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

Pittsburgh, PA 15213-2303

REPORT DATE: September 2018

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 2018	Annual	1 Sep 2017 - 31 Aug 2018

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
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	21	

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.

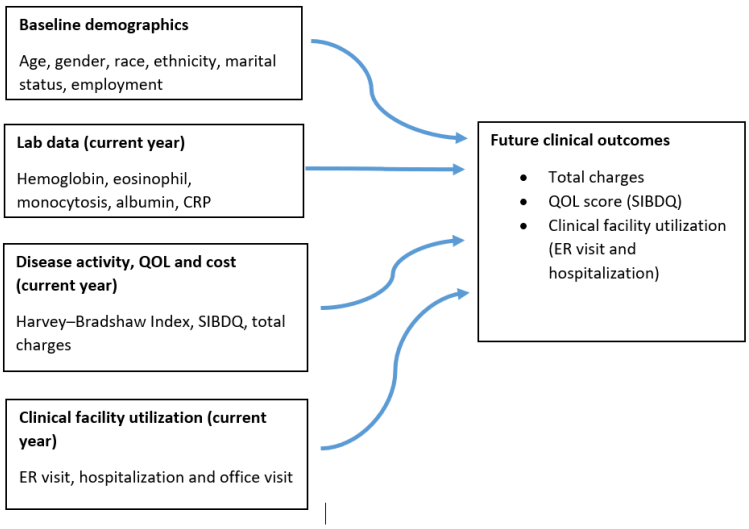
What were the major goals of the project?
List the major goals of the project as stated in the approved SOW. If the application listed milestones/target dates for important activities or phases of the project, identify these dates and show actual completion dates or the percentage of completion.

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

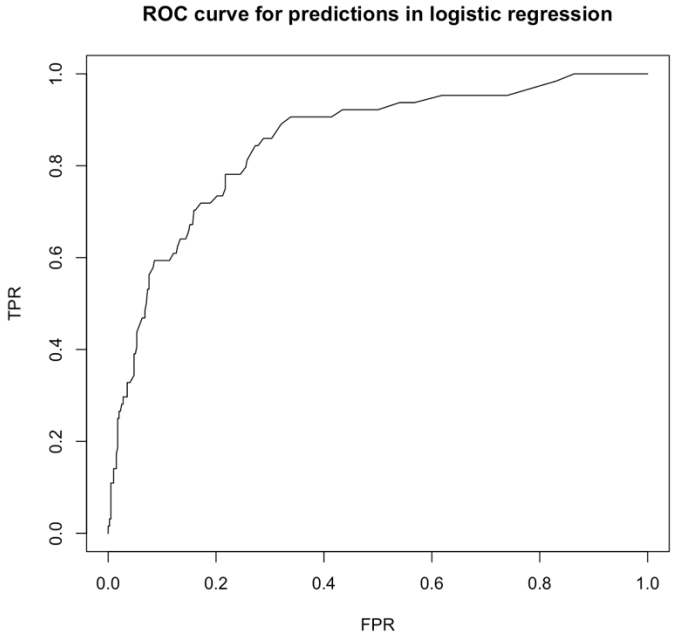
What was accomplished under these goals?
For this reporting period describe: 1) major activities; 2) specific objectives; 3) significant results or key outcomes, including major findings, developments, or conclusions (both positive and negative); and/or 4) other achievements. Include a discussion of 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. Development of statistical models to predict medical charges, quality of life outcome SIBDQ scores, and use of medical facility (emergency room visit and hospitalization) with historical data from a pre-determined training set and assessed its performance in a separate testing set.

