**INSTRUCTIONS: Complete Research Protocol (HRP-503)**

* Depending on the nature of what you are doing, some sections may not be applicable to your research. If so, you must provide the reason why the section is not applicable for the response. For example, most behavioral studies would answer all questions in section 30 with words to the effect of “drugs and medical devices are not used in this study.”
* When you write a protocol, keep an electronic copy. You will need to modify this copy when making changes.
* Do not remove the italics instructions or headings.
* If you are pasting information from other documents be sure to use the “Merge Formatting” paste option so that the formatting of the response boxes is not lost. If information is presented outside of the response boxes, it will not be accepted.
* If this study involves multiple participant groups who participate in different research procedures, consent processes, etc., be certain to provide information in each applicable section for each participant group and clearly label each participant group within a section or subsection.

**PROTOCOL TITLE:**

*Include the full protocol title.*

Quantifying ‘copy-and-paste’ events in outpatient electronic medical records

**PRINCIPAL INVESTIGATOR:**

*Name*

*Department*

*Telephone Number*

*Email Address*

Response:

Edwin Anand

Department of Biomedical Informatics

(716) 888-4858

eanand@buffalo.edu

**VERSION NUMBER:**

*Include the version number of this protocol.*

Response: 1.0

**DATE:**

*Include the date of submission or revision.*

**Grant Applicability:**

*Describe whether or not this protocol is funded by a grant or contract and if so, what portions of the grant this study covers.*

There isn’t funding for this project. It is being supported through departmental resources.

Table of Contents

[1.0 Objectives 4](#_Toc396981544)

[2.0 Background 4](#_Toc396981545)

[3.0 Inclusion and Exclusion Criteria 5](#_Toc396981546)

[4.0 Study-Wide Number of Subjects (Multisite/Multicenter Only) 6](#_Toc396981547)

[5.0 Study-Wide Recruitment Methods (Multisite/Multicenter Only) 6](#_Toc396981548)

[6.0 Multi-Site Research (Multisite/Multicenter Only) 7](#_Toc396981549)

[7.0 Study Timelines 7](#_Toc396981550)

[8.0 Study Endpoints 8](#_Toc396981551)

[9.0 Procedures Involved 8](#_Toc396981552)

[10.0 Data and Specimen Banking 9](#_Toc396981553)

[11.0 Data Management 9](#_Toc396981554)

[12.0 Provisions to Monitor the Data and Ensure the Safety of Subjects 10](#_Toc396981555)

[13.0 Withdrawal of Subjects 11](#_Toc396981556)

[14.0 Risks to Subjects 12](#_Toc396981557)

[15.0 Potential Benefits to Subjects 12](#_Toc396981558)

[16.0 Vulnerable Populations 12](#_Toc396981559)

[17.0 Community-Based Participatory Research 13](#_Toc396981560)

[18.0 Sharing of Results with Subjects 13](#_Toc396981561)

[19.0 Setting 14](#_Toc396981562)

[20.0 Resources Available 14](#_Toc396981563)

[21.0 Prior Approvals 15](#_Toc396981564)

[22.0 Recruitment Methods 16](#_Toc396981565)

[23.0 Local Number of Subjects 16](#_Toc396981566)

[24.0 Confidentiality 16](#_Toc396981567)

[25.0 Provisions to Protect the Privacy Interests of Subjects 17](#_Toc396981568)

[26.0 Compensation for Research-Related Injury 17](#_Toc396981569)

[27.0 Economic Burden to Subjects 18](#_Toc396981570)

[28.0 Consent Process 18](#_Toc396981571)

[29.0 Process to Document Consent in Writing 22](#_Toc396981572)

[30.0 Drugs or Devices 22](#_Toc396981573)

# Objectives

* 1. Describe the purpose, specific aims, or objectives.

The objective of the study is to quantify the frequency and prevalence of “copy and paste” events in narrative entries of electronic medical records (EMRs) using formal mathematical and information-theoretic criteria, e.g. text unification and winnowing, and statistical testing, using large EMR corpora.

* 1. State the hypotheses to be tested.

A significant fraction of narrative parts of EMR is generated via “copy and paste” behavior, and simple string-comparison-based methods underestimate or incorrectly detect such events.

# Background

* 1. Describe the relevant prior experience and gaps in current knowledge.

