AD_____

AWARD NUMBER: W81XWH-17-1-0556

TITLE: "Utilizing Clinical Metadata to Predict High-Cost Complications and Treatment Response in IBD: Development of Clinical Decision Support Tools"

PRINCIPAL INVESTIGATOR: David G. Binion

CONTRACTING ORGANIZATION: University of Pittsburgh Office of Research

Pittsburgh, PA 15213-2303

RECIPIENT: Catherine C. Henry

REPORT DATE: September 2019

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
Sept 2019	Annual	1 Sep 2018 - 31 Aug 2019

4. TITLE AND SUBTITLE

"Utilizing Clinical Metadata to Predict High-Cost Complications and Treatment Response in IBD: Development of Clinical Decision Support Tools"

5a. CONTRACT NUMBER

5b. GRANT NUMBER W81XWH-17-1-0556

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

David G. Binion and Claudia Ramos Rivers

E-Mail: binion@pitt.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

University of Pittsburgh Office of
Research
123 University Place
Pittsburgh, PA 15213-2303

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

IBD is a costly and debilitating disease, significantly affecting quality of life. Our research plans is to generate easy to use, internet based tools (similar to a calculator) to determine which patient will go on to have costly disease over the next several years, and/or is unlikely to respond to traditional biologic therapies with anti-TNF medications. We propose using an already available IBD patient registry database which has been developed by the P.I. and the research team at UPMC/University of Pittsburgh.

The short term goal is to use accessible patient information and routinely collected prospective clinical data derived from the electronic medical record from over 3,000 IBD patients followed for >7 years, to generate personalized prediction models and tools to assess response to biologic therapy and risk of high costs complications, including enteric infection and disability for the care of patients with IBD. We will generate a publically accessible computer based risk prediction calculator that allows for risk stratification after entering routinely collected patient information. The goal of this web-based technology will be to use routine clinical information to facilitate a personalized clinical approach for treatment and stratification of IBD patients based on severity and phenotype.

Personalized approaches for IBD treatment will help to avoid unnecessary exposure to biologic therapies and their associated risks in patients likely to fail a standard biologic treatment (i.e. anti-TNF) approach. Similarly, identifying patients that are at risk for future high-cost complications will provide a window of opportunity for cost-saving outpatient care, proactive lifestyle modifications and dietary interventions to prevent hospitalization, surgery, infectious complications, or disability.

This personalized approach to IBD treatment will positively impact patients and their experience with disease, avoiding risks and given the opportunity for early interventions to avoid debilitating disease complications. Personalization of care will also benefit those taking care of IBD patients, as it will provide insight into disease subgroups and treatment choices, saving time and financial resources from the health system.

15. SUBJECT TERMS

Inflammatory Bowel Disease, anti-Tumor Necrosis Factor, Electronic Medical Records, Short Inflammatory Bowel Disease Questionnaire, Hemoglobin, Crohn's Disease, Ulcerative Colitis

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
					USAMRMC
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER *(include area code)*
Unclassified	Unclassified	Unclassified	Unclassified	20	

TABLE OF CONTENTS

1. **INTRODUCTION:** Narrative that briefly (one paragraph) describes the subject, purpose and scope of the research.

Using readily accessible patient demographics and routinely collected prospective clinical data harvested from the electronic medical record (EMR) from >3,000 consented IBD patients followed for >7 years, to generate personalized prediction models to assess response to anti-TNF biologic therapy and risk of high cost complications, including enteric infection and disability for the care of patients with inflammatory bowel disease (IBD). We will generate an accessible computer-based risk prediction platform that allows for risk stratification after entering routinely collected patient demographic and clinical information.

2. **KEYWORDS:** Provide a brief list of keywords (limit to 20 words).

Inflammatory Bowel Disease, anti-Tumor Necrosis Factor, Electronic Medical Records, Short Inflammatory Bowel Disease Questionnaire, Hemoglobin, Crohn's Disease, Ulcerative Colitis

3. **ACCOMPLISHMENTS:** The PI is reminded that the recipient organization is required to obtain prior written approval from the awarding agency grants official whenever there are significant changes in the project or its direction.

What were the major goals of the project?

List the major goals of the project as stated in the approved SOW. If the application listed milestones/target dates for important activities or phases of the project, identify these dates and show actual completion dates or the percentage of completion.

1. Develop a clinical decision support tool for identifying IBD patients at risk of complicated disease.
2. Develop a clinical decision support tool to identify IBD patients at risk of poor response to anti-TNF biologic therapy

What was accomplished under these goals?

For this reporting period describe: 1) major activities; 2) specific objectives; 3) significant results or key outcomes, including major findings, developments, or conclusions (both positive and negative); and/or 4) other achievements. Include a discussion of stated goals not met. Description shall include pertinent data and graphs in sufficient detail to explain any significant results achieved. A succinct description of the methodology used shall be provided. As the project progresses to completion, the emphasis in reporting in this section should shift from reporting activities to reporting accomplishments.

