

Discover Data@Duke Clinical Data Assets and Tools

DEDUCE™ is a self-service, web-based query tool with an intuitive, user-friendly interface that allows for flexible searching of data hierarchies compiled through years of patient care. Users with no previous database experience can identify cohorts of patients and extract data while maintaining patient privacy.

The DEDUCE™ data set currently covers over 3.4 million patients, spanning 37 years. The data are refreshed daily with updates provided from a combination of current Epic Clarity and legacy EDW data elements. The data include more than 1.6 billion lab results crossing over 10,000+ socio-economic data elements based on U.S. census data. DEDUCE™ enables the quantification of potential study subjects at Duke based on varying inclusion and exclusion criteria available in clinical records. It can be used to estimate potential recruitment.

DEDUCE™ currently includes clinical information systems that serve DUHS and tools for

- Real time tracking of potential research participants (DISCERN)
- Geo-mapping of patients using DEDUCE Geo
- Text analytic searches of radiology and pathology reports

With DEDUCE™ you can also search clinical notes, including history and progress notes, operative notes, ED provider notes, consults, and discharge notes (inpatient and outpatient)

If you:

- Are looking for data that predates Maestro Care implementation
- Need discrete data or data contained in notes
- Need simple aggregate counts or fuller limited data sets
- Need geographic mapping of patients
- Need text analytics of radiology and pathology notes

then DEDUCE™ is the right tool.

DEDUCE™ Use Cases

Identifiable Dataset: Dr. Shepherd's clinical research coordinator is looking for a cohort of patients that may be eligible for a study Dr. Shepherd is conducting. She needs demographic and other information so she can reach out to them via email or phone call with information about the study. Data stays in PACE environment.

Review Preparatory to Research (Limited Dataset): Dr. Shepherd wants to do a retrospective chart review of all patients seen at the Southpoint Primary Care Clinic who are diabetic, under 50, and taking metformin.

Retrospective Data Analysis: Dr. Shepherd wants to know the HbA1C results for elderly, Hispanic or Latino, patients diagnosed with T2DM but not hypertension who had an outpatient visit between 7/1/2018 and 12/31/2018.

DISCERN: Dr. Shepherd's research coordinator wants to build a DISCERN report to be notified in real-time when the patients in the cohort she previously developed are coming for appointments in the following 4 weeks.

Retrospective Data Analysis: Dr. Shepherd is conducting a retrospective research project on patients with Addison's Disease. Because of the rarity of the condition, Dr. Shepherd wants as large of a cohort of patients as possible, so she is looking at all patients seen at Duke from 2007 – 2018. She first needs the ability to obtain a count of patients with the disease to determine if the study is feasible. If feasible, she will need a dataset with a large number of data elements for all patients diagnosed with Addison's Disease dating back to 2007.

Retrospective Chart Review: Dr. Allen is assisting on an NIH grant-funded study that involves a large number of manual chart reviews. The funding is tight and Dr. Allen does not have the hours in his schedule to devote to the large number of chart reviews necessary for the study. He is looking for an easier way to obtain progress notes for patients admitted to DRH in 2017 who were diagnosed with CVD. Preferably he would like to import the Progress Notes to PACE and have his statistician pull relevant information out using a Python program.