DEVELOPING A LOCAL CASUISTIC ADAPTED PROTOCOL FOR SELECTING COLORECTAL CANCER PATIENTS INCLUSION IN MMR GENETIC TESTING

LUCIAN NEGURĂ¹*, ANDREI CHICOȘ¹,2, VALERIU AURELIAN CHIRICA¹, ROXANA AVÂDANEI¹, ROXANA POSTOLICĂ³, ANCA NEGURĂ⁴

Received: 30 January 2017 / Revised: 20 February 2017 / Accepted: 20 March 2017 / Published: 4 April 2017

Keywords: Lynch syndrome, MMR genes, genetic testing, local adapted protocol

Abstract: Up to 30% of colorectal cancers (CRCs) have evidence of a familial component, and about 5% are thought to be due to inherited mutations in MMR (MisMatch Repair) genes such as MSH2, MLH1, MSH6. Hereditary Non-Polyposis Colorectal cancers (HNPCC or Lynch syndrome) are characterized by a very early onset, while the overall lifetime risk of CRC for those patients is estimated to be 80%. The identification of a germline mutation is the gold standard for the diagnosis of Lynch syndrome and is performed by complete Sanger sequencing of all 3 genes, which represents a very complex and expensive analysis. MMR mutations are detected in only 60% of criteria fulfilling families, while germline mutations in MMR genes are present in up to 20% of families not fulfilling these criteria and which are implicitly excluded from genetic counselling. Therefore, we propose an adapted algorithm, based on germline and tumor analysis, intended to increase molecular diagnostic efficiency and CRC casuistry coverage.

INTRODUCTION

Colorectal cancer (CRC) is the third most common cancer in men (746,000 cases per year, 10,0% of the total) and the second in women (614,000 cases per year, 9,2% of the total) worldwide (Ferlay et al., 2015), and the third most common cause of cancer death in the world (694,000 deaths per year, 8,5% of the total) (Patel et al., 2012; Center et al., 2009). The incidence of CRC in Romania (Ferlay et al., 2015) is estimated at 26,42 per 100,000 for both sexes, while mortality is estimated at 13,37 per 100,000. Epidemiological studies (Center et al., 2009; Valean et al., 2008) highlighted an alarming CRC incidence and mortality increase in Romania for the past decade, i.e. a 75%-fold incidence increase and mortality rates increasing 2.9% per year, Romania being one of the few countries worldwide where increases in mortality rates are still occurring. These epidemiologic trends are similar with those recorded in Western Europe in the early ‘90s, where the increase awareness about CRC and methods of screening and early diagnosis determined a stabilization of incidence and mortality of CRC in the late ‘90s (Jemal et al., 2005). This moment also corresponds to the apparition and development of Oncogenetics, as the medical and diagnostic follow-up of patients and their families presenting a hereditary monogenic risk to cancer. The prognosis for patients with CRC is heavily dependent on stage at diagnosis (de la Chapelle et al., 2004): 5-year survival is over 90% for early stages, but only 5% for late ones, which emphasis early diagnostic.

Up to 30% of CRCs have evidence of a familial component (Risch et al., 2001) and about 5% are thought to be due to inherited mutations in known genes (Patel et al., 2012), the rest being due to low-penetrance alleles (de la Chapelle et al., 2004). Lynch syndrome (LS) – named after Dr. Henry Lynch (Lynch et al., 1977) – has been proposed as a more appropriate name for the Hereditary Non-Polyposis Colorectal Cancer (HNPCC) (Schlussel et al., 2014) because the syndrome involves also predisposition to other sites (endometrium, stomach, ovaries, small bowel, hepatobiliary epithelium, urogenital epithelium and brain) and increased frequency of multiplicity for CRC. Germline mutations in the mismatch repair (MMR) genes MSH2 (OMIM 609309, 2p21), MLH1 (OMIM 120436, 3p22.2) and MSH6 (OMIM 600678, 2p16.3), and in a much lower proportion the PMS2 gene, cause LS with a penetrance of approximately 80% for CRC, 60% for endometrial cancer, and well below 20% for the other cancers (Lynch et al., 2003). MMR function as tumor suppressors to maintain genome integrity and fidelity of DNA replication (Chung et al., 2003). MMR germ-line mutations are the first hit in apparition of CRC (mutator phenotype), the second hit can occur by deletion (loss of heterozygosity), mutation, or methylation of CpG islands in the gene promoter (de la Chapelle et al., 2004). Mutations in MSH2 and MLH1 account for up to 90% and MSH6 about 10% of LS families. The mean age at diagnostic for sporadic CRC is 69 years, while hereditary CRC are characterized by a very early onset (27-46 years for MSH2 and MLH1-associated and 54-63 years for MSH6-associated CRC). The risk of CRC by age 70 is 27-74% in males and 22-53% in females for MSH2/MLH1 mutation carriers, 10-22% for MSH6, while only 5,5% in general population (Giardiello et al., 2014). The overall lifetime risk of CRC for patients with LS was estimated to be 80% (Vasen et al., 1996).