Electronic Medical Records are increasingly replacing traditional paper charts.1 While this process is advantageous in many aspects, including high availability of the data, the ease of monitoring for disease trends or the ability to provide live decision support to providers, it enables behaviors with a potentially adverse impact. One such behavior is copying and pasting by providers, parts of their own or other providers’ narrative entries of EMR.

Although “copy and paste” events often accelerate document creation, their potential negative effects cannot be neglected. A provider who has copied information may or may not have fully reviewed the information being transferred, which may lead to percolation of incomplete or incorrect data. Moreover, cognitive processing and summarization of the patient’s information is likely poorer in such notes.2,3

To this day, only a few limited studies have approached the topic.4,5 These studies: i) involved only a small number of EMR, ii) involved inadequate text processing methods designed for biological sequences comparison and not plagiarism or similarity detection in text corpora, iii) did not provide statistically justified measures to reliably assess scale of the “copy and paste” events. Consequently, the extent of “copy and paste” behavior remains unclear.

* 1. Describe any relevant preliminary data.

In the referenced paper by Wren et al. 54% of the information was duplicated from previous narratives among progress notes in inpatient encounters.4

Cohen et al. studied the effect of documentation redundancy and found that it significantly affected data mining results. In their study, about 40% of the charts exhibited high redundancy.5

* 1. Provide the scientific or scholarly background for, rationale for, and significance of the research based on the existing literature and how will it add to existing knowledge.

The literature is sparse on the extent of copy and paste behavior, and a few available studies are limited in their methodology. Here, for the first time we propose to use robust and formal techniques and large EMR corpora to capture and quantify “copy and paste” behavior. This study will provide much deeper insights into the issue than what is available in the current literature. Moreover, it will enable us to assess quality of the techniques used thus far, which we believe might be inadequate. Finally, it will lay the foundation for future studies to help better EMR development.

* 1. Include complete specific citations/references.

1. Hsiao CJ, Hing E, Ashman J. Trends in electronic health record system use among office-based physicians: United States, 2007-2012. Natl Health Stat Report. 2014(75):1-18.

2. Hirschtick RE. A piece of my mind. Copy-and-paste. JAMA. 2006;295(20):2335-2336.

3. Wrenn JO, Stein DM, Bakken S, Stetson PD. Quantifying clinical narrative redundancy in an electronic health record. J Am Med Inform Assoc. 2010;17(1):49-53.

4. Cohen R, Elhadad M, Elhadad N. Redundancy in electronic health record corpora: analysis, impact on text mining performance and mitigation strategies. BMC Bioinformatics. 2013;14:10.

# Inclusion and Exclusion Criteria

* 1. Describe the criteria that define who will be included or excluded in your final study sample.

2000 outpatient records of patients aged 18 or more, created by physicians of the different specialties at the UBMD clinics.

All outpatient visits that involve contact with a physician or a mid-level provider will be included

Exclusions:

Encounters where the patient was not physically seen by a provider will be excluded. These include nurse visits for vaccination, phone consults.

Charts of children, pregnant women and prisoners will be excluded.

* 1. Describe how individuals will be screened for eligibility.

Records will be considered if they fit into eligibility described above based on the visit codes.

* 1. Indicate specifically whether you will include or exclude each of the following special populations: (You may not include members of these populations as subjects in your research unless you indicate this in your inclusion criteria.)
     + Adults unable to consent
     + Individuals who are not yet adults (infants, children, teenagers)
     + Pregnant women
     + Prisoners

Children, pregnant women and prisoner records will be excluded from our study.

* 1. Indicate whether you will include non-English speaking individuals. Provide justification if you will exclude non-English speaking individuals.   
     (In order to meet one of the primary ethical principles of equitable selection of subjects, non-English speaking individuals may not be routinely excluded from research. In cases where the research is of therapeutic intent or is designed to investigate areas that would necessarily require certain populations who may not speak English, the researcher is required to make efforts to recruit and include non-English speaking individuals. However, there are studies in which it would be reasonable to limit subjects to those who speak English: e.g., pilot studies, small unfunded studies with validated instruments not available in other languages, numerous questionnaires, and some non-therapeutic studies which offer no direct benefit.)

There may be records of non-English speaking individuals included in the pool from which records will be drawn. However, only charts that are recorded in English will be used.

# Study-Wide Number of Subjects (Multisite/Multicenter Only)

* 1. If this is a multicenter study, indicate the total number of subjects to be accrued across all sites.

This is not a multicenter study.