Current modeling strategy:



Results from an internal validation of a prediction model on the total charges (>$15K) of a future year:



II. Completed data de-idenfication automation scripts
III. Completed automation of training dataset generation
IV. Completed initial predictor selection
Predictors (Independent variables):
    a. **Demographic**: gender, marital status, employment status, age, distance to clinic, median income
    b. **Medications**: 5-ASA, immunomodulators, systemic steroids, vitamin D, biologics

c. **Encounters**: procedure visit, contact with provider (office visit, email, phone)
d. **Labs**: Eosinophils, monocytes, inflammation markers, hemoglobin, vitamin D
e. **Quality of life (SIBDQ)**: SIBDQ total score, Question 1: feeling fatigue, question 8: feeling relaxed
f. **Psychiatric comorbidities** (including SIBDQ scores related to stress)

Response (Dependent variable):

a. Treatment cost

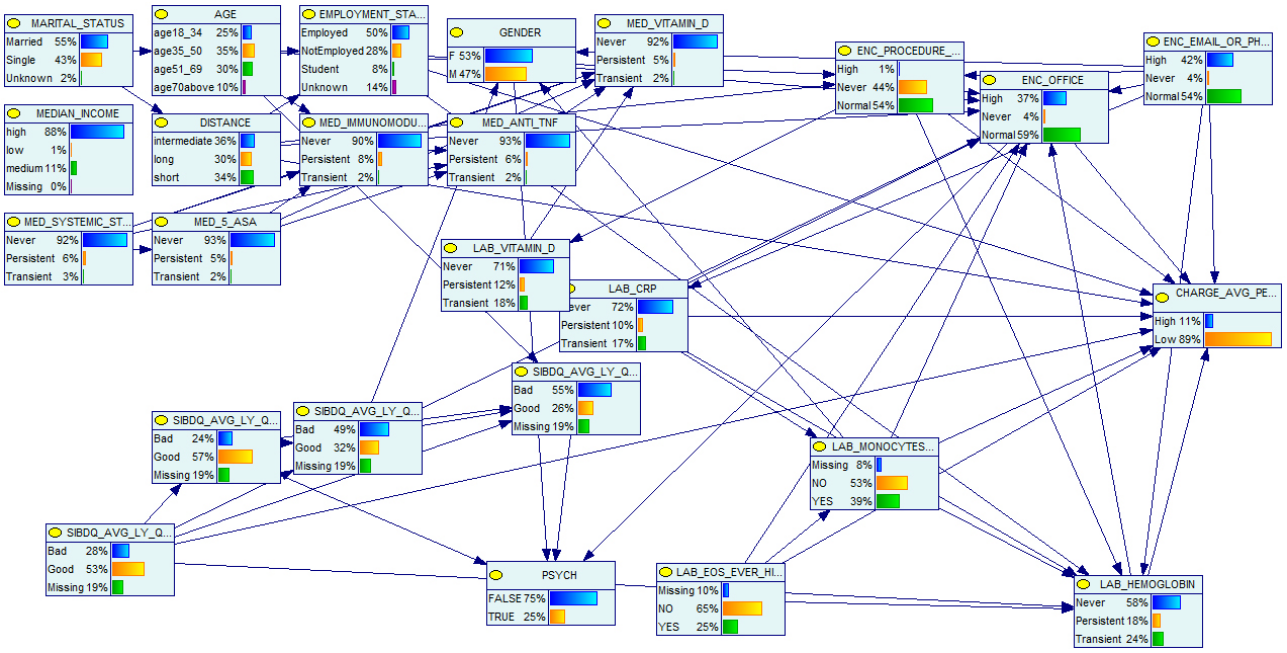
V. Created and evaluated a series of classification and probabilistic models trained from the IBD registry dataset

Preliminary Classification Models

Classification Algorithm	Accuracy Score	Cross-validated Score (k = 10)
SVM with rbf kernel	0.908	0.910
Naïve Bayes	0.858	0.853
SVM with linear kernel	0.896	0.901
SVM with poly kernel	0.923	0.914
KNN with 3 neighbors	0.910	0.87
Random Forest	0.917	0.891
Decision Tree	0.903	0.898
Logistic Regression	0.929	0.911