Mutations in MLH1 and MSH2 are located in all regions of these genes, without obvious hot spots (Peltomaki et al., 1997). Mutation-detection strategies must therefore cover the entirety of these genes. Large genomic rearrangements are common in MSH2 (10-20% of all mutations), less common in MLH1 and are thought to be rare in MSH6 (Gille et al., 2003).
Ligation-dependent Probe Amplification) for detecting large rearrangements in MMR genes was performed using the P003, P072 or P248 kits from MRC Holland performed by the MSI Analysis System™, 4 μm obtained by microtome on FFPE tissue sections of 3-
done using BigDye™ Terminator and purified using the BigDye™XTerminator, analyzed using the ME-043 kit,
VariantReporter™, Genemapper™ (Life Technologies on the ABI 3500 Genetic Analyser, Applied Biosystems of an amplicon containing the codon 600. Capillary electrophoresis for sequencing and fragment analysis was performed
MSH2, MLH1 and MSH6 genes (Bedeir et al., 2011), exons and exon-intron boundaries. This is a huge work to do, as the genes are very large (MLH1-19 exons on 100 kb genomic DNA, MSH2-16 exons on 73 kb, MSH6-10 exons on 24 kb) (Cunningham et al., 2001). The difficult task of interpretation is hardened by benign SNPs and unclassified sequence variants – UVs) (Perrez-Carbonero et al., 2013). LS is traditionally defined based on family history and age of onset (Umar et al., 2004). However, MMR mutations are detected in only 60% of criteria fulfilling families (Lynch et al., 2003). Moreover, although less frequent, germline mutations in MMR genes are present in up to 20% of families not fulfilling these criteria and which are therefore excluded from genetic counselling (Genuardi et al., 1998).

Taking in account local CRC casuistry, familial history, age at onset, as well as available technologies for somatic and germline investigations, we propose a locally adapted protocol for investigating genetic factors in CRC sporadic and familial patients. Our model is intended to improve efficiency, fastness, accuracy, and cost-effectiveness in CRC testing.

PATIENTS AND METHODS

Patients were recruited among CRC cancer casuistry of the Regional Oncology Institute of Iasi, Sf. Spiridon and CFR Hospitals in Iasi, as well as from Lynch families addressed to the Oncogenetics Department, University of Medicine and Pharmacy of Iași. All patients agreed to participate to this study by signing an informed consent, approved by the local Ethical Committee, UMF Iasi. Collected data included family history of neoplasms, age at diagnosis, histopathological type, risk factors linked to lifestyle, populational data. DNA extraction for germline molecular analysis (DNA sequencing, MLPA) was performed from peripheral blood, using the Wizard™ Genomic DNA Purification Kit, Promega, by adapted protocol. DNA extraction for somatic analysis (MSI, PHM, BRAF) was performed on FFPE tissue sections of 5-10μm obtained by microtome, using the ReliaPrep™ FFPE gDNA Miniprep System, Promega, according to producer’s protocol. DNA concentration was evaluated by spectrophotometry and aliquots were frozen at -20°C.