# Study-Wide Recruitment Methods (Multisite/Multicenter Only)

If this is a multicenter study and subjects will be recruited by methods not under the control of the local site (e.g., call centers, national advertisements) describe those methods. Local recruitment methods are described later in the protocol.

* 1. Describe when, where, and how potential subjects will be recruited.

This is not a multicenter study.

* 1. Describe the methods that will be used to identify potential subjects.

Any record that meets the above inclusion criteria previously described will be included.

* 1. Describe materials that will be used to recruit subjects. (Attach copies of these documents with the application. For advertisements, attach the final copy of printed advertisements. When advertisements are taped for broadcast, attach the final audio/video tape. You may submit the wording of the advertisement prior to taping to preclude re-taping because of inappropriate wording, provided the IRB reviews the final audio/video tape.)

There are no recruitment materials for this study.

# Multi-Site Research (Multisite/Multicenter Only)

* 1. If this is a multi-site study where you are the lead investigator, describe the processes to ensure communication among sites, such as:
     + All sites have the most current version of the protocol, consent document, and HIPAA authorization.
     + All required approvals have been obtained at each site (including approval by the site’s IRB of record).
     + All modifications have been communicated to sites, and approved (including approval by the site’s IRB of record) before the modification is implemented.
     + All engaged participating sites will safeguard data as required by local information security policies.
     + All local site investigators conduct the study appropriately.
     + All non-compliance with the study protocol or applicable requirements will reported in accordance with local policy.

This is a single center study.

* 1. Describe the method for communicating to engaged participating sites:
     + Problems.
     + Interim results.
     + The closure of a study

This is not a multi-site trial.

# Study Timelines

* 1. Describe the duration of an individual subject’s participation in the study.

The project is anticipated to be completed in 1 year from data release.

* 1. Describe the duration anticipated to enroll all study subjects.

Data will be extracted in a one-time transfer.

* 1. Describe the estimated date for the investigators to complete this study (complete primary analyses)

It will take approximately 1 year to complete this study.

# Study Endpoints

* 1. Describe the primary and secondary study endpoints.

The study end point is when the chart corpora are analyzed, statistically studied and interpreted.

* 1. Describe any primary or secondary safety endpoints.

There are no safety endpoints being studied.

# Procedures Involved

* 1. Describe and explain the study design.

Step 1: Charts of 2000 patients will be extracted into a secure database in a one step process.

Step 2: Computational methods for plagiarism detection and text analysis will be applied on the narrative sections of the chart. Obtained results will be processed to quantify “copy and paste’ behavior.

* 1. Provide a description of all research procedures being performed and when they are performed, including procedures being performed to monitor subjects for safety or minimize risks.

We intend to apply formal mathematical and information-theoretic criteria including text unification and winnowing as well as statistical testing. This is an application study not intended to evaluate patient treatments or safety/risk factors.

* 1. Describe procedures performed to lessen the probability or magnitude of risks.

The only perceivable risk is related to privacy. Data will be stored securely at the Institute for Healthcare Informatics (IHI), and all security policies (attached) will be adhered to.

* 1. Describe all drugs and devices used in the research and the purpose of their use, and their regulatory approval status.

This project does not involve drugs or devices.

* 1. Describe the source records that will be used to collect data about subjects. (Attach all surveys, scripts, and data collection forms.)

We intend to collect the charts needed for our study from the UBMD data stored at the Institute of Healthcare Informatics

* 1. What data will be collected including long-term follow-up.

Narrative entries in the chart - vital signs, laboratory data as well as metadata describing the narrative entries’ date and time, author’s information will be collected and stored. There is no long term follow up in our study.

* 1. For HUD uses provide a description of the device, a summary of how you propose to use the device, including a description of any screening procedures, the HUD procedure, and any patient follow-up visits, tests or procedures.

This project does not involve devices.

# Data and Specimen Banking

* 1. If data or specimens will be banked for future use, describe where the data/specimens will be stored, how long they will be stored, how the data/specimens will be accessed, and who will have access to the data/specimens.

The copy of patient level data will be destroyed when the study is completed and the results published.

* 1. List the data to be stored or associated with each specimen.

Only the above described data from the chart will be included. No specimens are involved in our study.

* 1. Describe the procedures to release data or specimens, including: the process to request a release, approvals required for release, who can obtain data or specimens, and the data to be provided with specimens.

Patient data will not be released to any other party.

# Data Management

* 1. Describe the data analysis plan, including any statistical procedures.

