Tab 1

**Project 1 Deliverables**

* Gain Access to Tabular IO Registry data
* ~~Access to deidentified tabular data~~
* Access to deidentified clinical notes (reach out to Subha)
* Confirm Schema and Cohort Inclusion/Exclusion
  + Confirm Schema
    - understand data dictionary/documentation (what tables and features are available), what are the data types, what are the data formats, how are outcomes recorded (binary, multiclass)
  + Cohort Inclusion/Exclusion Criteria
    - Inclusion Criteria
      * **Patients who received immune checkpoint inhibitors (ICI)**
        + Needed to define irAE risk
      * **ICI is 1st-line, monotherapy** 
        + Reduces confounding by indication
      * **Primary cancer is melanoma** 
        + Melanoma patients tend to be younger, less comorbid, and have a more consistent ICI treatment path (=> more homogeneous population)
      * **Has clinical and treatment data available prior to treatment start**
        + **baseline demographic**

Age, sex, stage, histology, etc.

* + - * + Clinical history

Smoking status, BMI, comorbidities (e.g. COPD, autoimmune disease)

* + - * + Baseline labs

Creatinine, albumin, hemoglobin, WBC, ANC, ALC

* + - * + Cancer details

Histology, stage, line of therapy

* + - * + Derived features

Charlson Comorbidity Index, time from diagnosis to ICI, etc.

* + - * + **& cancer characteristics**
        + labs?
        + ICI start date
      * **Patients with irAE outcome label**
      * **No irAE before ICI start** 
        + Ensures proper temporality for prediction
      * **At least one-follow up, 4 weeks after treatment**
        + Allows for sufficient observation time for irAE to manifest
    - Exclusion criteria
      * Combination therapies (e.g., ICI + chemo, ICI + targeted)
        + Complicates interpretation and increases heterogeneity
      * ICI used after prior therapies (2nd-line+)
        + Prior treatments may affect both irAE risk and biomarker profile
      * Missing treatment start date
        + Can’t assess timing or exposure
      * Non-melanoma cancers (for now)
        + Different cancer types → different risk, biology, outcomes
      * Duplicate records or re-treatment episodes
        + Avoid double-counting risk
      * ICI has to be first-line treatment
      * Patients with no structured data prior to ICI therapy
      * Patients lost to follow-up before irAE could be observed
      * Patients with ambiguous or missing irAE outcome label
      * Patients without **documented follow-up** after treatment
      * Patients without clear documentation of ICI therapy initiation
  + **Action Items**
    - Ask Neil Shah
      * How much observation time is sufficient in order to confirm irAE?
      * How confident are we in the no irAE label
        + Balance of type 2 error, less true negative label more true positive you capture and vice versa
      * Plot the CDF of irAE onsets (time between start and follow-up)
      * What is the distribution of no-irAEs?
    - Should we focus on one drug b/c of confounding by indication bias?
    - Should we summarize comorbidities with index like charleston comorbidity index to reduce dimensionality?
    - Read that one study since it’s an exact copy LMAO
      * Figure out from the documentation what poor performance is, also why it is necessary to control for vars like type of ICI treatment, length of exposure, dosage strength (especially if these are high signal features)
      * How does exposure criteria and response to steroids factor in?
      * Ask to join lab data
      * Figure out what data should be required before treatment start
    - Variables not sure to control for
      * Previous ICI therapies
        + That is a bias, and we need remove that potentially
      * Multicancer or one cancer?
        + Check the number of patients have a secondary cancer
        + If not a lot don’t worry abou it,

At least 300 patients to be good

Check distributions of drugs

Check manuscript (categorized treatments, pd-L1, Chemotherapy + ICI)

* Don’t isolate one, but use as a predictor

What is the existing prognostic to predict irAE

* Include all of those in the model

CHeck time from ici start to first irae onset time and then time from last followup

Include features (especially categorical cause you don’t know how to imput) to have less than 30%

* Data Cleaning and Filtering
  + Handle missingness, data types, scaling continuous variables and one-hot encoding categorical variables
  + Filter out post-treatment data
  + Define outcome label - presence/absence of irAE
  + Validation Split - Combine old cohort and new cohort data and then create validation split
  + Action Items
    - Figure out how to combine the old cohort and new cohort data
    - Need documentation for both
      * Figure out the differences from the documentation
      * Missing features, different names,
* Exploratory Data Analysis
  + Visualize class balance, feature distributions, and outliers
  + Identify most relevant features (e.g., labs, comorbidities, demographics)
* Modeling (Baseline and Advanced)
  + Train 9 models (same as Yili’s paper)
  + Tune hyperparameters and compare model variants
  + Address class imbalance (resampling, class weighting, etc.)
* Evaluation & Interpretation
  + Assess metrics (AUC, precision/recall, F1, calibration)
  + Analyze feature importance (e.g., SHAP)
  + Ensure clinical relevance of top predictors
* Deliverables & Documentation
  + Produce plots/tables of performance metrics
  + Create summary table of top features and model interpretation
  + Write model summary + limitations + recommendations