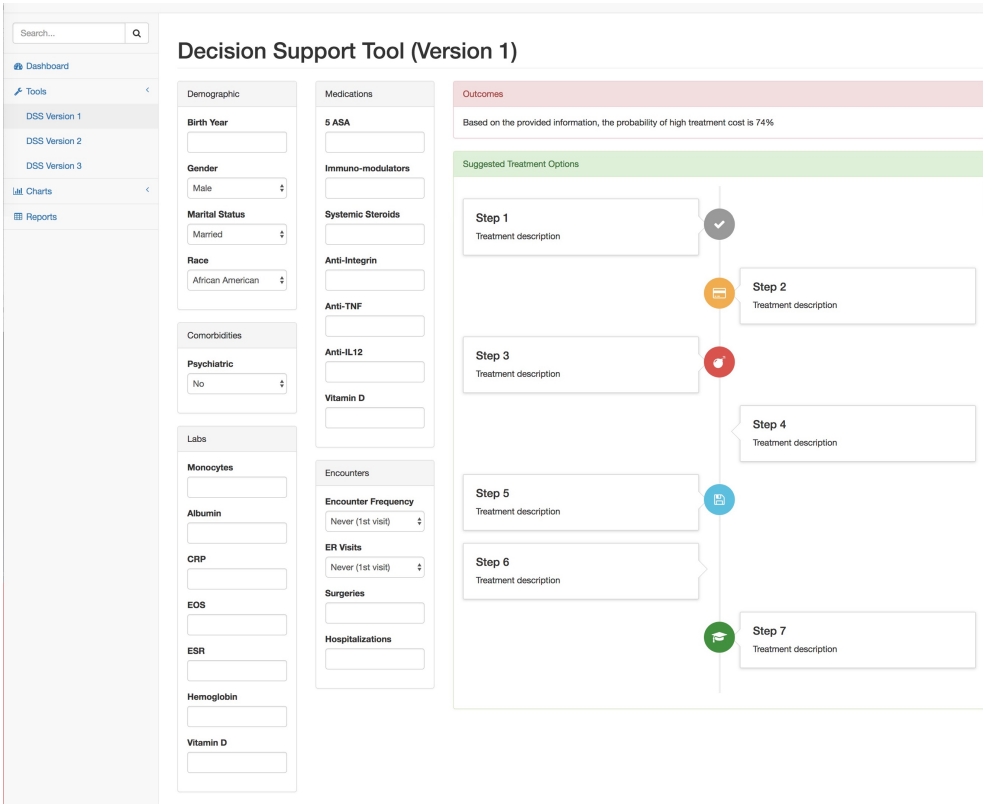
Preliminary Bayesian Network Model

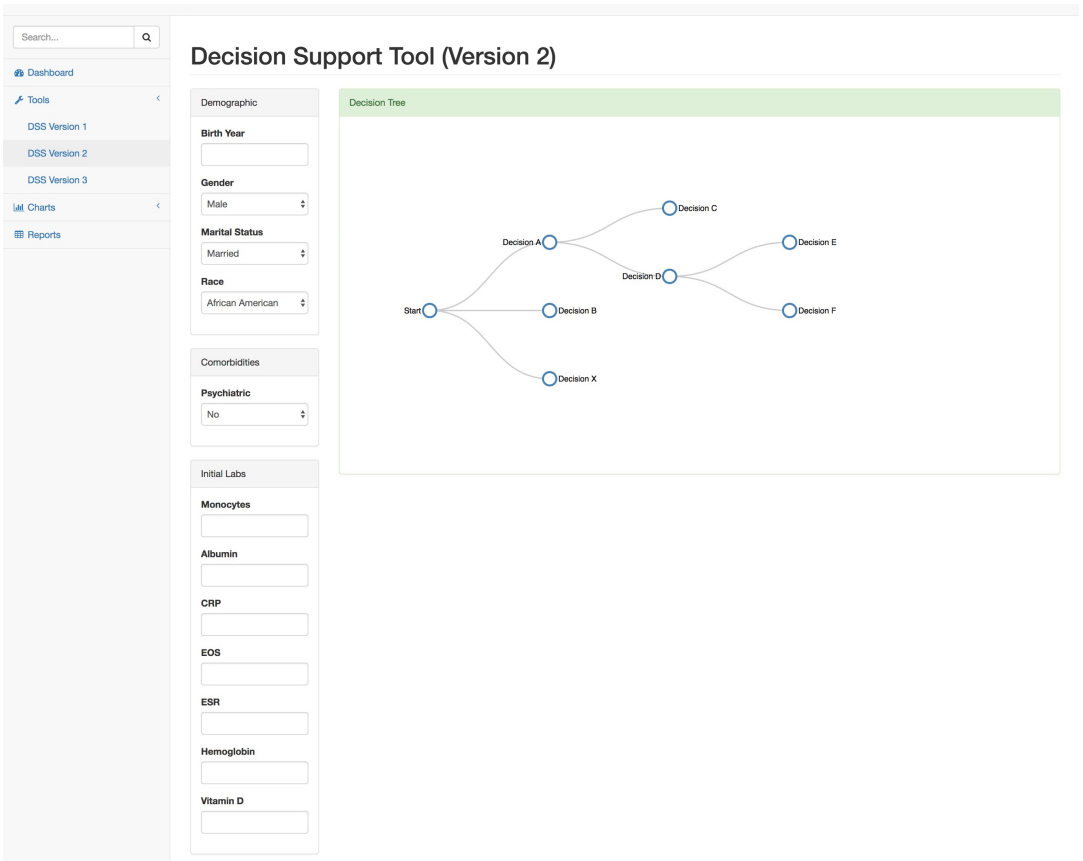
Training Algorithm	Accuracy Score	Cross-validated Score (k = 10)
PC	0.892	0.877
Tree-augmented Naïve Bayes (TAN)	0.901	0.899