Response: We will use a paired samples t-test to compare the average amount of copy/paste activity found using our methods with those found using the methods from previous studies in the literature. This will be computed:

1) daily - the amount of copy/paste on day n (from day n-1, n > 1) will be compared between the two methods of calculation. For this, there is 2000 samples visits 2 and 3. We will calculate the daily difference only for as many days as we have an appropriate amount of data to have statistically meaningful results.

2) globally - instead of stratifying the data by day, we will simply accumulate it all and compare the global means. This will result in at least 2000 samples

* 1. Provide a power analysis.

Response:

First, for the measurement of the difference globally of the two methods, having at least 2000 samples, we will be able to show, with p < 0.05, that there is a significant difference between the two means for mean differences as low as 1 even with relatively large standard deviations of the mean (up to 15.95). We expect to have many more than 2000 samples, making this test even more powerful.

These 2000 samples will allow us to calculate a mean difference of 1 for standard deviations of the mean up to 11.27. Beyond that we will continue to analyze the data as long as the power analysis for paired samples t-test shows that p > .05 for the number of subjects and the standard deviation of the mean.

* 1. Describe the steps that will be taken secure the data (e.g., training, authorization of access, password protection, encryption, physical controls, certificates of confidentiality, and separation of identifiers and data) during storage, use, and transmission.

IHI security procedures (attached) will be followed to ensure that only authorized users will be able to access the data.

* 1. Describe any procedures that will be used for quality control of collected data.

Departmental quality control procedures will be followed.

* 1. Describe how data and specimens will be handled study-wide:

Data will be handled electronically as plain text.

* 1. What information will be included in that data or associated with the specimens?

The following data will be included:

1. The narrative entries of the chart.

2. Laboratory and radiological reports.

* 1. Where and how data or specimens will be stored?

Data will be held at the Institute for Healthcare informatics.

* 1. How long the data or specimens will be stored?

Data will be stored for the duration of the study (approximately 1 year).

* 1. Who will have access to the data or specimens?

Data will be accessed by the PI and those listed on this protocol.

* 1. Who is responsible for receipt or transmission of the data or specimens?

The principal investigator, Edwin Anand, will be responsible for the receipt of the data from the Institute for Healthcare Informatics.

* 1. How data and specimens will be transported?

Response:

# Provisions to Monitor the Data and Ensure the Safety of Subjects

* 1. Describe the plan to periodically evaluate the data collected regarding both harms and benefits to determine whether subjects remain safe.

Data will not be reviewed to evaluate patient health, treatment or safety.

* 1. Describe what data are reviewed, including safety data, untoward events, and efficacy data.

Data will not be reviewed to evaluate patient health, treatment or safety.

* 1. Describe how the safety information will be collected (e.g., with case report forms, at study visits, by telephone calls with participants).

No safety data is collected for this project.

* 1. Describe the frequency of data collection, including when safety data collection starts.

Data collection is a one-time extraction.

* 1. Describe who will review the data.

Data is reviewed and interpreted by the PI and reviewers.

* 1. Describe the frequency or periodicity of review of cumulative data.

Data will be reviewed after extraction, then after application of computing techniques and after application of statistical methods.

* 1. Describe the statistical tests for analyzing the safety data to determine whether harm is occurring.

There are no statistical tests being applied to analyze safety data.

* 1. Describe any conditions that trigger an immediate suspension of the research.

Suspension of the study would occur if the data storage system was compromised.

# Withdrawal of Subjects

* 1. Describe anticipated circumstances under which subjects will be withdrawn from the research without their consent.

Since patient records are extracted, there are no conditions present to remove a case from the project.

* 1. Describe any procedures for orderly termination.

There are no procedures for subject termination.

* 1. Describe procedures that will be followed when subjects withdraw from the research, including partial withdrawal from procedures with continued data collection.

There are no withdrawal scenarios associated with this project.

# Risks to Subjects

* 1. List the reasonably foreseeable risks, discomforts, hazards, or inconveniences to the subjects related the subjects’ participation in the research. Include as may be useful for the IRB’s consideration, a description of the probability, magnitude, duration, and reversibility of the risks. Consider physical, psychological, social, legal, and economic risks.

As identifiable patient data is being extracted, there is minimal risk that patient identities could be released if a violation/damage to the system storing the data occurs. However, stringent procedures for accessing and working with the data will be adhered to including following IHI security policies.

* 1. If applicable, indicate which procedures may have risks to the subjects that are currently unforeseeable.