Questions

* Any ICI therapy?
* How to handle missingness in tabular data?
* How to verify tabular data is **before treatment?** 
  + Are there any sources that are after
* For the modeling
  + Do all the models need to be clinically interpretable (XGBoost)
  + Meaning you can run feature importance on them?
    - SHap values

Assumption that bad effects after treatment are treated as irAEs

Assumption that non-irAE bad effects discovered in the treatment is low/no probability

Some models might perofmr better (Since far more categorical data)

Project 2 Deliverables

* + **Action Items**
    - Wait to hear back from Subha deidentified clinical notes
    - How will model learn irAE specific terms
      * How to deal with low frequency irAEs in clinical notes

**Questions for Yili**

* **Where did you get Kader\_njs abstract from**

**Check Which columns have less than 70% missingness**

**Within those features**

* **AUROC score of them correlated with the adverse events**

**Automate that**

**FIrst stick to filter hte main question (yes/no)**

* **See if prevalent, or not a lot of missingness, and hten**
* **Move down to hte individual questions**

**Check the time period from the start of their immunotherapy to their first adverse event**

**Criteria**

* **irAEs**

**-AUROC curves**

**-Plots**

**Prevalence of different comoborbitids, disease histories**

**Present the columns that they might be interested**

Project 1

Project 1

* Goal: predict irAE occurrence
  + Validation cohort - Hackensack data set
* Known predictors of irAE
  + poor performance status, elevated LDH, Response to steroids

deidentified using the De-ID software

Project 2

* How to deal with low frequeqncy irAE (MSK)
* High-toxicity, more relevant (

Need to study Samir Gupta’s paper to figure out what the ground truth was he used and what NLP methods he used

Also if he used an LLM and if so how did he maintain HIPAA compliance

Short-term solution

* Create a clinical tool that is basically an advanced query search to filter sentences utilizing irAE’s, doesn’t have to be an exact match,
* Can have built-in feature for certain expert clinicians to validate the search results and further improve the model

**Addressing Lack of Sentence-Level irAE labeling**

“In this study, machine learning models were trained to predict the occurrence of immune-related adverse events (irAEs) using only unstructured clinical notes, while the ground truth labels indicating whether a patient experienced an irAE were assigned at the patient level through comprehensive clinician chart review. This introduces a classic **label–input mismatch**: the model attempts to infer a label that was derived from **multiple sources of evidence** (including structured data such as lab values, diagnoses, and imaging) while being exposed only to a **subset of those data—namely, filtered clinical text**. This mismatch creates potential risks, most notably in the form of **false negatives**, where a patient is correctly labeled as irAE-positive based on information unavailable to the model, such as a radiographic report or lab test interpreted by a clinician but not explicitly discussed in a note. Such a gap between input and label can undermine model performance and confidence in its predictions.

However, this risk is substantially mitigated in the case of irAEs due to their **unique clinical nature**. Unlike many adverse events that can be algorithmically flagged from abnormal labs or diagnosis codes, irAEs are often **diagnosed through clinical reasoning** and attributed **directly by providers** in narrative notes. These events require contextual interpretation—considering timing, treatment history, and response to interventions like corticosteroids—all of which are typically captured in **free-text documentation**. As a result, the most informative and definitive evidence of irAEs is often found in unstructured notes rather than structured fields. This makes clinical notes a particularly **appropriate and high-yield modality** for irAE identification, and helps explain why the models in this study—despite the apparent mismatch—were still able to achieve strong predictive performance.”

Inclusion Exclusion Criteria

# Inclusion Exclusion Criteria (var\_name)

1. **Primary cancer type = melanoma (currentcancer = ‘Melanoma’ multiple\_primary)**
   1. Melanoma patients tend to be younger, less comorbid, and have a more consistent ICI treatment path (=> more homogeneous population)
2. **Patients who have received ICI therapy previously (prevrx\_io\_therapy = ‘No’)**
3. **Has all desired clinical treatment data available prior to treatment start**
   1. Has to have all since very difficult to impute categorical data (which most of the dataset consists of)
   2. Including ICI start date **(io\_dose1date)**
4. **No irAE before ICI start (irae\_onset > io\_dose1date)**
   1. Confounding by indication if include this subset because inherently healthier folks if they endured at least one ICI cycle already
5. **At least one follow-up, 4 weeks after treatment (do\_followup > io\_dose1date + 4weeks)**
   1. Allows for sufficient time to be confident in ‘no-irAE’ label

Exclusion Criteria:

1. **Duplicate records or re-treatment episodes**

## Plots of interest

1. Bar chart: Proportion of melanoma patients out of primary cancer subtypes
2. Stacked Bar Chart: Monotherapy, first-line
3. Pie chart: Received ICI Therapy previously vs not
4. Pie Chart: received irAE Before ICI Start vs Not
5. CDF: irAE onset (first irAE date - ICI start time)
6. AUROC scores for each feature

