Purpose:
To determine if there is a significant difference between polyp rates of patients with the EPCAM deletion and patients with mismatch repair mutations of Lynch syndrome.

Background:
The 2009 discovery of the EPCAM deletion led to an addition to Lynch syndrome, traditionally characterized by mutations of mismatch repair mutations (MMR) MLH1, MSH2, MSH6 and PMS2. Patients positive for the EPCAM deletion show some of the phenotype of Lynch syndrome with the MSH2 mismatch repair mutation, but do not show a detectable mutation in MSH2. Patients with the EPCAM deletion also develop fewer extracolonic cancers than patients with MSH2 mismatch repair mutations. It was hypothesized that since EPCAM deletions are much more site-specific to the colon, there may be an increased development of colonic polyps compared to patients with other MMR mutations.

Methods:
Over 90 families with MMR mutations and one family with EPCAM mutation are recorded in the database at Creighton University Hereditary Cancer Center. Of these 90 families, over 400 individuals have tested positive for a MMR mutation and 24 have tested positive for an EPCAM mutation. Over 350 letters requesting HIPAA release to access past colonoscopies, flexible sigmoidoscopies, and related pathology reports were sent out, with over 70 were returned. Pertinent data including dates of procedure, pathology, and location of polyps was recorded.

Results:
No statistical difference was found between polyp rates of each group. However, there was a p-value for EPCAM vs. MSH2 of .08, suggesting that a greater lifetime polyp rate in patients with EPCAM deletion should be considered in future studies.

Conclusions:
From the data collected, no statistically significant conclusions can be made about lifetime polyp rates between patients with the EPCAM deletion and patients with MMR mutations. It is hoped that this study will be a pilot study to be used for further investigation of the EPCAM mutation and Lynch syndrome.
Reflections on the Research Process by the First Author

Every step, from recruiting mentors, seeking funding, data collection, and manuscript preparation was fraught with its unique set of unforeseen obstacles. Because of the struggles, however, we gained an appreciation for the importance of rigorous methodology, careful calculation, and asking the right kinds of questions. Overall, it is very satisfying to see a research project go from an idea to a manuscript. This collaborative experience between students, teachers, researchers, and mentors—all with diverse backgrounds—has boosted our respect for the research process and our passion for culturally competent patient care.

In my senior year of college, I was fortunate enough to start a summer internship in the pathology department at Mercy Medical Center under Dr. Shirin Nash. As part of my final project for the internship, I wrote a paper on Lynch Syndrome. One year later, I began school at Creighton University School of Medicine where Dr. Henry Lynch established his career.

During my first year of medical school, I was interested in pursuing a summer research program. When scanning the list of physicians who offered to mentor medical students, I saw Dr. Lynch’s name and jumped at the opportunity. This was a unique chance to work with one of the most well known physicians in hereditary cancers.

This research project was challenging yet rewarding throughout the entire experience. One of the most difficult challenges was incomplete records of colonoscopy reports. Throughout the summer, I made many phone calls and faxes to hospitals for release of medical records in attempts to fill the gaps of colonoscopy reports.

Something else I had to consider was how to keep all of my data organized. We sent out over 350 letters asking for release of medical records. Every day throughout the summer, mail would come in containing signed medical releases or medical reports. I kept track not only of which patient’s medical releases came in, but also which years they were from and which years they may be missing given their first or last colonoscopy. I learned about organizing significantly large volumes of data and about the entire retrospective research process. Although our data was not statistically significant, there was definitely a trend to look into given our p-value of .08 in comparing lifetime polyp rates between EPCAM patients and traditional Lynch syndrome patients. Looking back on all of my hard work that summer and seeing that my study may spark other research that will benefit patient care and management, made it all a worthwhile experience.

Introduction

Lynch syndrome is an autosomal dominant genetic disorder which predisposes individuals to colorectal cancer and other types of extracolonic cancers such as ovarian and endometrial cancer. It is the result of mutations in the mismatch repair (MMR) genes MLH1, MSH2, MSH6 or PMS2. In the United States, approximately 50% of Lynch syndrome cases are a result of MLH1 mutation and approximately 40% of are a result of MSH2 mutation. Microsatellite instability, a characteristic molecular marker of Lynch syndrome, is found in over 90% of the cases.1

Recently, it was discovered that some families may exhibit the phenotype of Lynch syndrome, in addition to a silenced MSH2 gene, but will not show any mutations in the MSH2 gene. Deletion of the 3’ end of the epithelial cell adhesion molecule (EPCAM) gene, which is immediately upstream to MSH2, causes transcription read-through (transcription beyond the normal termination signal) and consequential silencing of MSH2 via hypermethylation of the promoter.1