Security breaches at the IHI could potentially lead to loss of privacy.

* 1. If applicable, indicate which procedures may have risks to an embryo or fetus should the subject be or become pregnant.

This is not applicable to this study.

* 1. If applicable, describe risks to others who are not subjects.

There are no risks to others outside of the patient records.

# Potential Benefits to Subjects

* 1. Describe the potential benefits that individual subjects may experience from taking part in the research. Include as may be useful for the IRB’s consideration, the probability, magnitude, and duration of the potential benefits.

There are no direct benefits to patients.

* 1. Indicate if there is no direct benefit. Do not include benefits to society or others.

There are no direct benefits to others in the project. However, a formal study of “copy and paste’ function could help federal and state regulators, physician societies and EMR developers to draw appropriate conclusions and take actions that would enhance physician notes and thereby the quality of care to patients.

# Vulnerable Populations

* 1. If the research involves individuals who are vulnerable to coercion or undue influence, describe additional safeguards included to protect their rights and welfare.
     + If the research involves pregnant women, review “CHECKLIST: Pregnant Women (HRP-412)” to ensure that you have provided sufficient information.
     + If the research involves neonates of uncertain viability or non-viable neonates, review “CHECKLIST: Neonates (HRP-413)” or “HRP-414 – CHECKLIST: Neonates of Uncertain Viability (HRP-414)” to ensure that you have provided sufficient information.
     + If the research involves prisoners, review “CHECKLIST: Prisoners (HRP-415)” to ensure that you have provided sufficient information.
     + If the research involves persons who have not attained the legal age for consent to treatments or procedures involved in the research (“children”), review the “CHECKLIST: Children (HRP-416)” to ensure that you have provided sufficient information.
     + If the research involves cognitively impaired adults, review “CHECKLIST: Cognitively Impaired Adults (HRP-417)” to ensure that you have provided sufficient information.
     + Consider if other specifically targeted populations such as students, employees of a specific firm or educationally/economically disadvantaged persons are vulnerable to coercion or undue influence. The checklists listed above for other populations should be used as a guide to ensure that you have provided sufficient information.

Data from patients in vulnerable populations will be excluded from the study, therefore there are no risks to these populations.

# Community-Based Participatory Research

* 1. Describe involvement of the community in the design and conduct of the research.

There is no community involvement in this project.

Note: “Community-based Participatory Research” is a collaborative approach to research that equitably involves all partners in the research process and recognizes the unique strengths that each brings. Community-based Participatory Research begins with a research topic of importance to the community, has the aim of combining knowledge with action and achieving social change to improve health outcomes and eliminate health disparities.

# Sharing of Results with Subjects

* 1. Describe whether or not results (study results or individual subject results, such as results of investigational diagnostic tests, genetic tests, or incidental findings) will be shared with subjects or others (e.g., the subject’s primary care physicians) and if so, describe how it will be shared.

There will not be any results per se to be shared with those whose records are reviewed.

# Setting

* 1. Describe the sites or locations where your research team will conduct the research.

Project work will take place in the Department of Biomedical Informatics.

* 1. Identify where your research team will identify and recruit potential subjects.

This project does not require recruitment.

* 1. Identify where research procedures will be performed.

Coding and data evaluation will be done in the Department of Biomedical Informatics.

* 1. Describe the composition and involvement of any community advisory board.

There is no advisory board for this project.

* 1. For research conducted outside of the organization and its affiliates describe:
* Site-specific regulations or customs affecting the research for research outside the organization.
* Local scientific and ethical review structure outside the organization.

This does not apply.

# Resources Available

* 1. Describe the qualifications (e.g., training, experience, oversight) of you and your staff as required to perform their role. When applicable describe their knowledge of the local study sites, culture, and society. Provide enough information to convince the IRB that you have qualified staff for the proposed research. Note- If you specify a person by name, a change to that person will require prior approval by the IRB. If you specify people by role (e.g., coordinator, research assistant, co-investigator, or pharmacist), a change to that person will not usually require prior approval by the IRB, provided that person meets the qualifications described to fulfill their roles.

Edwin Anand is a first-year fellow in Clinical Informatics with a background as a trained Nephrologist. He is currently a fellow in the Department of Biomedical Informatics, with a focus on the role of the narrative text on the modern EMR.

Jaroslaw Zola is an Assistant Professor in the Department of Computer Science and Engineering and Research Assistant Professor in the department of Biomedical Informatics. He is a recognized leader in big data analytics in life sciences, including application of high performance, parallel and distributed computing, machine learning and information theory.