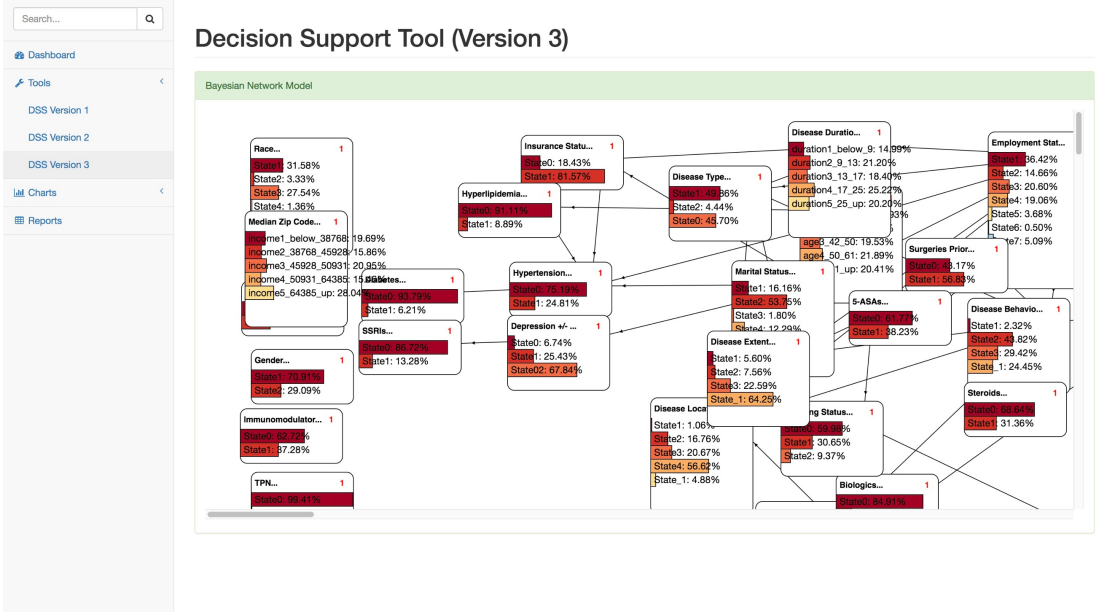


VI. Completed initial decision support system user interface designs

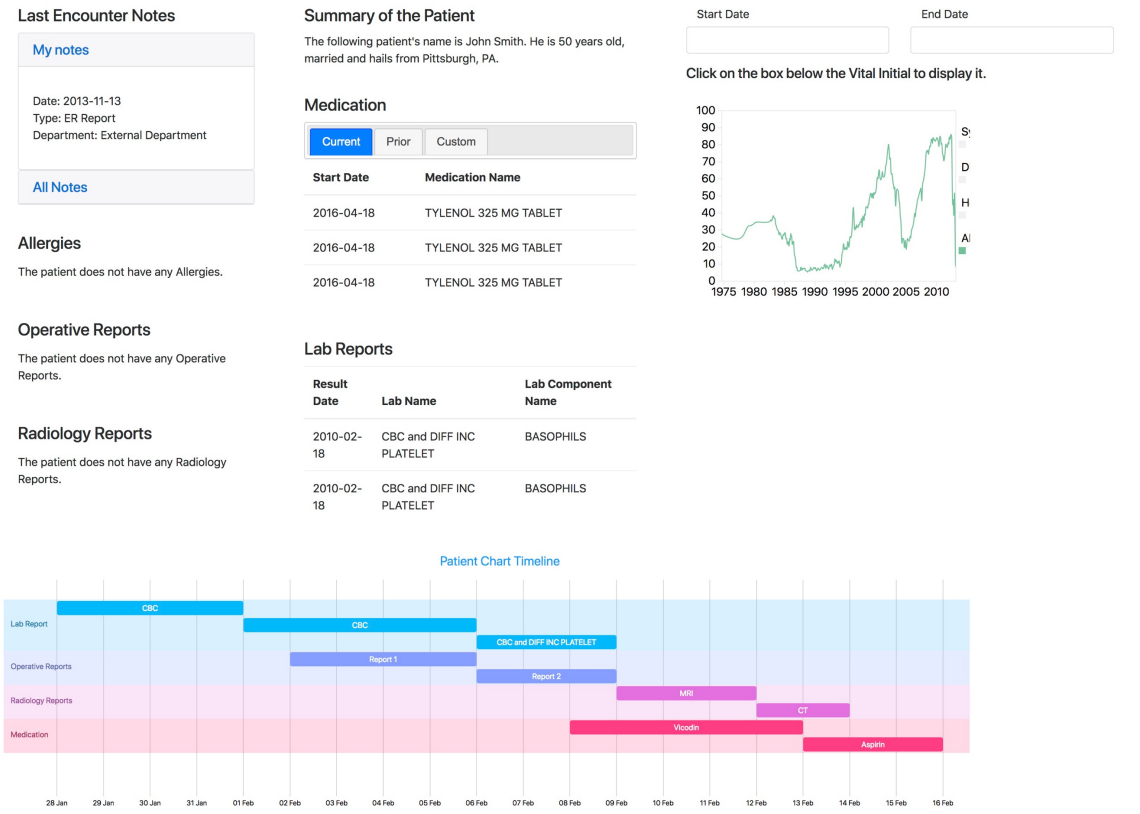
Proposed Decision Support System (DSS) Interfaces




Search...
Dashboard
Tools
DSS Version 1
DSS Version 2
DSS Version 3
Charts
Reports
Decision Support Tool (Version 2)
Demographic
Birth Year
Gender
Male
Marital Status
Married
Race
African American
Comorbidities
Psychiatric
No
Initial Labs
Monocytes
Albumin
CRP
EOS
ESR
Hemoglobin
Vitamin D
Decision Tree
Start
Decision A
Decision B
Decision C
Decision D
Decision E
Decision F
Decision X


Search...
Dashboard
Tools
DSS Version 1
DSS Version 2
DSS Version 3
Charts
Reports
Decision Support Tool (Version 3)
Bayesian Network Model
Race...
Median Zip Code...
Insurance Statu...
Hyperlipidemia...
Disease Type...
Disease Duratio...
Employment Stat...
Surgeries Prior...
Marital Status...
Hypertension...
Depression +/- ...
SSRIs...
Gender...
Immunomodulator...
TPN...
Disease Extent...
5-ASAs...
Disease Behavio...
Steroids...
Biologics...

VII. Conducted focus groups with clinicians about decision support system design



VIII. Evaluated FIHR API standard (https://www.hl7.org/fhir/http.html) for connecting models and the decision support system to systems such as EPIC and CERNER

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.

1. Expand the modelling effort to predict medical charges to include dimension reduction, regularization and causal pathway analysis.
2. Investigate the genetic mechanism behind the severity of IBD symptom, which may improve the performance of the prediction model by adding an additional dimension.
3. Improve models' predictive accuracy
    a. Explore different feature selection algorithms
    b. Explore different dimension reduction algorithms
4. Train probabilistic models using additional machine learning algorithms such as greedy thick thinning
5. Identify 3 levels of predictors, based on how much data is available to different providers to build three sets of models for different users

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.

