://guide.hfea.gov.uk/pgd/>

30. Dewanwala, A. *et al.* Attitudes toward childbearing and prenatal testing in individuals undergoing genetic testing for Lynch Syndrome. *Fam. Cancer* **10,** 549–556 (2011).

31. Wilding, A. *et al.* Life expectancy in hereditary cancer predisposing diseases: an observational study. *J. Med. Genet.* **49,** 264–269 (2012).

32. Magnetic Resonance Imaging Screening in Li Fraumeni Syndrome - Full Text View - ClinicalTrials.gov. at <https://clinicaltrials.gov/ct2/show/NCT01737255>

33. Menko, F. H. *et al.* Hereditary leiomyomatosis and renal cell cancer (HLRCC): renal cancer risk, surveillance and treatment. *Fam. Cancer* 637–644 (2014). doi:10.1007/s10689-014-9735-2

34. Rebbeck, T. R. *et al.* Bilateral prophylactic mastectomy reduces breast cancer risk in BRCA1 and BRCA2 mutation carriers: the PROSE Study Group. *J. Clin. Oncol.* **22,** 1055–62 (2004).

35. Jasperson, K. W. & Burt, R. W. APC-Associated Polyposis Conditions. (2014). at <http://www.ncbi.nlm.nih.gov/books/NBK1345/>

36. Schuler, M. Aspirin and the risk of colorectal cancer in women. *N. Engl. J. Med.* **334,** 121–122 (1996).

37. Burn, J. *et al.* Long-term effect of aspirin on cancer risk in carriers of hereditary colorectal cancer: an analysis from the CAPP2 randomised controlled trial. *Lancet* **378,** 2081–7 (2011).

38. Johnson, R. L. *et al.* Human homolog of patched, a candidate gene for the basal cell nevus syndrome. *Science* **272,** 1668–1671 (1996).

39. Hahn, H. *et al.* Mutations of the human homolog of Drosophila patched in the nevoid basal cell carcinoma syndrome. *Cell* **85,** 841–851 (1996).

40. Gailani, M. R. *et al.* Developmental defects in Gorlin syndrome related to a putative tumor suppressor gene on chromosome 9. *Cell* **69,** 111–117 (1992).

41. Aszterbaum, M. *et al.* Inhibiting the Hedgehog Pathway in Patients with the Basal-Cell Nevus Syndrome. (2012).

42. Farmer, H. *et al.* Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. *Nature* **434,** 917–921 (2005).

43. Nik-Zainal, S. *et al.* Mutational processes molding the genomes of 21 breast cancers. *Cell* **149,** 979–93 (2012).

44. Rahman, N. Realizing the promise of cancer predisposition genes. *Nature* **505,** 302–8 (2014).

45. Al-Tassan, N. *et al.* Inherited variants of MYH associated with somatic G:C-->T:A mutations in colorectal tumors. *Nat. Genet.* **30,** 227–232 (2002).

46. Astuti, D. *et al.* Gene Mutations in the Succinate Dehydrogenase Subunit SDHB Cause Susceptibility to Familial Pheochromocytoma and to Familial Paraganglioma. **3,** 49–54 (2001).

47. Alston, C. L. *et al.* Recessive germline SDHA and SDHB mutations causing leukodystrophy and isolated mitochondrial complex II deficiency. *Journal of Medical Genetics* **49,** 569–577 (2012).

48. Gatti, R. Ataxia-Telangiectasia. (2010). at <http://www.ncbi.nlm.nih.gov/books/NBK26468/>

49. Schneider, K., Zelley, K., Nichols, K. & Garber, J. *Li-Fraumeni Syndrome*. *GeneReviews™ [Internet]* (University of Washington, Seattle, 2013). at <http://www.ncbi.nlm.nih.gov/books/NBK1311/>

50. Vasen, H. F., Mecklin, J. P., Khan, P. M. & Lynch, H. T. The International Collaborative Group on Hereditary Non-Polyposis Colorectal Cancer (ICG-HNPCC). *Diseases of the colon and rectum* **34,** 424–425 (1991).