I. We rewrote much of the computational logic for rating predictors using Random Forest regressor and classifier from Python's sklearn machine learning library. This approach allowed us to compare procedurally-selected predictors with predictors identified with experts and through literature reviews. Through iterative testing of machine learning algorithm and through in-depth review of classification results we discovered that all tested ML classifiers were incorrectly classifying some high-charge patients as low-charge patients with high confidence. Through additional data review, we discovered that while charge data provides a good proxy for clinical outcomes, we have to take extra steps to

identify patients whose charges are higher due to treatments unrelated to IBD. After removing patients with cancer-related diagnosis and organ transplant patients, the Random Forest and XGBoost models' accuracy improved to 92%, with misclassification rates significantly reduced.
While classification-modeling approach is producing reasonable results with predicting which patients will likely have poor response to anti-TNF biologic therapy (79% accuracy), we are expanding our efforts to test causal probabilistic models to explore how treatment pathways affect response to anti-TNF biologic therapy. Initial Bayesian network (BN) models have accuracy comparable to classification models but may provide more insight into contributing factors.

II. We have applied cluster analyses to screen patients into groups with different prognosis and differential response to standard IBD therapies and anti-TNF treatments.

III. We have explored several non-modeling approaches for identifying poor clinical outcomes patients. A patient cohort and treatment pathways visualization system that integrates elements of exploratory search has shown the most promise as a viable approach to developing a useful clinical decision support system. The two current approaches include using a Sankey chart to illustrate how different clinical decisions result in different outcomes (Figure 1) and an exploratory search decision tree that illustrates immediate outcomes of every clinical decision (Figure 2).

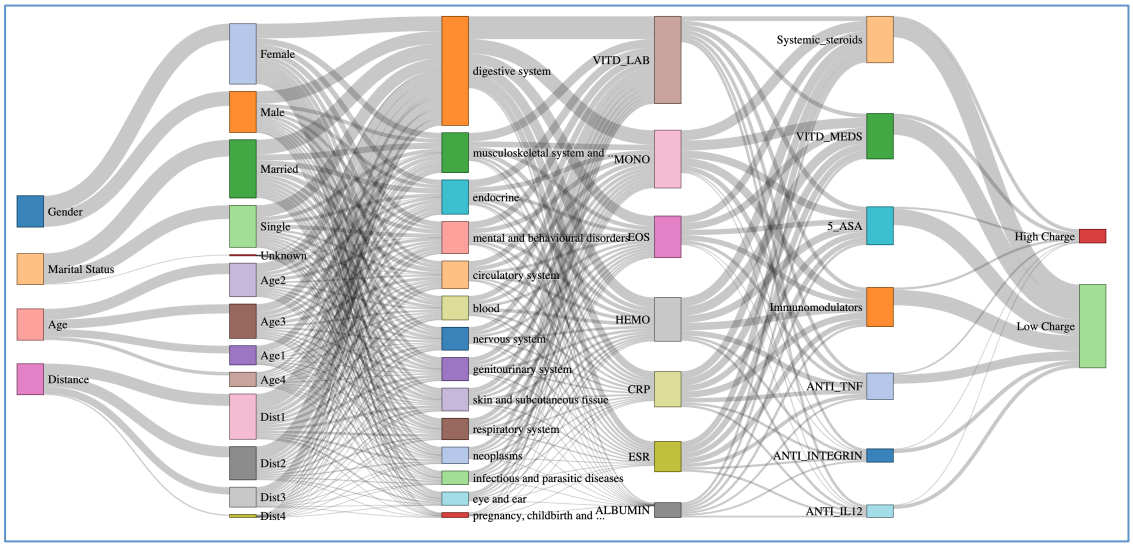

Figure 1: Sankey chart to illustrate how different clinical decisions result in different outcomes