Next Steps: get actual new cohort set

Variables of Interest

# Selecting Variables of Interest

Initial Filtering

* Identifier: No
* No Field Annotation: Descriptor

Maybes:

* Comorbidities like HIV
  + Should keep treatment, treatment efficacy, and stage of disease to ascertain the degree to which the comorbidity is affecting the immune strength of the patient
  + Or is it sufficient to say yes/no to whether they have it
* Subtypes of Melanoma
  + Important?
* Comorbidities
  + Yes, will summarize with CCI
  + But should I include specific info on
    - autoimmune diseases like
      * RA, IBD, lupus, psoriasis, etc.
    - Comorbidities tied to immune function or fraility
      * e.g., diabetes, COPD, CKD, CHF, MI
  + Immunosuppressive meds?

Old Data Dictionary VOI Mismatches:

* bal (Baseline albumin (BA) level)
* cps (Combined positive score (CPS))
* dmmr (Deficient mismatch repair (dMMR) status)
* rmh\_score (RMH prognostic score rmh(RMH score = 0-1 good prognosis, RMH score = 2-3 poor prognosis))
* tmb (Tumor Mutational Burden)
* tps (Tumor Proportional Score)

Engineered VOI

* Create date of first irAE onset from earliest Date of Onset from irAE choices
* Create Categories of ICI drugs from Drug name
* Create Charleson Comorbidity Index (CCI) from the Comorbidities list
* Current Cancer and other cancer (needs to be grouped cause there’s too many of them)

Questions to ask of Physicians tomorrow:

Clinical Question to ask of Physicians tomorrow:

* Questions about Inclusion/Exclusion Criteria
  + Should we only consider primary cancer = melanoma?
    - ~~What about cases with multiple primary cancers?~~ (column is empty, so don’t worry about it)
    - What about cases with metastatic disease present before IO therapy (met\_iotx)
    - What about patients with a history of previous cancers
    - Does it matter whether IO treatment was discontinued? (tx\_disc\_yn)
* Questions about Variables of Interest
  + Want to include labs?
    - Which ones are the most important
  + Is my outcome variable correct? (io\_toxicity, field label = “Did the patient experience any I-O toxicities?” cause couldn’t non-immune-related toxicity occur?
  + Was patient on antibiotics/antivirals at the start of IO therapy?
  + Was the patient previously treated with an Immuno-therapy?
* Questions about good predictors for irAE occurence
  + What are the current prognostic tools to predict irAE occurrence
  + What are the factors in the known literature that contribute irAE occurrence
  + I know you said response to steroids can be a predictor for immune system involvement, and therefore can be a predictor for immune-related adverse events. However, what would the variable be in the dataset that would represent the response to steroids?
  + Does the strength of IO therapy dosage not vary? Why is there no variable for the strength of the dose?
  + Good to drop these vars due to missingness?
    - If active, does this patient currently take steroids (aid\_steroids)
    - Is this patients autoimmune disease active? (aid\_active)
    - Date of Progression (progression\_date)
    - Does this patient have disease progression as per RECIST? (progression\_recist\_yn) high AUROC
    - Please specify which other cancer (prev\_cancer\_history\_other)
    - Upper limit of normal (ULN) of LDH (brainmets\_baseline\_ldh\_uln)
    - Baseline LDH (brainmets\_baseline\_ldh)
    - If tested, report the percent of PD L -1 expression (mut\_pdl1)
    - If other, specify the comorbidity (other\_comorb)
    - How many members of patient's family have a history of cancer? (num\_cancer\_fam)
    - Does this patient have previous history of cancer prior to the current cancer that IO is given for? (prev\_cancer\_hx)
    - Pre treatment BMI (pre\_bmi\_calculated) High AUROC
  + **Would any of the variables in AUROC graph serve as good predictors based on domain knowledge**

Tab 6

Bevacizumab

Level of Dose of Steroids

* Sddsdsd
* High dose vs Low dose
* Topical vs IV steroids vs Peels
* Check Immune related guildelines for adverse events
  + CDC pdf
  + Grades 1 - 4
    - Grades 2,
* High dose
* University of New York
  + Blood of antibodies to predict antibodies to predict irAEs
  + C3, c4 elements
  + Join lab data

Immune-related adverse events are permanent for life

* So then Doesn’t matter steroids?
* Glands are not treted with steroids
* Adrenlitis, autoimmunite
* Grade
* Biologic treatments
* Check published work on steroids and imune supresive agents in IO data agents
  + Join lab data

Reasons to prescribe steroids

* Brain metastasis
* Organ transplants

Metastatic sites as proxy

* Schedule Shaked Lev Ari

**Update plot**

* **To include only vars only available at baseline**
* **Rename variable**

**Send via Email**

**The impact of immunosuppressive agents on immune checkpoint inhibitor efficacy in patients with advanced melanoma: A real-world, multicenter, retrospective study**

**mailto:mail4shaked@gmail.com**