51. Umar, A. *et al.* Revised Bethesda Guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability. *Journal of the National Cancer Institute* **96,** 261–268 (2004).

52. Aarnio, M. *et al.* Cancer risk in mutation carriers of DNA-mismatch-repair genes. *Int. J. Cancer* **81,** 214–218 (1999).

53. Stoffel, E. *et al.* Calculation of risk of colorectal and endometrial cancer among patients with Lynch syndrome. *Gastroenterology* **137,** 1621–7 (2009).

54. Harbour, J. W. *et al.* Frequent mutation of BAP1 in metastasizing uveal melanomas. *Science* **330,** 1410–1413 (2010).

55. Horsman, D. E. & White, V. a. Cytogenetic analysis of uveal melanoma: Consistent occurrence of monosomy 3 and trisomy 8q. *Cancer* **71,** 811–819 (1993).

56. Abdel-Rahman, M. H. *et al.* Germline BAP1 mutation predisposes to uveal melanoma, lung adenocarcinoma, meningioma, and other cancers. *J. Med. Genet.* **48,** 856-9 (2011).

57. Popova, T. *et al.* Germline BAP1 mutations predispose to renal cell carcinomas. *Am. J. Hum. Genet.* **92,** 974–80 (2013).

58. Giusti, F., Marini, F. & Brandi, M. L. in *GeneReviews^{TM} [Internet]* (ed. Pagon RA, Adam MP, Bird TD, et al.) (University of Washington, Seattle, 2013). at <http://www.ncbi.nlm.nih.gov/books/NBK1257/>

59. Kloos, R. T. *et al.* Medullary thyroid cancer: management guidelines of the American Thyroid Association. *Thyroid* **19,** 565–612 (2009).

60. Seri, M. *et al.* A Cys634Gly substitution of the RET proto-oncogene in a family with recurrence of multiple endocrine neoplasia type 2A and cutaneous lichen amyloidosis. *Clin. Genet.* **51,** 86–90 (1997).

61. Hofstra, R. M. *et al.* A mutation in the RET proto-oncogene associated with multiple endocrine neoplasia type 2B and sporadic medullary thyroid carcinoma. *Nature* **367,** 375–376 (1994).

62. NHGRI: Breast Cancer Information Core. at <https://research.nhgri.nih.gov/projects/bic/>

63. Billroth, T. *Die Allgemeine Chirurgische Pathologie and Therapie. 51 Vorlesungen. Ein Handbuch fur Studierende and Artze*. (G Reimer, 1889).

64. Berrino, F. *et al.* Survival for eight major cancers and all cancers combined for European adults diagnosed in 1995-99: results of the EUROCARE-4 study. *Lancet Oncol.* **8,** 773–783 (2007).

65. Mariotto, A. B., Rowland, J. H., Ries, L. a G., Scoppa, S. & Feuer, E. J. Multiple cancer prevalence: a growing challenge in long-term survivorship. *Cancer Epidemiol. Biomarkers Prev.* **16,** 566–71 (2007).

66. Rosso, S. *et al.* Multiple tumours in survival estimates. *Eur. J. Cancer* **45,** 1080–1094 (2009).

67. Travis, L. B. The epidemiology of second primary cancers. *Cancer Epidemiol Biomarkers Prev* **15,** 2020–2026 (2006).

68. Dong, C. & Hemminki, K. Second primary neoplasms in 633,964 cancer patients in Sweden, 1958-1996. *Int. J. Cancer* **93,** 155–161 (2001).