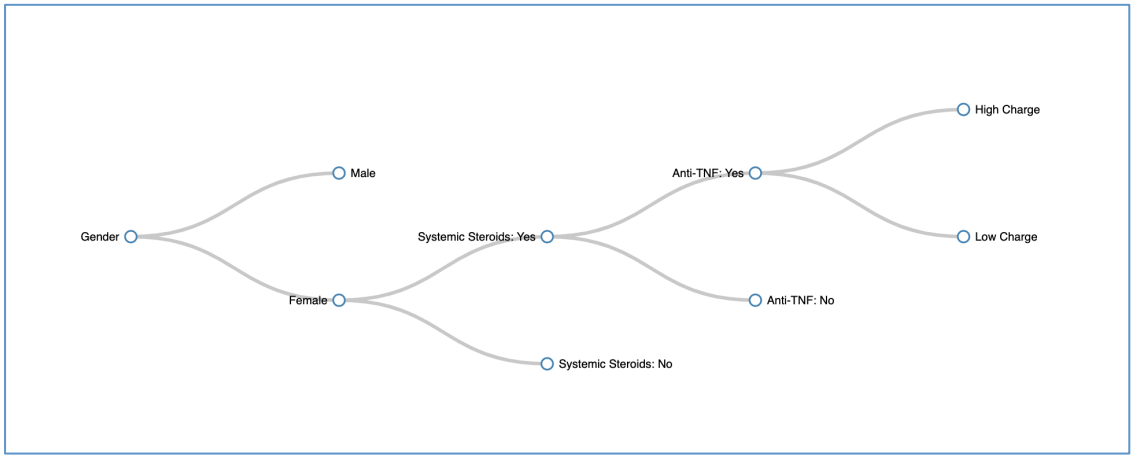

Figure 2: Decision tree-based user interface to support exploratory search of patient outcomes based on clinical decisions and demographic information

We also began developing a web-based decision support system that combines the visualization approaches described above with predictions (classifications) produced by XGBoost models. The current version of the user interface is shown in Figure 3.

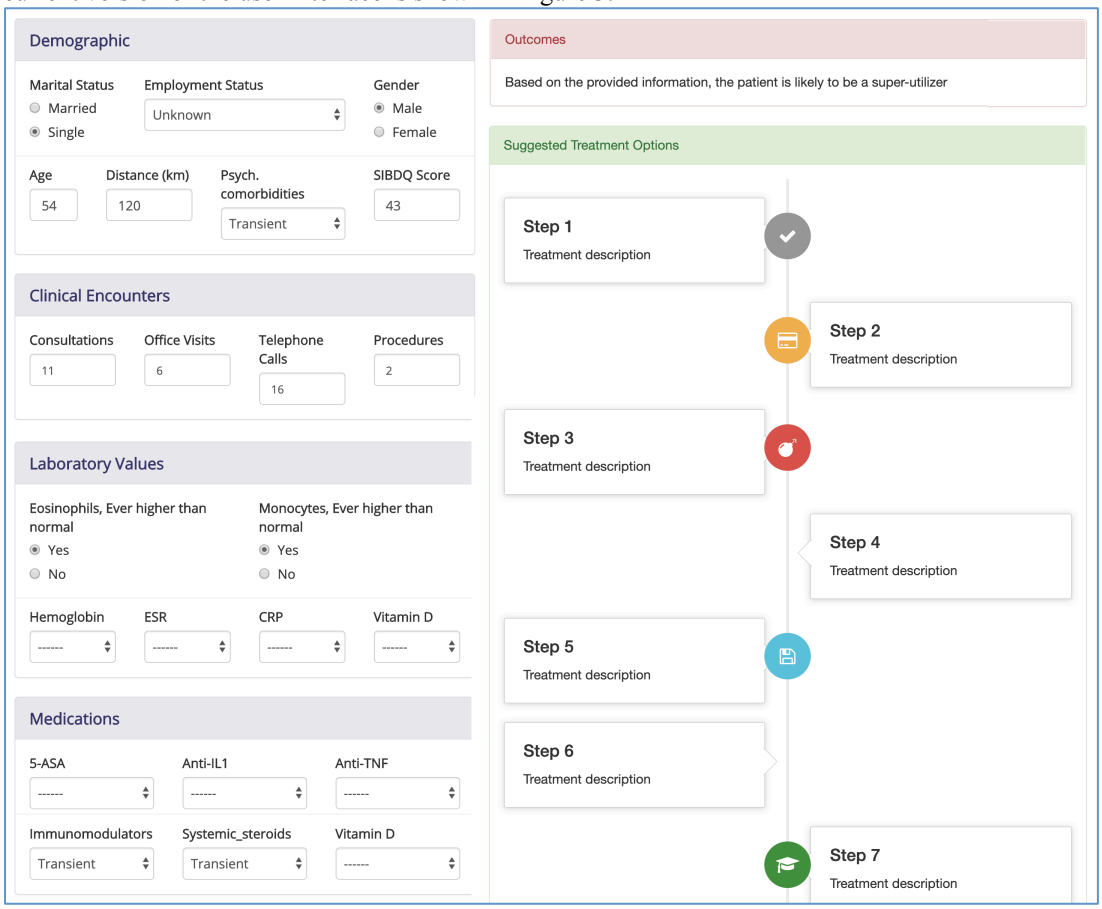

Figure 3: Proposed decision support system user interface

IV. We began working exploring alternative approaches of representing temporal data and of rapidly generating / grouping patient cohorts. In order to achieve this, we need to make significant changes to how patient data is stored and structured.

Electronic Medical Record Systems (EMRS) such as Epic and Cerner represent patient data using relational data models. In such models, each object and event relevant to a patient, including the patient themselves, is represented by a table (an entity). Examples of such entities in EMRS include diagnosis, medications, visits, allergies, procedures, etc... In turn, each table is described by a set of attributes (columns) that provide additional detail about the entity in question.