Sanger dyeoxy sequencing was performed on amplicons covering the whole coding sequence of MSH2, MLH1 and MSH6, including junctions with introns. PCR reactions were tested by agarose gel electrophoresis and UV visualization after ethidium bromide staining. PCR products were purified using ExoSAP-IT™, Affymetrix. Sequencing reactions were done using BigDye™ Terminator and purified using the BigDye™XTerminator, Life Technologies. MLPA (Multiplex Ligation-dependent Probe Amplification) for detecting large rearrangements in MMR genes was performed using the P003, P072 or P248 kits from MRC Holland. Immunohistochemistry (IHC) analysis of MMR protein expression was performed on FFPE tissue sections of 3-4 μm obtained by microtome. Microsatellite instability (MSI) analysis is intended to be performed by the MSI Analysis System™, Promega. Promoter hypermethylation (PHM) of MLH-1 is intended to be analyzed using the ME-043 kit, MRC Holland. BRAF gene mutation V600E will be identified by somatic DNA sequencing of an ampicon containing the codon 600. Capillary electrophoresis for sequencing and fragment analysis was performed on the ABI 3500 Genetic Analyser, Applied Biosystems. Raw data were exported and interpreted with the VariantReporter™, Genemapper™ (Life Technologies) and SeqMan™ (DNA Star Lasergene) softwares.

RESULTS

The medical oncogenetic approach is based on molecular diagnostic, which is extremely expensive and time consuming, therefore limiting the access of all patients needing it. Although molecular diagnostic by entire gene DNA sequencing is addressed only to patients fulfilling strict inclusion criteria, MMR mutations are detected in only 60% of criteria fulfilling families, while germline mutations in MMR genes are present in up to 20% of families not fulfilling these criteria and which are therefore excluded from genetic counselling. Moreover, MMR germline mutations predispose to early-onset CRC in absence of family history, considering that early-onset CRC has increasing incidence and represents a not negligible target group. Besides family history, tumoral somatic investigations are the only possible filter for excluding non-germline events, therefore avoiding unnecessary complex gene sequencing.

There is no consensus worldwide regarding a definitive model of molecular diagnostic. The methodology varies in different oncogenetic centers, and there is an additionally variation linked
to distribution, proportion and penetrance of the MMR mutations in different populations. Therefore, local adapted detection models are more efficient and reduce the cost of investigations. We proceeded to develop a locally adapted model for CCR casuistry, following three main steps:

1. Identification and recruitment of three batches of patients: familial Lynch syndrome, early-onset CCR, sporadic unselected CCR. This step allows an evaluation of local casuistry and of biological samples / clinical data availability;
2. Evaluation of accuracy and cost-effectiveness of each technological approach for molecular analysis, in order to define the best technical combination for each batch;
3. Framing a local algorithm adapted to each batch, for a better efficiency and inclusion coverage.

**Identification and recruitment of three batches of patients**

We aimed the recruitment of three lots of patients: (1) according to Amsterdam criteria for HNPCC; (2) with early-onset CRC (<50 years); (3) with sporadic CRC.

In 1990, Amsterdam Criteria were established for HNPCC (Vasen et al., 1991), based on a strong family history of the disease and the early onset age. Amsterdam criteria referred to: ≥ 3 relatives with CRC, 1 being first-degree relative of the other 2; CRC involving ≥ 2 generations; ≥ 1 CRC diagnosed <50 years. It is estimated that Amsterdam criteria have a 22% sensitivity and 98% specificity for diagnosis of LS. Another simple 3-question CRC risk assessment tool was developed and validated (Kastrinos et al., 2009), able to identify 95% of individuals with germline mutations with 77% sensitivity. Finally, a series of clinical computational prediction models exist to determine an individual’s risk for LS, including MMRpredict, MMRpro, and the PREMM1,2,6 models (Umar et al., 2004). However, utilization of clinical criteria and modeling to identify patients with LS has been criticized for less than optimal sensitivity and efficiency (Julie et al., 2008; Perez-Carbonell et al., 2012). Patients justifying MMR analysis are not included for the test, while the test is applied on other patients which do not present *a posteriori* MMR mutations. In the future, germline testing might be the most cost-effective universal testing approach. In the meantime, filtering testing on tumor tissue can be done on archived formalin-fixed tissue (FFPE), in order to evaluate with higher precision and sensitivity either a LS or sporadic CCR.

We managed to identify so far 10 families fulfilling Amsterdam criteria for HNPCC. An example of HNPCC family is presented in Figure 1. One can observe in this family all requirement for Lynch syndrome inclusion according to Amsterdam criteria: (1) At least three CCR cases in the same family line; (2) At least 2 generations involved; (3) At least one of these cases developed before age 50.