69. UK, C. R. Lifetime risk of cancer. (2014). at <http://www.cancerresearchuk.org/cancer-info/cancerstats/incidence/risk/>

70. UK, C. R. Cancer incidence by age. at <http://www.cancerresearchuk.org/cancer-info/cancerstats/incidence/age/>

71. UK, C. R. Skin cancer incidence statistics. (2014). at <http://www.cancerresearchuk.org/cancer-info/cancerstats/types/skin/incidence/>

72. Brobeil, A. *et al.* Multiple primary melanomas: implications for screening and follow-up programs for melanoma. *Ann. Surg. Oncol.* **4,** 19–23 (1997).

73. Bagni, O., Filippi, L. & Schillaci, O. Incidental Detection of Colorectal Cancer Via 18F-Choline PET/CT in a Patient With Recurrent Prostate Cancer: Usefulness of Early Images. *Clin Nucl Med* **40,** e328–30 (2015).

74. Bertagna, F. *et al.* Incidental 11C-choline PET/CT uptake due to esophageal carcinoma in a patient studied for prostate cancer. *Clin. Nucl. Med.* (2014). doi:10.1097/RLU.0000000000000334

75. Zaino, R. *et al.* Simultaneously detected endometrial and ovarian carcinomas--a prospective clinicopathologic study of 74 cases: a gynecologic oncology group study. *Gynecol. Oncol.* **83,** 355–362 (2001).

76. Abe, O., Abe, R., Enomoto, K. & Kikuchi, K. Tamoxifen for early breast cancer: an overview of the randomised trials. *Lancet* **351,** 1451–1467 (1998).

77. De Meerleer, G. *et al.* Radiotherapy for renal-cell carcinoma. *The Lancet Oncology* **15,** (2014).

78. Allan, J. M. & Travis, L. B. Mechanisms of therapy-related carcinogenesis. *Nat. Rev. Cancer* **5,** 943–955 (2005).

79. Ng, J. & Shuryak, I. Minimizing second cancer risk following radiotherapy: current perspectives. *Cancer Manag. Res.* 1 (2014). doi:10.2147/CMAR.S47220

80. Preston, D. L. *et al.* Solid cancer incidence in atomic bomb survivors: 1958-1998. *Radiat. Res.* **168,** 1–64 (2007).

81. Nikiforov, Y. E. Radiation-induced thyroid cancer: What we have learned from chernobyl. *Endocrine Pathology* **17,** 307–317 (2006).

82. Inskid, P. D. in *Multiple Primary Cancers* (eds. Neugut, A., Meadows, A. & Robinson, E.) 91–136 (Lippincott Williams and Wilkins, 1999).

83. Curtis, R. E. *et al.* Relationship of leukemia risk to radiation dose following cancer of the uterine corpus. *J. Natl. Cancer Inst.* **86,** 1315–1324 (1994).

84. Yahalom, J. & Petrek, J. Breast cancer in patients irradiated for Hodgkin's disease: a clinical and pathologic analysis of 45 events in 37 patients. *J. Clin Oncol.***10,** 1674–1681 (1992).

85. Crump, M. & Hodgson, D. Secondary breast cancer in Hodgkin's lymphoma survivors. *Journal of Clinical Oncology* **27,** 4229–4231 (2009).

86. De Bruin, M. L. *et al.* Breast cancer risk in female survivors of Hodgkin's lymphoma: Lower risk after smaller radiation volumes. *J. Clin. Oncol.* **27,** 4239–4246 (2009).

87. Ezoe, S. Secondary leukemia associated with the anti-cancer agent, etoposide, a topoisomerase II inhibitor. *Int. J. Environ. Res. Public Health* **9,** 2444–2453 (2012).

88. Travis, L. B. *et al.* Bladder and kidney cancer following cyclophosphamide therapy for non-Hodgkin's lymphoma. *J. Natl. Cancer Inst.* **87,** 524–530 (1995).

89. Rubino, C. *et al.* Long-term risk of second malignant neoplasms after neuroblastoma in childhood: Role of treatment. *Int. J. Cancer* **107,** 791–796 (2003).