For example, a *Medication* table would contain columns such as a medication's unique identifier, brand name, generic name, therapeutic class, etc. Each row of data in such a table represents a single record (row of data) describing some part of the patient's treatment history. A row of data in the *Medication* table may look like the example shown in Table 1 and indicate that a patient with the medical record number (MRN) of 54321 was prescribed Tylenol on April 11, 2018.

Medication ID	**Medication Name**	**Generic Name**	**Therapeutic Class**	**Patient MRN**	**Order Date**
1	Tylenol	Acetaminophen	Analgesics	54321	04/11/2018

In other words, storing patient data in relational data models facilitates state-based representations, where each record in relevant database tables represents the state of a patient at the time that the data was created.

While using relational models to store patient data is a common approach that is widely used by EMRS vendors such as Epic and Cerner, relational database management systems (RDBMS) have several shortcomings when it comes to helping answer questions about temporal events, especially when these questions concern exploring and understanding causality. Another key shortcoming is the difficulty of rapidly identifying patient cohorts based on temporal events. For example, retrieving a sub-cohort of CF patients based on criteria such as "CF patients ages 12-19 with cystic fibrosis-related diabetes (CFRD) that had multiple adverse reactions to Kalydeco" from a relational database, one would need perform computationally expensive connections between Patient, Diagnoses, Medication, and Allergy tables. Once the data were retrieved, further computations to extract temporal relationships would be required in order to extract the desired patient cohort.

To address these shortcomings, we propose representing patient data as a graph, a mathematical structure and a data model often used to represent, study, and model credit card fraud patterns, power consumption patterns, and information and influence propagation in social networks.

Figure 4 shows a possible representation of patient clinical data as a graph with three types of nodes: patient (P), event (E), and doctor (D). This example shows five types of events - medication (EM), encounter (EE), procedure (EP), laboratory test (EL), and adverse / allergic reaction (EA). Directed edges (arrows) indicate temporal precedence (i.e. event 1 occured before event 2). Undirected edges (lines without arrows) indicate co-occurrence (i.e. event 1 and event 2 occured at the same time).

By examining the pattern of nodes and edges in Figure 4, we can quickly identify that patients 1 and 2 (P1 and P2) had encounters with the same doctor (D1) and that patient 2 (P2) was on two medications simultaneously and had an adverse reaction following these medication events. Representing patient clinical data as a graph affords numerous computational advantages over the relational model, including abilities to (1) quickly identify and extract patient cohorts based on similarity metrics, (2) visualize patient disease progression trajectories, (3) identify possible causal relationships in the data.