The risk of developing colorectal cancer in patients who are positive for the EPCAM deletion is comparable to Lynch syndrome patients who are carriers of the MSH2 mutation. A 2011 study examined 93 individuals with EPCAM and compared the age of onset and cumulative risk by age 70 for colorectal cancer to Lynch syndrome patients with MLH1, MSH2, and MSH6 mutations. This study found that individuals positive for the EPCAM deletion did not differ from MLH1 or MSH2 in terms of cumulative risk or age of onset, but did have a greater cumulative risk and earlier age of onset of colorectal cancer than MSH6 mutations.2 The average age at diagnosis of colorectal cancer for EPCAM deletion carriers was 43 years old,2 which is fairly consistent compared to other patients with Lynch syndrome.
However, *EPCAM* patients may differ from the traditional Lynch syndrome cases in that relatively fewer extracolonic cancers have been reported. In particular, endometrial cancer is quite rare and the cumulative risk is significantly lower.\(^2\)

Certain extracolonic cancers such as endometrial, ovarian, and stomach cancer are not as common in *EPCAM* patients. However, *EPCAM* is not totally devoid of extracolonic cancers, as there is a relatively higher incidence of pancreatic and duodenal cancers among *EPCAM* patients.\(^2\) Patients with *EPCAM* deletions have a relatively high level of *EPCAM* expression in colorectal stem cells.\(^3,4\) There is a suspicion that since the *EPCAM* deletions appear to be more site-specific to the colon that they may play a role in an increased development of colonic polyps.

Considering that *EPCAM* deletions may be more site-specific to the colon, we hypothesized that there would be an increased rate of development for colonic polyps compared to patients with other MMR mutations.

**Methods**

This study was a retrospective observational study, using medical records and pathology reports released by individuals in Creighton University’s Heredity Cancer Center database. In May 2013, 350 letters were sent out to individuals in the database who tested positive for *MSH2* and *MLH1* MMR mutations, and *EPCAM* deletions. Inclusion criteria for this study included family members who tested positive for germ-line, pathogenic mutations in MMR mutations and *EPCAM* deletions, as well an age requirement of at least 19 years old. Exclusion criteria included those in the database who did not have a viable mailing address and those who were deceased. Other exclusion criteria include the circumstances of the existing medical records and failure to follow-up.

All patients signed a HIPAA release for all of colonoscopies, flexible sigmoidoscopies, and associated pathology reports. Medical records were examined to determine polyp counts from each colonoscopy and flexible sigmoidoscopy.

Statistical analysis packages R and SAS 9.3 were used to analyze the collected data. The statistical methods were used in attempt to detect a statistical difference in the lifetime polyp rate between the two groups.

The measure of interest was lifetime polyp count; however, to account for the effect of age, lifetime polyp count was converted into lifetime polyp rate (LTPR) by dividing the lifetime polyp count by the age at last colonoscopy. Two approaches to the data analysis were

<table>
<thead>
<tr>
<th></th>
<th>EPCAM</th>
<th>MLH1</th>
<th>MSH2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Participants</td>
<td>10</td>
<td>28</td>
<td>26</td>
</tr>
<tr>
<td>Male</td>
<td>5</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>Female</td>
<td>5</td>
<td>20</td>
<td>17</td>
</tr>
<tr>
<td>Unique families</td>
<td>1</td>
<td>17</td>
<td>12</td>
</tr>
<tr>
<td>Total number of recorded colonoscopies</td>
<td>74</td>
<td>247</td>
<td>166</td>
</tr>
<tr>
<td>Average number of recorded colonoscopies per patient</td>
<td>7.4</td>
<td>8.8</td>
<td>6.4</td>
</tr>
<tr>
<td>Average span of recorded colonoscopy screening (years)</td>
<td>11.0</td>
<td>17.1</td>
<td>13.9</td>
</tr>
<tr>
<td>Average age at first colonoscopy</td>
<td>41.9</td>
<td>42.1</td>
<td>45.4</td>
</tr>
<tr>
<td>Total number of polyps</td>
<td>60</td>
<td>167</td>
<td>86</td>
</tr>
<tr>
<td>Average number of lifetime acolic polyps discovered</td>
<td>6.0</td>
<td>6.0</td>
<td>3.3</td>
</tr>
<tr>
<td>Developed colorectal cancer (n)</td>
<td>2 (20%)</td>
<td>10 (36%)</td>
<td>5 (19%)</td>
</tr>
<tr>
<td>Average age at cancer diagnosis</td>
<td>34.0</td>
<td>42.5</td>
<td>43.0</td>
</tr>
</tbody>
</table>

Table 1. Patient Characteristics
taken with regard to detecting differences between the \textit{EPCAM}, \textit{MSH2}, and \textit{MLH1} mutations. First, since the three mutations are hereditary, subjects from the same biological family were correlated. To account for this unknown correlation, the lifetime polyp rates were averaged for all members of a family, thus yielding a single observation per family. This yielded a single observation for \textit{EPCAM}; 17 observations for \textit{MLH1}; and 12 observations for \textit{MSH2}. Having only one observation from the \textit{EPCAM} family necessitated the use of a randomization test. We randomized averaged lifetime polyp rates for families, as opposed to individuals.