90. Roychoudhuri, R., Evans, H., Robinson, D. & Møller, H. Radiation-induced malignancies following radiotherapy for breast cancer. *Br. J. Cancer* **91,** 868–872 (2004).

91. Breslow, N. E. *et al.* Second malignant neoplasms following treatment for Wilms' tumor: A report from the National Wilms' Tumor Study Group. *J. Clin. Oncol.* **13,** 1851–1859 (1995).

92. Birdwell, S. H., Hancock, S. L., Varghese, A., Cox, R. S. & Hoppe, R. T. Gastrointestinal cancer after treatment of Hodgkin's disease. *Int. J. Radiat. Oncol. Biol. Phys.* **37,** 67–73 (1997).

93. Marín-Gutzke, M. *et al. Basal cell carcinoma in childhood after radiation therapy: case report and review. Annals of plastic surgery* **53,** 593–595 (2004).

94. Evans, D. G. R., Birch, J. M., Ramsden, R. T., Sharif, S. & Baser, M. E. Malignant transformation and new primary tumours after therapeutic radiation for benign disease: substantial risks in certain tumour prone syndromes. *J. Med. Genet.* **43,** 289–294 (2006).

95. Rodriguez-Antona, C. & Ingelman-Sundberg, M. Cytochrome P450 pharmacogenetics and cancer. *Oncogene* **25,** 1679–1691 (2006).

96. Levi, F., Randimbison, L., Te, V. C. & La Vecchia, C. Second primary cancers in patients with lung carcinoma. *Cancer* **86,** 186–90 (1999).

97. Health and Social Care Information Centre. *Statistics on Smoking - England 2014*. (2014). at <http://www.hscic.gov.uk/catalogue/PUB14988/smok-eng-2014-rep.pdf>

98. Health Survey for England. *Health Survey for England - 2013*. (2014). at <http://www.hscic.gov.uk/catalogue/PUB16076/HSE2013-Ch10-Adult-anth-meas.pdf>

99. De Leon, J. *et al.* Association between smoking and alcohol use in the general population: Stable and unstable odds ratios across two years in two different countries. *Alcohol Alcohol.* **42,** 252–257 (2007).

100. Hemminki, K., Zhang, H., Sundquist, J. & Lorenzo Bermejo, J. Modification of risk for subsequent cancer after female breast cancer by a family history of breast cancer. *Breast Cancer Res Treat* **111,** 165–169 (2008).

101. Hemminki, K., Li, X. & Dong, C. Second primary cancers after sporadic and familial colorectal cancer. *Cancer Epidemiol. Biomarkers Prev.* **10,** 793–798 (2001).

102. Gonzalez, K. D. *et al.* Beyond Li Fraumeni Syndrome: clinical characteristics of families with p53 germline mutations. *J. Clin. Oncol.* **27,** 1250–6 (2009).

103. Kleinerman, R. *et al.* Variation of second cancer risk by family history of retinoblastoma among long-term survivors. *J. Clin. Oncol.* **30,** 950–7 (2012).

104. Lim, W. *et al.* Relative frequency and morphology of cancers in STK11 mutation carriers1 ☆. *Gastroenterology* **126,** 1788–1794 (2004).

105. Metcalfe, K. a *et al.* The risk of ovarian cancer after breast cancer in BRCA1 and BRCA2 carriers. *Gynecol. Oncol.* **96,** 222–6 (2005).

106. Win, A. K. *et al.* Risks of colorectal and other cancers after endometrial cancer for women with Lynch syndrome. *J. Natl. Cancer Inst.* **105,** 274–9 (2013).

107. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4,** 44–57 (2009).

108. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* **37,** 1–13 (2009).

109. Eng, C. *PTEN Hamartoma Tumor Syndrome (PHTS)*. (University of Washington, Seattle, 2012). at <http://www.ncbi.nlm.nih.gov/books/NBK1488/>

110. Wonderling, D. *et al.* A descriptive study of UK cancer genetics services: an emerging clinical response to the new genetics. *Br. J. Cancer* **85,** 166–170 (2001).