---

**Figure 1. Example of HNPCC family fulfilling Amsterdam criteria.**
We further identified and recruited 53 sporadic CCR cases, in a consecutive series unselected for familial history or age at diagnostic. Collected data included family history of neoplasms, age at diagnosis, histopathological type, risk factors linked to lifestyle, populational data. 43 of these patients (81%) come from urban environment, while 10 (19%) from rural environment. Sex distribution was relative equivalent, with 57% male and 43% female patients. The diagnostic within this batch was of caecum neoplasm for 2 patients, ascendant colon for 11, descendent colon for 1, rectum for 16, recto-sigmoidian for 7, sigmoid for 15, and transverse colon for 1 CRC patient.

Interestingly, the age at diagnostic (Figure 2) showed non-uniform distribution in our batch. Less than 2% of the patients developed CCR after age 70, and only 13% were diagnosed in the sixty decade. The vast majority of cases were diagnosed before age 60, with 28% between 50 and 60 years, over 43% between 40 and 50, and even 13% before age 40. Therefore, we defined an early-onset batch based on diagnostic before age 50, composed of 30 out of 53 patients (56%).

Regarding the familial history of cancer (Figure 2), while 30% of our patients had no familial history, 23 of them (44%) declared 1 or 2 relatives with different neoplasms, and 11 of the cases (almost 27%) declared more than 3 relatives with cancer. Therefore, we defined those as a separate batch of CRC cases with family history of cancer.

Nonetheless, 6 out of the 30 early-onset patients declared to have at least 3 relatives with neoplastic pathologies. Although not fulfilling completely Amsterdam criteria for Lynch syndrome, this group of patients needs a particular attention and separate investigation algorithm comparing to completely sporadic CRC cases.

![Figure 2. Distribution of CRC casuistry according to the age at diagnostic (left) and familial history of cancer (right).](image)

### Evaluation of accuracy and cost-effectiveness of each technological approach

Immunohistochemistry (IHC) is targeting the loss of expression of MMR proteins, utilizing antibodies to MLH1, MSH2, and MSH6. Deleterious alterations (either germline or somatic) in specific DNA MMR are indicated by loss or partial production of the MMR protein. MSH2 and MSH6 proteins are often lost concurrently and indicate MSH2 mutation. Loss of MLH1 function can either be due to germline mutation or somatic silencing of the MLH1 gene. Therefore, IHC on MMR can identify tumors susceptible to carry germ-line MMR mutations. Figure 4 presents an example of IHC staining for MMR proteins.

![Figure 4. Examples of IHC positive staining for MLH1 (left), MSH2 (center) and MSH6 (right).](image)
Microsatellites (MS) are simple repetitive DNA sequences, liable for errors during DNA replication, resulting in insertion/deletion loops. MMR genes play a critical role in the identification and correction of these errors. Failure of the MMR apparatus leads to persistence of errors and an alteration in the length of a microsatellite sequence, a process described as microsatellite instability (MSI). MSI is a phenomenon manifested by ubiquitous MS mutations found in the tumor DNA (but not in the non-tumoral DNA) of individuals with MMR gene mutations. MSI is found in most (90%) CRC in patients with LS (due to germline MMR mutation) and in 12% of patients with sporadic CRC. Most CRCs in LS are MSI-high (>30% of markers are unstable). Investigation of MSI is a reliable method to identify defects in MMR and therefore it is a strong pre-analytical tool to select samples to be MMR sequenced.

Aberrant MLH1 gene promoter methylation (PHM) is a somatic event that is confined to the CRC and is rarely inherited. It is responsible for causing loss of MLH1 protein expression and results in MSI found in approximately 12% of sporadic cancers. Since it is a somatic event inactivating one MMR protein, it could be considered as exclusion for other MMR gene germ-line mutations, and therefore a filter for molecular analysis.

Somatic mutations in the BRAF gene, largely at codon 600, are noted in 15% of sporadic CRCs and up to 50% of sporadic MSI-CRC due to PHM. Somatic BRAF V600E (c.1709T>A) mutation have been detected predominantly in sporadic CRC. Consequently, the presence of a BRAF mutation in an MSI-high CRC is usually, but not always, evidence against LS.