Our second method for detecting differences between the \textit{EPCAM}, \textit{MSH2}, and \textit{MLH1} mutation was to treat each patient as an independent observation. In contrast to the above randomization, which used the means of entire families to compare, this second method compared individuals. This was done utilizing median one-way analysis, a nonparametric test based on the relative ranks of the LPR between groups.

Results

When comparing the empirical distribution for LTPR for individuals with either \textit{EPCAM}, \textit{MLH1}, or \textit{MSH2} mutation, the proportion of individuals with a given mutation increases with LTPR. However, the difference between polyp rates between \textit{MSH2} and \textit{EPCAM} is more clearly visualized in figure 1 by examining the large gap between the two mutations as the LTPR increases from 0.1 to 0.2 polyps per year.

Approximately 60\% of the \textit{EPCAM} deletion patients have a LTPR of 0.1 or less, while almost 80\% of \textit{MSH2} mutations have a polyp rate of 0.1 or less. This trend generally continues up through polyp rates of 0.1 to 0.2 polyps per year.

Figure 2 demonstrates that \textit{EPCAM} had the highest proportion of patients who had a LTPR above the median for all three groups. The frequency represents the number of individual observations, while the shaded red area represents the frequency that LTPR for a given individual fell above the median LTPR for the

![Figure 1. Empirical Distribution for Lifetime Polyps](image)

The p-value for \textit{EPCAM} vs \textit{MLH1} = 0.216 and the p-value for \textit{EPCAM} vs \textit{MSH2} = 0.0831.
group (EPCAM deletion, MLH1 mutation, and MSH2 mutation) as a whole. When comparing the patients as individual observations using median one-way analysis rather than the randomization method there was no significant difference detected between the LTPRs, median p-value = 0.7614.

Discussion
Tracking polyp counts is important, as an increase in the number of polyps heightens the risk for colorectal cancer, given the adenoma-carcinoma sequence. From the EPCAM family, 20% patients developed colon cancer, with one of the patients developing colorectal cancer three separate times. From the families with the MLH1 mutation, 36% developed colorectal cancer, and one patient developed cancer four separate times. From the MSH2 family, 19% of patients developed cancer, with one patient developing cancer on two separate occasions.

Although there was not a statistically significant difference of LTPR between EPCAM, MLH1, and MSH2, there are certainly trends to investigate. When comparing the difference between polyp rates between EPCAM and MSH2, there was a p-value of 0.08 suggesting that LTPR among EPCAM patients may be higher. This can be seen in figure 1 by comparing the proportion of patients who have a LTPR between 0.1 and 0.2. Almost 80% of MSH2 patients have a LTPR of 0.1 or less, while only 60% of EPCAM patients have a LTPR of 0.1 or less.

Although there is a similar distribution at an LTPR of 0.2 between both groups, the empiric distribution is helpful in that it shows the large gap that runs between EPCAM and MSH2 between 0.1 and 0.2, showing that a greater distribution of MSH2 patients have a smaller LTPR.

Figure 2 shows that EPCAM patients had the greatest percentage of patients who fell above the LTPR median. This is of interest because it demonstrates the consistency of polyps among EPCAM patients despite

Figure 2. Frequencies Above and Below the Overall Medial for Lifetime Polyp Rate.
Fadus, Lifetime Colonic Polyp Rate

a smaller sample size.

There were some interesting findings among EPCAM positive individuals. One EPCAM patient had seven polyps on the first colonoscopy, and another patient was diagnosed with colon cancer at age 24 on first colonoscopy. Patients in the EPCAM group who did develop cancer developed it at a much younger age (34.0) than MLH1 (42.5) and MSH2 (43.0), despite having a similar age at first recorded colonoscopy.

Among the 10 patients positive for the EPCAM mutation, all of the patients had at least one polyp in their lifetime. This includes several younger patients in the EPCAM family that only had one colonoscopy completed at the time of this study. There were six MMR Lynch syndrome patients who had no polyps found at all, despite many recorded colonoscopies. Our findings show a trend in polyp rates, and at this moment, should not guide any management in colorectal cancer surveillance. We hope that future polyp studies such as our own will more accurately assess the need for aggressive colonic surveillance in Lynch syndrome and other polyposis syndromes.