111. Hisada, M., Garber, J. E., Fung, C. Y., Fraumeni, J. F. & Li, F. P. Multiple primary cancers in families with Li-Fraumeni syndrome. *J. Natl. Cancer Inst.* **90,** 606–11 (1998).

112. Tan, M.-H. *et al.* Lifetime cancer risks in individuals with germline PTEN mutations. *Clin. Cancer Res.* **18,** 400–7 (2012).

113. Jafri, M. *et al.* Germline Mutations in the CDKN2B tumor suppressor gene predispose to renal cell carcinoma. *Cancer Discov.* **Epub ahead,** (2015).

114. Platz, A. *et al.* Screening of germline mutations in the CDKN2A and CDKN2B genes in Swedish families with hereditary cutaneous melanoma. *J. Natl. Cancer Inst.* **89,** 697–702 (1997).

115. Laud, K. *et al.* Comprehensive analysis of CDKN2A (p16INK4A/p14ARF) and CDKN2B genes in 53 melanoma index cases considered to be at heightened risk of melanoma. *J. Med. Genet.* **43,** 39–47 (2006).

116. Boultwood, J. & Wainscoat, J. S. Gene silencing by DNA methylation in haematological malignancies. *British Journal of Haematology* **138,** 3–11 (2007).

117. Birmingham Women's Hospital NHS Foundation Trust. West Midlands Family Cancer Service - The West Midlands Region. (2013). at <http://www.bwhct.nhs.uk/west-midlands-family-cancer-service/the-west-midlands-region>

118. Manchester Centre for Genomic Medicine. Manchester Centre for Genomic Medicine. (2013). at <http://www.manchesterbrc.org/OurFacilities/ManchesterCentreforGenomicMedicine.php>

119. International Agency for Research on Cancer. *International Rules for Multiple Primary Cancers (ICD-O Third Edition)*. (2004).

120. Evans, D. G. R. A new scoring system for the chances of identifying a BRCA1/2 mutation outperforms existing models including BRCAPRO. *J. Med. Genet.* **41,** 474–480 (2004).

121. Cancer Research UK. Cervical cancer incidence statistics. (2013). at <http://www.cancerresearchuk.org/cancer-info/cancerstats/types/cervix/incidence/#source1>

122. Baseman, J. G. & Koutsky, L. A. The epidemiology of human papillomavirus infections. *J. Clin. Virol.* **32 Suppl 1,** S16–24 (2005).

123. Institute of Cancer Research. Protocol 2 BRCA1 and BRCA2 mutation testing guidelines Woman with Woman with Male with breast cancer who Individual unaffected. 2013 (2013). at <http://www.icr.ac.uk/research/team_leaders/Rahman_Nazneen/Rahman_Nazneen_Protocols/Protocols/22778.pdf>

124. REPLI-g Mini Kit - QIAGEN. at <https://www.qiagen.com/gb/products/catalog/sample-technologies/dna-sample-technologies/genomic-dna/repli-g-mini-kit/#productdetails>

125. Flicek, P. *et al.* Ensembl 2013. *Nucleic Acids Res.* **41,** (2013).

126. Untergasser, A. *et al.* Primer3-new capabilities and interfaces. *Nucleic Acids Res.* **40,** (2012).

127. Ye, J. *et al.* Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics* **13,** 134 (2012).

128. James Kent, W. *et al.* The human genome browser at UCSC. *Genome Res.* **12,** 996–1006 (2002).

129. Cancer Research UK. Cancer incidence for common cancers. (2013). at <http://www.cancerresearchuk.org/cancer-info/cancerstats/incidence/commoncancers/>

130. Applications, G. & Group, W. Recommendations from the EGAPP Working Group: genetic testing strategies in newly diagnosed individuals with colorectal cancer aimed at reducing morbidity and mortality from Lynch syndrome in relatives. *Genet. Med.* **11,** 35–41 (2009).