Overall, tumor IHC analysis MMR proteins, MSI testing, MLH1 PHM and BRAF mutations analysis should provide a highly sensitive strategy for pre-identifying MMR-carrying mutation samples. Tumor-based approaches for early-onset CRC cancer patients for MMR gene mutation testing, irrespective of family history, appear to be an efficient screening strategy for HNPCC. However, the relationships between all markers above, family history, and germline MMR gene mutation status have not been studied on a population basis. Several guidelines exist but they are generally controversial and should definitely be adapted on a local casuistry.

Framing a locally adapted algorithm

We didn’t aim to investigate independently the same parameters in all 3 patient study batches, but more to correlate clinical, familial, histological and genetic data, for germ-line as well as for somatic level. Therefore, it is essential to collect paired data, for those patients for whom are available simultaneously clinical and familial information, blood samples for germline analysis, and FFPE tumor blocks for somatic analysis. We set up our work based on the lack in Romania of a complete CCR casuistry registry, and identified as main difficulty the data traceability and follow-up, prospectively and retrospectively, in a two-way information flow.

For the patients identified directly, all data are available together with blood samples, after signing of informed consent. For those cases, the main difficulty showed to be the identification and obtaining of tumoral blocks. On the other side, when patients are identified from anatomopathological archives, FFPE blocks are available, but it becomes challenging to retrospectively contact the patients for clinical data and blood samples.

From our knowledge, there isn’t so far in Romania a fluent and ergonomic system for simultaneously obtaining, from the same patient, all data and samples mentioned above, and there is an important lack, even at a regional level, of a complete database paired with germline and tumoral biobank. Our work is the first example in North-Eastern Romanian CCR casuistry of a complete biobank doubled with all necessary clinical and familial data. Therefore, in order to answer such a major challenge, we imagined an integrated circuit for tracing and targeting all data and samples. Our model, presented in Figure 5, is already implemented and functional.
The description above of our patient batches shows the possibility of investigating by molecular analysis not only families fulfilling Amsterdam inclusion criteria, but also patients recruited for early-onset diagnostic, important family history of cancer, or even both. An important proportion of patients unselected specifically for age or history showed however an early-onset diagnostic (before age 50), and/or family cancer history not sufficient for including in Amsterdam criteria, but not negligible at all from a molecular investigation point of view. The arising question is what kind of investigations should be performed for each type of group.

Usually, MMR genetic testing by complete gene sequencing is addressed exclusively to Amsterdam criteria patients. However, it should be highlighted that MMR mutations are detected in only 60% of Amsterdam criteria fulfilling families, while germline mutations in MMR genes are present in up to 20% of families not fulfilling these criteria and which are therefore excluded from genetic counseling. Beyond optimizing and implementing each technology, our principal aim was to elaborate an investigating algorithm, adapted to each batch of patients, depending on overall data and samples available. If complete gene sequencing is inappropriate for all CCR casuistry, the different methods presented allow filtering patients for sequencing, but also partial and targeted investigations for each patient. We propose below (Table I) a complete model for CCR casuistry coverage by adapted partial technology algorithms. By applying the adapted protocol to each patient batch, we estimate an important increase in diagnostic efficiency, as well as in investigation coverage.

<table>
<thead>
<tr>
<th>Investigation technique</th>
<th>Addressability</th>
<th>Final estimated batch</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complete MMR analysis by DNA sequencing</td>
<td>Amsterdam + familial history early-onset CCRs</td>
<td>10 HNPCC + 10 non-HNPCC</td>
</tr>
<tr>
<td>Partial MMR analysis by DNA sequencing (known mutations)</td>
<td>No familial history early-onset CCRs; Familial history normal-onset CCRs</td>
<td>50 early-onset + 20 familial history CCRs</td>
</tr>
<tr>
<td>MMR MLPA analysis for large genomic rearrangements</td>
<td>Amsterdam + familial history early-onset CCRs</td>
<td>10 HNPCC + 10 non-HNPCC</td>
</tr>
<tr>
<td>MMR protein analysis by immunohistochemistry (IHC)</td>
<td>Amsterdam + early-onset CCRs With or without familial history</td>
<td>10 HNPCC + 50 early-onset</td>
</tr>
<tr>
<td>Microsatellite instability (MSI) analysis</td>
<td>Amsterdam + early-onset CCRs With or without familial history</td>
<td>10 HNPCC + 50 early-onset</td>
</tr>
<tr>
<td>Promoter hypermethylation (PHM) analysis on MLH1</td>
<td>Amsterdam + early-onset CCRs With or without familial history</td>
<td>10 HNPCC + 50 early-onset</td>
</tr>
<tr>
<td>BRAF mutations tumor analysis</td>
<td>All CCRs</td>
<td>All CCRs (100)</td>
</tr>
</tbody>
</table>