Research and findings in EPCAM deletions and the mismatch repair enzymes involved in Lynch syndrome are important. It has been found that there are certain risks involved with each specific mutated gene in Lynch syndrome. For example, studies have shown that MSH2 mutation carriers develop extracolonic cancers more frequently than carrier of MLH1 mutation. Patients with mutations in MLH1 have higher rates of colorectal cancer, and patients with a mutation in MSH6 gene have an increased risk of endometrial cancer but a decreased risk of colorectal cancer.

Patients with EPCAM deletions have a more site-specific risk of colorectal cancer because of a relatively high level of EPCAM expression in colorectal stem cells. The literature suggests that there are trends in the tumor spread of Lynch syndrome patients that may depend on the type of mutation leading to the development of Lynch Syndrome. Management and prevention strategies could be maximally implemented by tailoring them according to the mutation type in a patient with Lynch Syndrome. The 2009 discovery of the EPCAM deletion is important because it allows for more reasonable preventive strategies to be used compared to other MMR mutations. For example, female patients with the EPCAM deletion will no longer undergo unnecessary prophylactic hysterectomies and bilateral salpingooophorectomies as a result of their family history of colorectal cancers. Instead, these patients with EPCAM deletion can focus almost exclusively on colorectal cancer. Although it has been found that there are generally fewer extracolonic cancers in EPCAM deletions, there have been increased rates of pancreatic and duodenal cancers, therefore, it is not possible to only focus on colorectal cancer risk in EPCAM deletion patients. Other cancer locations are possible, but the enhanced risk of colorectal cancer makes screening essential in this patient population.

An important trend in hereditary cancers, especially Lynch syndrome, is that the surveillance and prophylactic procedures become tailored to the risks of certain mutations. Our study reveals that there may be a difference in LTPR between patients with EPCAM deletion and patients with MMR mutations. LTPR might be a viable surveillance mechanism to screen EPCAM deletion patients for colorectal cancer. By analyzing the pedigrees and tumor spectrums of families who are carriers for gene mutations and other inherited forms of colorectal cancer, the risk and prevention strategies can be assessed on an individual basis. Not only would this be in the best interest of the patients, but it would also prevent seemingly unnecessary diagnostic tests thus saving valuable healthcare resources.

Limitations
The nature of the study was a retrospective observational study, and some records, especially records before 1980, were incomplete. The most limiting factor in the study was the EPCAM positive sample size (n=10). There were 24 EPCAM positive individuals available to request records, and 10 responses returned. The number of EPCAM positive deletion individuals was limited because several of the individuals were too young to be considered for colonoscopy screening (less than 20 years old). The data for some of the EPCAM positive patients was also limited to one colonoscopy in three of the records because these patients were also particularly young.

Another limitation for this report was that the
data was only recorded from one institution and one particular family in the United States. Although there are several known families of the EPCAM deletion throughout the world, the fairly recent discovery of EPCAM limited the resources available.

The mutation spread in the Creighton University Hereditary Cancer Center’s databank was not a complete representation of Lynch syndrome mutations. Although PMS2 and MSH6 account for 10% of Lynch syndrome mutations (MLH1 and MSH2 account for 90%), there were no MSH6 or PMS2 positive mutations present among the 350 positive individuals in the database. MSH6 traditionally has a lower rate of colorectal cancer than the other MMR mutations and EPCAM deletions, so a lack of MSH6 mutations in the data could lead to an overestimation of the rate of colorectal cancer in the MMR mutations compared to EPCAM deletions.

Another limitation in this study was the change of screening technology that was used in patients over the length of the study. Lynch syndrome was discovered before the advent of colonoscopies (flexible sigmoidoscopies was utilized) so some older patients were limited in that they only had a portion of the colon viewed during their earliest screening procedures.

A final and very important limitation of the statistical analysis was the use of surgical procedures that are typically used to treat colorectal cancers. Treatment often involves removing the portion of the affected colon, decreasing the length of the colon, and thereby decreasing the length of colon where polyps can develop. This is important to note, but despite the differences in colon length, all patients were weighted the same in the statistical analysis as long as there were segments of colon that were still at risk for polyp development.

The population that was studied for both MMR mutations and EPCAM deletions was Caucasian. This was not necessarily limiting for the study itself because it compared two similar groups, but it is limiting in the ability to generalize the results of this particular study to the entire population, especially since rates of colorectal cancer and Lynch syndrome vary among population groups.

References