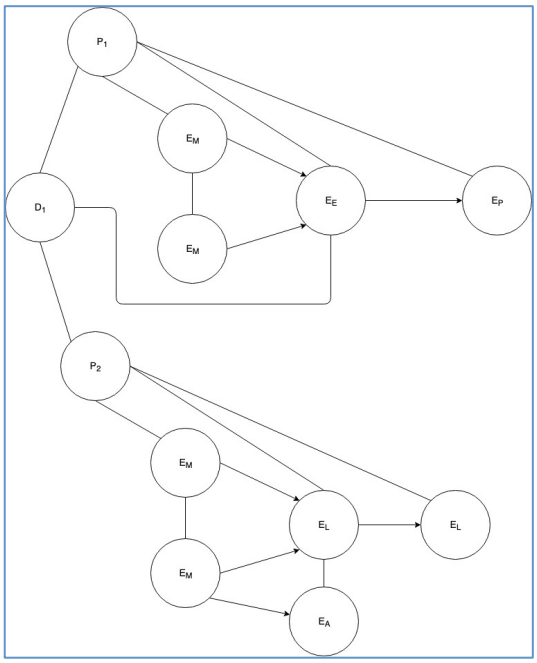

Figure 4: Example of patient data model represented as a graph

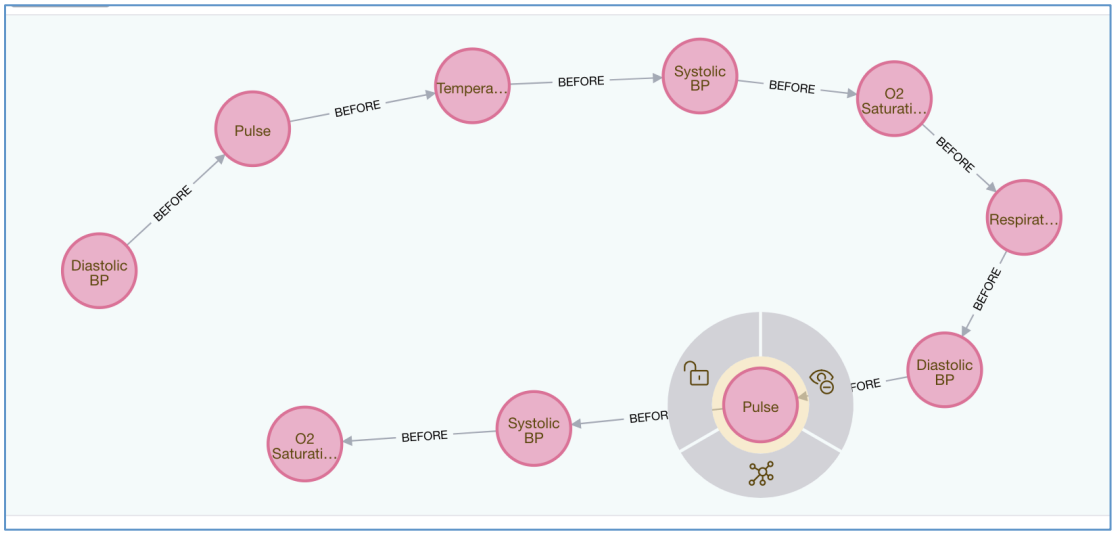

Figure 5: A temporal chain of events for a single patient represented by a graph. This data is modeled as a graph database using Neo4j graph DB engine

V. We also have discovered two novel genes (ABCG2 and PKD2) that are associated with disease severity (measured by the mean medical charges).

What opportunities for training and professional development has the project provided?
If the project was not intended to provide training and professional development opportunities or there is nothing significant to report during this reporting period, state "Nothing to Report."

Describe opportunities for training and professional development provided to anyone who worked on the project or anyone who was involved in the activities supported by the project. "Training" activities are those in which individuals with advanced professional skills and experience assist others in attaining greater proficiency. Training activities may include, for example, courses or one-on-one work with a mentor. "Professional development" activities result in increased knowledge or skill in one's area of expertise and may include workshops, conferences, seminars, study groups, and individual study. Include participation in conferences, workshops, and seminars not listed under major activities.

Nothing to report.

How were the results disseminated to communities of interest?
If there is nothing significant to report during this reporting period, state "Nothing to Report."

Describe how the results were disseminated to communities of interest. Include any outreach activities that were undertaken to reach members of communities who are not usually aware of these project activities, for the purpose of enhancing public understanding and increasing interest in learning and careers in science, technology, and the humanities.

Nothing to report.

What do you plan to do during the next reporting period to accomplish the goals?
If this is the final report, state "Nothing to Report."

Describe briefly what you plan to do during the next reporting period to accomplish the goals and objectives.

We are in the process of exploring genetic factors of other outcomes, such as pain and narcotic use.

4. **IMPACT:** Describe distinctive contributions, major accomplishments, innovations, successes, or any change in practice or behavior that has come about as a result of the project relative to:

What was the impact on the development of the principal discipline(s) of the project?
If there is nothing significant to report during this reporting period, state "Nothing to Report."

Describe how findings, results, techniques that were developed or extended, or other products from the project made an impact or are likely to make an impact on the base of knowledge, theory, and research in the principal disciplinary field(s) of the project. Summarize using language that an intelligent lay audience can understand (Scientific American style).

Nothing to report.

What was the impact on other disciplines?
If there is nothing significant to report during this reporting period, state "Nothing to Report."

Describe how the findings, results, or techniques that were developed or improved, or other products from the project made an impact or are likely to make an impact on other disciplines.

Nothing to report

What was the impact on technology transfer?
If there is nothing significant to report during this reporting period, state "Nothing to Report."

Describe ways in which the project made an impact, or is likely to make an impact, on commercial technology or public use, including:

- *transfer of results to entities in government or industry;*
- *instances where the research has led to the initiation of a start-up company; or*
- *adoption of new practices.*

Nothing to report.

What was the impact on society beyond science and technology?
If there is nothing significant to report during this reporting period, state "Nothing to Report."

Describe how results from the project made an impact, or are likely to make an impact, beyond the bounds of science, engineering, and the academic world on areas such as:

- *improving public knowledge, attitudes, skills, and abilities;*
- *changing behavior, practices, decision making, policies (including regulatory policies), or social actions; or*
- *improving social, economic, civic, or environmental conditions.*

Nothing to report.

5. **CHANGES/PROBLEMS:** The PD/PI is reminded that the recipient organization is required to obtain prior written approval from the awarding agency grants official whenever there are significant changes in the project or its direction. If not previously reported in writing, provide the following additional information or state, "Nothing to Report," if applicable:

Nothing to report.

Describe problems or delays encountered during the reporting period and actions or plans to resolve them.

Nothing to report.

Changes that had a significant impact on expenditures

Describe changes during the reporting period that may have had a significant impact on expenditures, for example, delays in hiring staff or favorable developments that enable meeting objectives at less cost than anticipated.

Purchase of a development server.

Significant changes in use or care of human subjects, vertebrate animals, biohazards, and/or select agents

Describe significant deviations, unexpected outcomes, or changes in approved protocols for the use or care of human subjects, vertebrate animals, biohazards, and/or select agents during the reporting period. If required, were these changes approved by the applicable institution committee (or equivalent) and reported to the agency? Also specify the applicable Institutional Review Board/Institutional Animal Care and Use Committee approval dates.