Table I. Establishment of the best working protocol for each patient or batch.
DISCUSSION AND CONCLUSIONS

Detection of the germ-line defect in HNPCC families allows identification of relatives who require appropriate surveillance and prevents useless surveillance in non-carrier relatives. Of all cancer risk factors, the hereditary factor is the only one to reach the positive or negative predictive value threshold in order to justify a medical follow-up. Oncogenetics is focusing either early diagnostic of the disease, prevention and prophylactic surgery. Patients with LS are at increased risk for the development of colorectal and extracolonic cancers at early ages. Screening for CRC by colonoscopy is recommended in persons at risk or affected with LS every 1 to 2 years, beginning at ages 20-25. Screening for endometrial, ovarian, gastric cancer should be offered for patients affected with LS. Hysterectomy and bilateral salpingooophorectomy should be recommended to women with LS who have finished childbearing or at age 40 years.

In this context, our purpose was to develop and implement an efficient diagnostic tool in order to reduce the costs of molecular investigations, to decrease the waste of inappropriate investigations and to enlarge the coverage of patients to benefit of the genetic testing. We identified two major problems in molecular oncogenetic diagnostic of CRC:

1. Genetic testing is proposed to the patients according to inclusion criteria (Amsterdam and Bethesda criteria, see above) and using computational models, but unfortunately these models often underestimate the probability of finding a mutation, are validated only in some countries and no universal consensus exist regarding their common use. Moreover, familial history is also absent or unknown in some mutation positive families. Testing is not available to all patients that need it.

2. MMR mutations are detected in about 60% of criteria fulfilling families, which mean that fully 40% of families that meet the criteria do not have LS. From an economic and efficiency point of view, this means 40% waste of the molecular diagnostic and it is rising the question of filtering the patients to test in order to rise the proportion of mutations identification.

Both problems are rising the question of Where limiting the access to the testing and Who should be the beneficiary of the molecular testing. One may consider that a reasonable purpose is to simultaneously reduce the 40% waste and to extend the accessibility to the testing. This can be done by generating an analytic algorithm of inclusion and testing and by optimizing it to the local condition and population particularities.

When used appropriately, genetic testing for LS can confirm the diagnosis at the molecular level, justify surveillance of at-risk persons, decrease the cost of surveillance by risk stratification, aid in surgical and chemoprevention management, and help in decisions concerning family and career planning. However, when used inappropriately, genetic testing can misinform affected patients with false-negative results and waste patient and societal resources. Our work has an important impact on: (1) reducing the cost and time of molecular diagnostic; (2) include for the diagnostic more patients than usually selected by inclusion criteria; (3) perform a more efficient diagnostic with >80% mutation finding probability; (4) provide “best practice” evidence-based recommendations for oncogenetic diagnostic.

REFERENCES


1) University of Medicine and Pharmacy « Gr.T.Popa », Iaşi, Romania
2) CFR Hospital, Iaşi, Romania
3) Regional Institute of Oncology, Iaşi, Romania
4) University «Alexandru Ioan Cuza», Iaşi, Romania
* corresponding author: luciannegura@yahoo.fr

ACKNOWLEDGEMENT
This study was possible with financial support from the Romanian Ministry for Education and Research, by the CNCSIS research grant PN-II-RU-TE-2014-4-2257, Contract 203/01/10/2015, UEFISCDI.
We kindly thank Center Jean Perrin in Clermont-Ferrand, France, for sharing Gataca software intended to draw familial anamnesis of HNPCC patients.