Sashank Kaushik is a Clinical Informatics fellow (Year II) in the department of Biomedical Informatics. He trained as an Internist and practices medicine as a Hospitalist physician at Kaleida Health. Earlier in 2015, he completed Masters in Business Administration at University at Buffalo. His clinical informatics interests include clinical decision support systems, EHR interoperability, population health analytics, HIT’s role in healthcare process improvement, governance and economics.

Daniel Schlegel is a postdoctoral associate in the department of Biomedical Informatics. He received his PhD from the department of Computer Science and Engineering at UB earlier in 2015. Daniel has published several papers on natural language processing (and automated understanding), reasoning systems, and the use of terminologies and ontologies in healthcare and related disciplines.

Jonathan Bona is a postdoctoral associate in the Department of Biomedical Informatics. His research focuses on ontology-based knowledge representation and reasoning in the biomedical domain, and on the analysis, interpretation, and use of healthcare data.

Peter Elkin is considered a leader in the field of Biomedical Informatics and has the credentials, experience and skills to sponsor this project. As chair of the University at Buffalo’s Department of Biomedical Informatics, with over 30 years history of Research and Education in Biomedical Informatics, he has specialized in building biomedical phenotypic and genotypic common data infrastructure design, systems and protocols, creating and implementing bioinformatics standards for big data storage and exchange.

All the study members above have received training and education in clinical informatics so is able to review and clinically interpret the data.

Describe other resources available to conduct the research: For example, as appropriate:

* 1. Justify the feasibility of recruiting the required number of suitable subjects within the agreed recruitment period. For example, how many potential subjects do you have access to? What percentage of those potential subjects do you need to recruit?

Obtaining 2000 patient charts from the IHI is practical and feasible.

* 1. Describe the time that you will devote to conducting and completing the research.

The project should take 1 year from data release.

* 1. Describe your facilities.

The Department has the required computer technologies to apply the de-identification software and secure storage space required to house the data.

* 1. Describe the availability of medical or psychological resources that subjects might need as a result of an anticipated consequences of the human research.

These resources are not required for this project.

* 1. Describe your process to ensure that all persons assisting with the research are adequately informed about the protocol, the research procedures, and their duties and functions.

The protocol has been discussed and developed with contribution from all team members who will be well informed of the protocol and their respective roles will be clarified in our weekly meeting.

# Prior Approvals

* 1. Describe any approvals that will be obtained prior to commencing the research. (E.g., school, external site. funding agency, laboratory, radiation safety, or biosafety approval.)

Apart from consent from the CIO of UBMD, no other approvals are required for this study.

# Recruitment Methods

* 1. Describe when, where, and how potential subjects will be recruited.

A random set of 2000 inpatient records meeting eligibility criteria will be selected.

* 1. Describe the source of subjects.

Clinical data will be drawn from EMR records from UBMD clinics

* 1. Describe the methods that will be used to identify potential subjects.

We will query the EMR database for identifying patient records, which meet the study criteria.

* 1. Describe materials that will be used to recruit subjects. (Attach copies of these documents with the application. For advertisements, attach the final copy of printed advertisements. When advertisements are taped for broadcast, attach the final audio/video tape. You may submit the wording of the advertisement prior to taping to preclude re-taping because of inappropriate wording, provided the IRB reviews the final audio/video tape.)

There are no recruitment materials needed for this project.

* 1. Describe the amount and timing of any payments to subjects.

No payment will be made to any patient whose chart is studied.

# Local Number of Subjects

* 1. Indicate the total number of subjects to be accrued locally.

2000

* 1. If applicable, distinguish between the number of subjects who are expected to be enrolled and screened, and the number of subjects needed to complete the research procedures (i.e., numbers of subjects excluding screen failures.)

All admissions during a calendar year will be screened for eligibility. 2000 charts will be randomly selected.

# Confidentiality

Describe the local procedures for maintenance of confidentiality.

* 1. Where and how data or specimens will be stored locally?

The data will be held at the Institute for Health Informatics (IHI), co-located within the Center for Computational Research (CCR), UB’s supercomputing center. The IHI, funded through a New York State HEAL grant, is designed to provide a secure infrastructure for storing protected health information (PHI). Security features are built to HIPAA standards.

* 1. How long the data or specimens will be stored locally?

Data will be stored for the duration of the study, approximately 1 year.

* 1. Who will have access to the data or specimens locally?