Significant changes in use or care of human subjects

Nothing to report.

Significant changes in use or care of vertebrate animals

Not applicable.

Significant changes in use of biohazards and/or select agents

Not applicable.

6. **PRODUCTS:** List any products resulting from the project during the reporting period. If there is nothing to report under a particular item, state "Nothing to Report."

- **Publications, conference papers, and presentations**
  Report only the major publication(s) resulting from the work under this award.

  Journal publications. *List peer-reviewed articles or papers appearing in scientific, technical, or professional journals. Identify for each publication: Author(s); title; journal; volume: year; page numbers; status of publication (published; accepted, awaiting publication; submitted, under review; other); acknowledgement of federal support (yes/no).*

  1. *Hassieb Din, Alyce J. Anderson, Claudia Ramos Rivers, Siobhan Proksell, Tariq Salim, Dmitriy Babichenko, Gong Tang, Ioannis E. Koutroubakis, Marc Schwartz, Elyse Johnston, Arthur Barrie, Janet Harrison, Michael A. Dunn, Douglas J. Hartman, David G. Binion.* Natural History Of Diabetes Mellitus And Inflammatory Bowel Disease: Increased Disease Severity, Worse Quality Of Life, And Under-Treatment With Immunomodulator And/Or Biologic Agents. Inflammatory Bowel Diseases. 2019. Accepted.
  2. *Krishnapriya Marangattu Prathapan, Claudia Ramos Rivers, Sandra C. Kim, Ioannis Koutroubakis, Dmitriy Babichenko, Gong Tang, Marc Schwartz, Siobhan Proksell, Elyse Johnston, Jana G. Hashash, Michael A. Dunn, Annette Wilson, Alyce J. Anderson, Arthur Barrie, Janet Harrison, Douglas J. Hartman, David G. Binion.* Peripheral Blood Eosinophilia Is A Biomarker Of Long-Term Severity In Pediatric-Onset Inflammatory Bowel Disease Patients. Inflammatory Bowel Diseases. 2019. Accepted.
  3. *Filippos Koutroumpakis, Anna Evans Phillips, Dhiraj Yadav, Jorge D Machicado, Claudia Ramos-Rivers, Marc Schwartz, Siobhan Proksell, Elyse Johnston, Jeffrey Dueker, Jana G Hashash, Arthur Barrie, Janet Harrison, Michael A Dunn, Liza Konnikova, Douglas J Hartman, Hasieb Din, Dmitriy Babichenko, Gong Tang, David G Binion.* Low Serum Levels of IgG4 Antibodies Define a Commonly Encountered, Severe Inflammatory Bowel Disease Subtype. Clinical gastroenterology and Hepatology. 2019. Submitted.

  Books or other non-periodical, one-time publications. *Report any book, monograph, dissertation, abstract, or the like published as or in a separate publication, rather than a periodical or series. Include any significant publication in the proceedings of a one-time conference or in the report of a one-time study, commission, or the like. Identify for each one-time publication: author(s); title; editor; title of collection, if applicable; bibliographic information; year; type of publication (e.g., book, thesis or dissertation);*

status of publication (published; accepted, awaiting publication; submitted, under review; other); acknowledgement of federal support (yes/no).

Nothing to report.

Other publications, conference papers and presentations. *Identify any other publications, conference papers and/or presentations not reported above. Specify the status of the publication as noted above. List presentations made during the last year (international, national, local societies, military meetings, etc.). Use an asterisk (*) if presentation produced a manuscript.*

American College of Gastroenterology Annual meeting 2019:

1. Low Serum Levels of IgG4 Antibodies May Function as a Biomarker of Severity in Inflammatory Bowel Disease. *Filippos Koutroumpakis, Anna Evans Phillips, Dhiraj Yadav, Jorge D Machicado, Claudia Ramos-Rivers, Marc Schwartz, Siobhan Proksell, Elyse Johnston, Jeffrey Dueker, Jana G Hashash, Arthur Barrie, Janet Harrison, Michael A Dunn, Liza Konnikova, Douglas J Hartman, Hasieb Din, Dmitriy Babichenko, Gong Tang, David G Binion.* ACG 2019*

Digestive Diseases Week 2019:

1. Candidate And Exploratory Genetic Association Study Of Ibd Severity Using A Novel Phenotype (Multiyear Mean Healthcare Charges). *Tanvi Nagpal, David G. Binion, Claudia Ramos Rivers, Yan Lin.* DDW 2019
2. Peripheral Blood Eosinophilia Is A Biomarker Of Long-Term Severity In Pediatric-Onset Inflammatory Bowel Disease Patients. *Krishnapriya Marangattu Prathapan, Claudia Ramos Rivers, Sandra C. Kim, Ioannis Koutroubakis, Dmitriy Babichenko, Gong Tang, Marc Schwartz, Siobhan Proksell, Elyse Johnston, Jana G. Hashash, Michael A. Dunn, Annette Wilson, Alyce J. Anderson, Arthur Barrie, Janet Harrison, Douglas J. Hartman, David G. Binion.* DDW 2019*
3. Characterizing The Colonoscopic Features Of Clostridium Difficile Infection In Inflammatory Bowel Disease. *Vance Hartke, Siobhan Proksell, Alyce J. Anderson, Claudia Ramos Rivers, Marc Schwartz, Elyse Johnston, Jana G. Hashash, Arthur Barrie, Janet Harrison, Ioannis E. Koutroubakis, Douglas J. Hartman, Dmitriy Babichenko, Michael A. Dunn, David G. Binion.* DDW 2019
4. Peripheral Blood Eosinophilia Functions As A Candidate Biomarker Of Decreased Response To Anti-Tnf Therapy In Crohn's Disease. *Scott Friedberg, Weston Bettner, Xianling Wang, Claudia Ramos Rivers, Ioannis E. Koutroubakis, Gong Tang, Dmitriy Babichenko, Siobhan Proksell, Elyse Johnston, Marc Schwartz, Jana G. Hashash, Arthur Barrie, Janet Harrison, Douglas J. Hartman, Michael A. Dunn1, David G. Binion.* DDW 2019
5. Does Inflammation In Ibd "Burnout" Over Time? *Hassieb Din, Alyce J. Anderson, Claudia Ramos Rivers, Dmitriy Babichenko, Gong Tang, Ioannis E. Koutroubakis, Marc Schwartz, Siobhan Proksell, Elyse Johnston, Arthur Barrie, Janet Harrison, Jana G. Hashash, Michael A. Dunn, Douglas J. Hartman, Tariq Salim, Eva Szigethy, David G. Binion.* DDW 2019
6. Big Data Analytics Identifies Ulcerative Colitis Patients At Increased Risk For Incident Colorectal Neoplasia Using Multiyear Patterns Of Routine Clinical Lab Values. *Carlita Shen, Claudia Ramos Rivers, Dmitriy Babichenko, Douglas J. Hartman, Ioannis Koutroubakis, Marc Schwartz, Siobhan Proksell, Elyse Johnston, Jana G. Hashash, Arthur Barrie2, Janet Harrison, Gong Tang, Andrew R. Watson, David G. Binion.* DDW 2019
7. Natural History Of Diabetes Mellitus And Inflammatory Bowel Disease: Increased Disease Severity, Worse Quality Of Life, And Under-Treatment With Immunomodulator And/Or Biologic Agents. *Hassieb Din, Alyce J. Anderson, Claudia Ramos Rivers, Siobhan Proksell, Tariq Salim, Dmitriy Babichenko, Gong Tang, Ioannis E. Koutroubakis, Marc Schwartz, Elyse Johnston, Arthur Barrie, Janet Harrison, Michael A. Dunn, Douglas J. Hartman, David G. Binion.* DDW 2019*
8. *Endoscopic Patterns And Location Of* Post-Operative Recurrence In Crohn's Disease Patients With Side To Side Anastomosis Following Ileocecal Resection. *Furkan Ertem, Andrew R. Watson, Claudia Ramos Rivers, Dmitriy Babichenko, Gong Tang, Marc Schwartz, Siobhan Proksell, Elyse Johnston, Jana G. Hashash, Arthur Barrie, Janet Harrison, Ioannis E. Koutroubakis, Michael A. Dunn, Douglas J. Hartman, David G. Binion.* DDW 2019

List the URL for any Internet site(s) that disseminates the results of the research activities. A short description of each site should be provided. It is not necessary to include the publications already specified above in this section.

Nothing to report.

- **Technologies or techniques**
  Identify technologies or techniques that resulted from the research activities. Describe the technologies or techniques were shared.

Nothing to report.

- **Inventions, patent applications, and/or licenses**
  Identify inventions, patent applications with date, and/or licenses that have resulted from the research. Submission of this information as part of an interim research performance progress report is not a substitute for any other invention reporting required under the terms and conditions of an award.

Nothing to report.

- **Other Products**
  Identify any other reportable outcomes that were developed under this project. Reportable outcomes are defined as a research result that is or relates to a product, scientific advance, or research tool that makes a meaningful contribution toward the understanding, prevention, diagnosis, prognosis, treatment and /or rehabilitation of a disease, injury or condition, or to improve the quality of life. Examples include:
  - *data or databases;*
  - *physical collections;*
  - *audio or video products;*
  - *software;*

- *models;*
- *educational aids or curricula;*
- *instruments or equipment;*
- *research material (e.g., Germplasm; cell lines, DNA probes, animal models);*
- *clinical interventions;*
- *new business creation; and*
- *other.*

Nothing to report.