131. SNPedia. Rs1042522 - SNPedia. (2013). at <http://snpedia.com/index.php/Rs1042522>

132. NHLBI Exome Sequencing Project (ESP). Exome Variant Server. (2013). at <http://evs.gs.washington.edu/EVS/>

133. Davydov, E. V. *et al.* Identifying a high fraction of the human genome to be under selective constraint using GERP++. *PLoS Comput. Biol.* **6,** (2010).

134. Adzhubei, I. a *et al.* A method and server for predicting damaging missense mutations. *Nat. Methods* **7,** 248–9 (2010).

135. Landrum, M. J. *et al.* ClinVar: Public archive of relationships among sequence variation and human phenotype. *Nucleic Acids Res.* **42,** (2014).

136. Stenson, P. D. *et al.* Human Gene Mutation Database (HGMD®): 2003 Update. *Human Mutation* **21,** 577–581 (2003).

137. CDKN2B homepage - LOVD - Leiden Open Variation Database - Leiden Open Variation Database. at <http://databases.lovd.nl/shared/genes/CDKN2B>

138. The National Cancer Registration Service - Eastern Office. at <http://ecric.nhs.uk/>

139. Pujol, P. *et al.* Lack of referral for genetic counseling and testing in BRCA1/2 and Lynch syndromes: a nationwide study based on 240,134 consultations and 134,652 genetic tests. *Breast Cancer Res. Treat.* **141,** 135–44 (2013).

140. Shiseki, M. *et al.* Germ-line p53 mutation is uncommon in patients with triple primary cancers. *Cancer Lett.* **73,** 51–57 (1993).

141. Malkin, D. *et al. Germline mutations of the p53 tumor-suppressor gene in children and young adults with second malignant neoplasms. The New England journal of medicine* **326,** 1309–1315 (1992).

142. Chompret, A. *et al.* Sensitivity and predictive value of criteria for p53 germline mutation. *J. Med. Genet.* **38,** 43–47 (2001).

143. Evans, D. *et al.* Low rate of TP53 germline mutations in breast cancer/sarcoma families not fulfilling classical criteria for Li-Fraumeni syndrome. *J. Med. Genet.* **39,** 941–944 (2002).

144. Martin, A. *et al.* Germline TP53 mutation in breast cancer families with multiple primary cancers: is Tp53 a modifier of BRCA1? *J. Med. Genet.* **40,** (2003).

145. De Vivo, I. *et al.* Novel germline mutations in the PTEN tumour suppressor gene found in women with multiple cancers. *J. Med. Genet.* **37,** 336–341 (2000).

146. Palles, C. *et al.* Germline mutations affecting the proofreading domains of POLE and POLD1 predispose to colorectal adenomas and carcinomas. *Nat. Genet.* **45,** 136–44 (2013).

147. Moyhuddin, A *et al.* Somatic mosaicism in neurofibromatosis 2: prevalence and risk of disease transmission to offspring. *J. Med. Genet.* **40,** 459–463 (2003).

148. Zhuang, Z. *et al.* Somatic HIF2A gain-of-function mutations in paraganglioma with polycythemia. *N. Engl. J. Med.* **367,** 922–30 (2012).

149. Tan, R. *et al.* An Evaluation of Copy Number Variation Detection Tools from Whole-Exome Sequencing Data. *Hum. Mutat.* **35,** 899–907 (2014).

150. Kuiper, R. P. *et al.* Recurrence and variability of germline EPCAM deletions in Lynch syndrome. *Hum. Mutat.* **32,** 407–414 (2011).

151. Ligtenberg, M. J. L. *et al.* Heritable somatic methylation and inactivation of MSH2 in families with Lynch syndrome due to deletion of the 3' exons of TACSTD1. *Nat. Genet.* **41,** 112–117 (2009).