The PI and project staff.

* 1. Who is responsible for receipt or transmission of the data or specimens locally?

Edwin Anand, the principal investigator will be responsible for receipt and transmission of the data described.

* 1. How data and specimens will be transported locally?

Data is to be transferred electronically on a hard drive.

# Provisions to Protect the Privacy Interests of Subjects

* 1. Describe the steps that will be taken to protect subjects’ privacy interests. “Privacy interest” refers to a person’s desire to place limits on whom they interact or whom they provide personal information.

All data will be de-identified as part of the study. After completion of the study, the original identifiable records will be destroyed.

* 1. Describe what steps you will take to make the subjects feel at ease with the research situation in terms of the questions being asked and the procedures being performed. “At ease” does not refer to physical discomfort, but the sense of intrusiveness a subject might experience in response to questions, examinations, and procedures.

This is a retrospective chart analysis. There is no direct patient contact in our study.

* 1. Indicate how the research team is permitted to access any sources of information about the subjects.

De-identified records will be accessed locally or through a VPN, in accordance with IHI security policies (attached).

# Compensation for Research-Related Injury

* 1. If the research involves more than Minimal Risk to subjects, describe the available compensation in the event of research related injury.

Compensation is not part of this study.

* 1. Provide a copy of contract language, if any, relevant to compensation for research-related injury.

Not applicable.

# Economic Burden to Subjects

* 1. Describe any costs that subjects may be responsible for because of participation in the research.

This is a study on documentation quality, hence no burden on the part of any patient is anticipated.

# Consent Process

* 1. Indicate whether you will be obtaining consent

No consent will be obtained for this project.

* 1. Describe where the consent process take place

Not applicable.

* 1. Describe any waiting period available between informing the prospective subject and obtaining the consent.

Not applicable.

* 1. Describe any process to ensure ongoing consent.

Not applicable.

* 1. Describe whether you will be following “SOP: Informed Consent Process for Research (HRP-090).” If not, describe:
     + The role of the individuals listed in the application as being involved in the consent process.
     + The time that will be devoted to the consent discussion.
     + Steps that will be taken to minimize the possibility of coercion or undue influence.
     + Steps that will be taken to ensure the subjects’ understanding.

Not applicable.

**Non-English Speaking Subjects**

* 1. Indicate what language(s) other than English are likely to be spoken/understood by your prospective study population or their legally authorized representatives.

No direct contact with subjects will take place in this project.

* 1. If subjects who do not speak English will be enrolled, describe the process to ensure that the oral and written information provided to those subjects will be in that language. Indicate the language that will be used by those obtaining consent.

Not applicable.

**Waiver or Alteration of Consent Process (consent will not be obtained, required information will not be disclosed, or the research involves deception)**

* 1. Review the “CHECKLIST: Waiver or Alteration of Consent Process (HRP-410)” to ensure you have provided sufficient information for the IRB to make these determinations. Provide any additional information necessary here:

This is a data extraction and application project only. No direct contact with subjects or review of data which could impact their health status or treatment will be reviewed or disseminated.

* 1. If the research involves a waiver the consent process for planned emergency research, please review the “CHECKLIST: Waiver of Consent for Emergency Research (HRP-419)” to ensure you have provided sufficient information for the IRB to make these determinations. Provide any additional information necessary here:

Not applicable.

**Subjects who are not yet adults (infants, children, teenagers)**

* 1. Describe the criteria that will be used to determine whether a prospective subject has not attained the legal age for consent to treatments or procedures involved in the research under the applicable law of the jurisdiction in which the research will be conducted. (E.g., individuals under the age of 18 years.) For research conducted in NY state, review “SOP: Legally Authorized Representatives, Children, and Guardians (HRP-013)” to be aware of which individuals in the state meet the definition of “children.”

Pediatric data will be excluded.

* 1. For research conducted outside of NY state, provide information that describes which persons have not attained the legal age for consent to treatments or procedures involved the research, under the applicable law of the jurisdiction in which research will be conducted. One method of obtaining this information is to have a legal counsel or authority review your protocol along the definition of “children” in “SOP: Legally Authorized Representatives, Children, and Guardians (HRP-013).”

This project will be fully conducted in New York State.

* 1. Describe whether parental permission will be obtained from:
     + Both parents unless one parent is deceased, unknown, incompetent, or not reasonably available, or when only one parent has legal responsibility for the care and custody of the child.
     + One parent even if the other parent is alive, known, competent, reasonably available, and shares legal responsibility for the care and custody of the child.