Actual or anticipated problems or delays and actions or plans to resolve them
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.

We hired a clinical research coordinator Beata Pasek.

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).*

Nothing to report.

  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.*

Digestive Disease Week 2018:

1. DOES SURGICAL ANASTAMOSIS TYPE IMPACT RATES OF ENDOSCOPIC RECURRENCE IN POST-OPERATIVE CROHN'S DISEASE? AN 8-YEAR OBSERVATIONAL COHORT STUDY. ***Furkan Ubeydullah Ertem, Claudia Ramos Rivers, Miguel D. Regueiro, Andrew R. Watson, Marc Schwartz, Ioannis Koutroubakis, Jana G. Hashash, Benjamin H. Click, Michael A Dunn, Dmitriy Babichenko, David G. Binion.***
2. BIOMAKERS ASSOCIATED WITH EARLY ENDOSCOPIC RECURRENCE OF POSTOPERATIVE CROHN'S DISEASE. ***Furkan Ertem, Claudia Ramos Rivers, Miguel Regueiro, Benjamin Click, Ioannis Koutroubakis, Marc Schwartz, Andrew Watson, Jana Hashash, Michael Dunn, Dmitriy Babichenko, David Binion.***
3. FINDING THE SWEET SPOT: THE ASSOCIATION BETWEEN ADDED DIETARY SUGARS AND INFLAMMATORY BOWEL DISEASE SEVERITY. ***Maaz Ahsan, Alyce Anderson, Dmitriy Babichenko, Claudia Ramos Rivers, Stephen O'Keefe, Miguel Regueiro, Marc Schwartz, Jana Hashash, Benjamin Click, Ioannis Koutroubakis, Michael Dunn, David Binion***
4. SEVEN-YEAR PERIOD PREVALENCE AND CHARACTERISTICS OF HYPERTENSION IN A LARGE US INFLAMMATORY BOWEL DISEASE COHORT. ***Kelly Gibbs, Alyce Anderson, Claudia Ramos Rivers, Benjamin Click, Ioannis Koutroubakis, Miguel Regueiro, Michael Dunn, Marc Schwartz, Jana Hashash, Dmitriy Babichenko, David Binion.***
5. THE IMPACT OF SERUM CORTISOL ON QUALITY OF LIFE AND DYSAUTONOMIA IN INFLAMMATORY BOWEL DISEASE. ***Benjamin Click, Alyce Anderson, Claudia Ramos Rivers, Marc Schwartz, Arthur Barrie, Michael Dunn, David Levinthal, Miguel Regueiro, David Binion.***

American College of Gastroenterology Annual meeting 2018:

1. IMPACT OF CHOLECYSTECTOMY ON CLINICAL COURSE OF IBD, INCREASED DIARRHEA, HEALTHCARE CHARGES, NARCOTIC USE AND DECREASED QUALITY OF LIFE INDEPENDENT OF INFLAMMATION. ***Maham Lodhi, Claudia Ramos-Rivers, Marc Schwartz, Dmitriy Babichenko, Gong Tang, Tanvi Nagpal, Michael Dunn, David Binion.***

- **Website(s) or other Internet site(s)**

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

What individuals have worked on the project?

Name: David G. Binion
Project Role: PI
Nearest person month(s) worked: 3
Contribution to Project: Dr. Binion oversaw all research 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. In addition, he has interviewed candidates for the Nurse Research Coordinator position.

Name: Gong Tang
Project Role: Co-Investigator
Nearest person month(s) worked: 2
Contribution to Project: Dr. Tang has developed statistical models to predict medical charges, the quality of life outcome SIBDQ scores, and use of medical facility (emergency room visit and hospitalization) 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: 1
Contribution to Project: Dr. Babichenko completed data de-idenfication 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: 1
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: *Marcin Kozniewski*
Project Role: *Graduate Student Researcher*
Nearest person month(s) worked: 11
Contribution to Project: *Under Marek Drudzel supervision, Marcin Kozniewski has performed exploration of data to determine analysis to evaluate complicated/ high cost disease using Bayesian networks*.

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.