7. PARTICIPANTS & OTHER COLLABORATING ORGANIZATIONS

Name: David G. Binion
Project Role: PI
Nearest person month(s) worked: 3
Contribution to Project: Dr. Binion oversaw all research activities in this project. Bi- weekly research meetings were held to disseminate progress. Dr. Binion has performed work providing strategies for extraction and preparation of clinically relevant variables and expert advice on decision tool modeling and interface development.

Name: Gong Tang
Project Role: Co-Investigator
Nearest person month(s) worked: 3
Contribution to Project: Dr. Tang has developed statistical models to identify patients with high medical charges and patients with low medical charges in the future with historical data from a pre-determined training set and assessed its performance in a separate testing set.

Name: *Dmitriy Babichenko*
Project Role: Co- Investigator
Nearest person month(s) worked: 2
Contribution to Project: Dr. Babichenko completed data de-identification automation scripts Created and evaluated a series of classification and probabilistic models trained from the IBD registry dataset. Completed data de-identification automation scripts. Worked on creating the initial decision support system user interface designs.

Name: *Marek Drudzel*
Project Role: *Co-investigator*
Nearest person month(s) worked: 2
Contribution to Project: Dr. Drudzel has performed extensive exploration of data to determine analysis to evaluate complicated/ high cost disease using Bayesian networks. Created and evaluated a series of classification and probabilistic models trained from the IBD registry dataset.

Name: *Mark Roberts*
Project Role: Co- Investigator
Nearest person month(s) worked: 1
Contribution to Project: Roberts has been providing advice about best strategies for extraction and preparation of clinical data. He has also provided advice on epidemiological relevance.

Name: *Michael Dunn*
Project Role: Co-Investigator
Nearest person month(s) worked: 1
Contribution to Project: Dr. Dunn has been providing advice and expertise about best strategies for extraction and preparation of clinically relevant variables.

Name: Claudia Ramos Rivers
Project Role: *Key personnel- Research Scientist*
Nearest person month(s) worked: 7
Contribution to Project: Dr. Ramos Rivers has overseen protocol submission for IRB approval as well as preparing progress reports. Dr. Ramos Rivers has also coordinated and attended to bi- weekly meetings to develop strategies on data extraction and preparation for analysis.

Name: *Annette Wilson*
Project Role: *Key personnel- Lab. Manager*
Nearest person month(s) worked: 2
Contribution to Project: Dr. Wilson has been responsible for the post award administrative work.

Name: *Yan Lin*
Project Role: *Key personnel - Faculty*
Nearest person month(s) worked: 2
Contribution to Project: *Dr. Lin has participated in developing prediction models for future clinical outcomes and worked with Dr. Binion and a research staff on bioinformatics analyses of genetic data from those IBD patients.*

Name: *Krauland, Mary G*
Project Role: *Graduate Student. Graduate School of Public Health*
Nearest person month(s) worked: 10
Contribution to Project: Under the supervision of Dr. Roberts, Mrs. Krauland has been providing advice about best strategies for extraction and preparation of clinical data. He has also provided advice on epidemiological relevance.

Name: *Behnam Rahdari*
Project Role: *Graduate Student Researcher*
Nearest person month(s) worked: 9
Contribution to Project: *Under Marek Drudzel and Dmitriy Babichenko supervision, Behnam Rahdari has continued the work performed by Marcin Kozniewski on exploration of data to determine analysis to evaluate complicated/ high cost disease using Bayesian network.*

Name: *Xianling Wang*
Project Role: *Graduate Student Researcher*
Nearest person month(s) worked: 11
Contribution to Project: Under the supervision of Dr. Tang, Ms. Wang has performed extensive analyses to predict future clinical outcomes of IBD patients based on demographics, historical lab data and other medical records. Ms. Wang will explore more comprehensive modelling to include dimension reduction, regularization and causal pathway analysis.

Name: *Beata Pasek*
Project Role: Clinical Research coordinator
Nearest person month(s) worked: 2
Contribution to Project: Mrs. Pasek has consented patients currently in the study and has been responsible for regulatory activities.

Has there been a change in the active other support of the PD/PI(s) or senior/key personnel since the last reporting period?

Nothing to Report.

What other organizations were involved as partners?

Nothing to Report.

8. SPECIAL REPORTING REQUIREMENTS

COLLABORATIVE AWARDS: For collaborative awards, independent reports are required from BOTH the Initiating Principal Investigator (PI) and the Collaborating/Partnering PI. A duplicative report is acceptable; however, tasks shall be clearly marked with the responsible PI and research site. A report shall be submitted to **https://ers.amedd.army.mil** for each unique award.

QUAD CHARTS: If applicable, the Quad Chart (available on https://www.usamraa.army.mil) should be updated and submitted with attachments.

9. APPENDICES:

Attach all appendices that contain information that supplements, clarifies or supports the text. Examples include original copies of journal articles, reprints of manuscripts and abstracts, a curriculum vitae, patent applications, study questionnaires, and surveys, etc.