Parental permission is not required for this project.

* 1. Describe whether permission will be obtained from individuals other than parents, and if so, who will be allowed to provide permission. Describe the process used to determine these individuals’ authority to consent to each child’s general medical care.

Not applicable.

* 1. Indicate whether assent will be obtained from all, some, or none of the children. If assent will be obtained from some children, indicate which children will be required to assent.

Children are excluded from the study.

* 1. When assent of children is obtained describe whether and how it will be documented.

Not applicable.

**Cognitively Impaired Adults**

* 1. Describe the process to determine whether an individual is capable of consent. The IRB sometimes allows the person obtaining assent to document assent on the consent document and does not automatically require assent documents to be used.

This study does not require consenting.

**Adults Unable to Consent**

When a person is not capable of consent due to cognitive impairment, a legally authorized representative should be used to provide consent and, where possible, assent of the individual should also be solicited.

* 1. List the individuals from whom permission will be obtained in order of priority. (e.g., durable power of attorney for health care, court appointed guardian for health care decisions, spouse, and adult child.) For research conducted in NY state, review “SOP: Legally Authorized Representatives, Children, and Guardians (HRP-013)” to be aware of which individuals in the state meet the definition of “legally authorized representative.” The list in the consent template signature section corresponds to the priority list for NYS.

We will not be seeking consent from any patient since this is records based study.

* 1. For research conducted outside of NY state, provide information that describes which individuals are authorized under applicable law to consent on behalf of a prospective subject to their participation in the procedure(s) involved in this research. One method of obtaining this information is to have a legal counsel or authority review your protocol along the definition of “legally authorized representative” in “SOP: Legally Authorized Representatives, Children, and Guardians (HRP-013).”

All study procedures will be conducted in New York State.

* 1. Describe the process for assent of the subjects. Indicate whether:
     + Assent will be required of all, some, or none of the subjects. If some, indicated, which subjects will be required to assent and which will not.
     + If assent will not be obtained from some or all subjects, an explanation of why not.
     + Describe whether assent of the subjects will be documented and the process to document assent. The IRB allows the person obtaining assent to document assent on the consent document and does not routinely require assent documents and does not routinely require subjects to sign assent documents.

No assenting is required for this project.

* 1. For HUD uses provide a description of how the patient will be informed of the potential risks and benefits of the HUD and any procedures associated with its use.

Not applicable.

# Process to Document Consent in Writing

If your research presents no more than minimal risk of harm to subjects and involves no procedures for which written documentation of consent is normally required outside of the research context, the IRB will generally waive the requirement to obtain written documentation of consent.

(If you will document consent in writing, attach a consent document. If you will obtain consent, but not document consent in writing, attach a consent script. Review “CHECKLIST: Waiver of Written Documentation of Consent (HRP-411)” to ensure that you have provided sufficient information. You may use “TEMPLATE CONSENT DOCUMENT (HRP-502)”to create the consent document or script.)

* 1. Describe whether you will be following “SOP: Written Documentation of Consent (HRP-091).” If not, describe whether and how consent of the subject will be obtained including whether or not it will be documented in writing.

There is not a consent process for this project.

# Drugs or Devices

* 1. If the research involves drugs or device, describe your plans to store, handle, and administer those drugs or devices so that they will be used only on subjects and be used only by authorized investigators.

This is not a drug study.

If the drug is investigational (has an IND) or the device has an IDE or a claim of abbreviated IDE (non-significant risk device), include the following information:

* 1. Identify the holder of the IND/IDE/Abbreviated IDE.

No IND/IDE is allowed.

* 1. Explain procedures followed to comply with FDA sponsor requirements for the following:

|  |  |  |  |
| --- | --- | --- | --- |
|  | ***Applicable to:*** | | |
| ***FDA Regulation*** | ***IND Studies*** | ***IDE studies*** | ***Abbreviated IDE studies*** |
| ***21 CFR 11*** | ***X*** | ***X*** |  |
| ***21 CFR 54*** | ***X*** | ***X*** |  |
| ***21 CFR 210*** | ***X*** |  |  |
| ***21 CFR 211*** | ***X*** |  |  |
| ***21 CFR 312*** | ***X*** |  |  |
| ***21 CFR 812*** |  | ***X*** | ***X*** |
| ***21 CFR 820*** |  | ***X*** |  |

Response:

Not applicable.