Children's Hospital Colorado

HDAA HEALTHCARE DATA & ANALYTICS ASSOCIATION

NIST INFORMATION TECHNOLOGY LABORATORY

An Infrastructure for Secure Sharing of Clinical Data

Joshua Roberts and David Ferraiolo
National Institute of Standards and Technology

Analytics with Altitude

2023 HDAA Conference

Denver, Colorado

AGENDA

- Overview: Secure Federated Data Sharing System (SFDS)
- SFDS value in performing clinical research
- Demonstration

SFDS Overview

Analytics with Altitude

2023 HDAA Conference

Denver, Colorado

Children's Hospital Colorado

HEALTHCARE DATA & ANALYTICS ASSOCIATION

Problem

How can organizations securely share data?

- The ability to share database resources among collaborating organizations is highly desirable – this is especially true in the performance of clinical research.

- However, challenges persist regarding interoperability in the exchange of resources among organizations and preservation of organization's distinct protection policies.

- Hard for users in different organizations to share DBMS data. Because the data is
  - from different systems,
  - in different formats,
  - organized under different schemas
  - protected under the host's access control policies.

Solution

Exchange attributes not data

- A standard means of providing policy-preserving access to the data where it currently resides, rather than its exchange, or centralized storage.
- Transparent to the otherwise normal business operations of participating organizations.
- Accomplished using two NIST developed technologies
  - **Data Block Matrix (DBM) –** Verify user's attributes across a federation of organizations
  - **Next generation Database Access Control (NDAC) –** Control access to SQL databases with cell level access control
- Through consent, previously unknown users are onboarded into local **NDAC** systems using their **DBM** validated attributes, allowing them policy preserving access to local database resources.

Data Block Matrix (DBM)

A permissioned distributed ledger with the ability to edit or delete data.

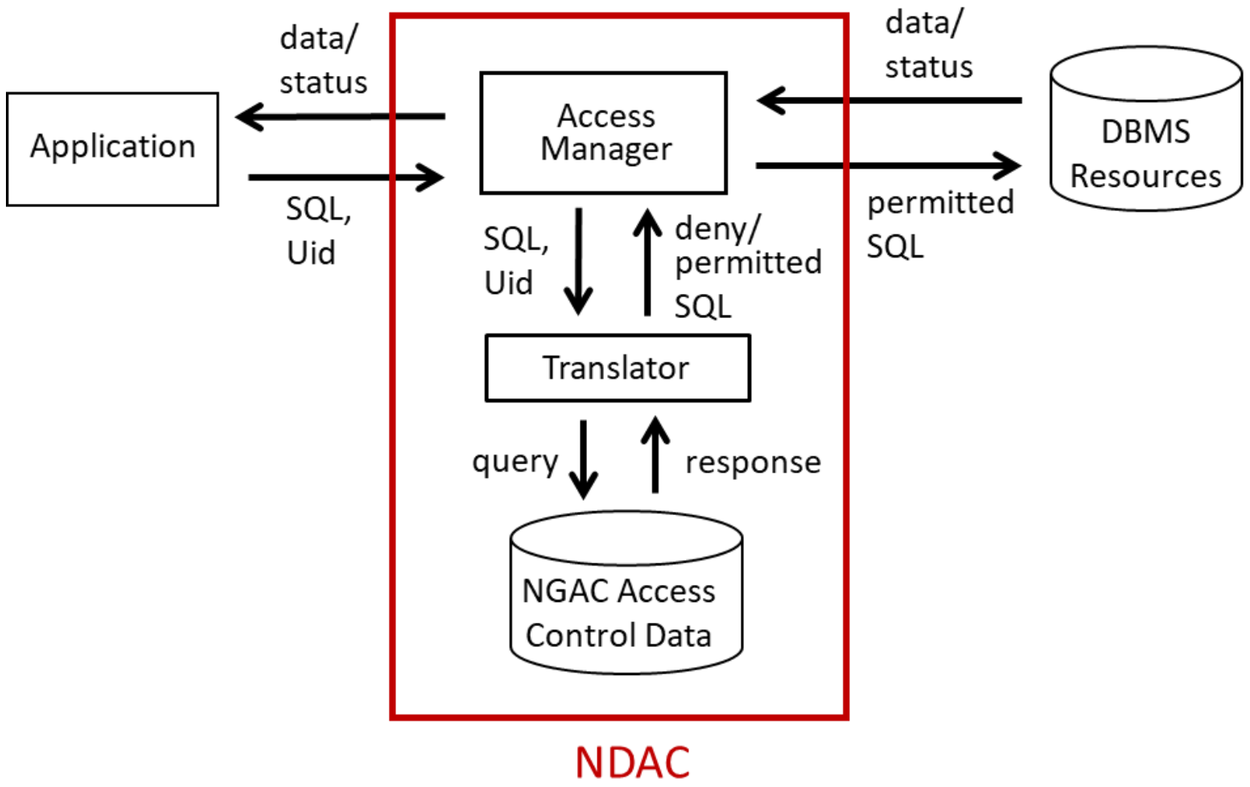
- A permissioned DLT, with the hashed data integrity protection of a blockchain, but with the additional ability to **edit or delete data**.
- Manage and verify user attributes for all federation users.
- A **standard catalog** of attributes agreed upon by the federation is used in their local **NGAC** policies (i.e. SNOMED).
- NGAC governance policy controls who can update user attribute assignments.

	0	1	2	3	4	
0	•	1	3	7	13	$H_{0,-}$
1	2	•	5	9	15	$H_{1,-}$
2	4	6	•	11	17	$H_{2,-}$
3	8	10	12	•	19	$H_{3,-}$
4	14	16	18	20	•	$H_{4,-}$
	$H_{-,0}$	$H_{-,1}$	$H_{-,2}$	$H_{-,3}$	$H_{-,4}$	etc.

Editing or deleting block **12** only requires the update of row hash $H_{3,-}$ and column hash $H_{-,2}$

Next generation Database Access Control (NDAC)

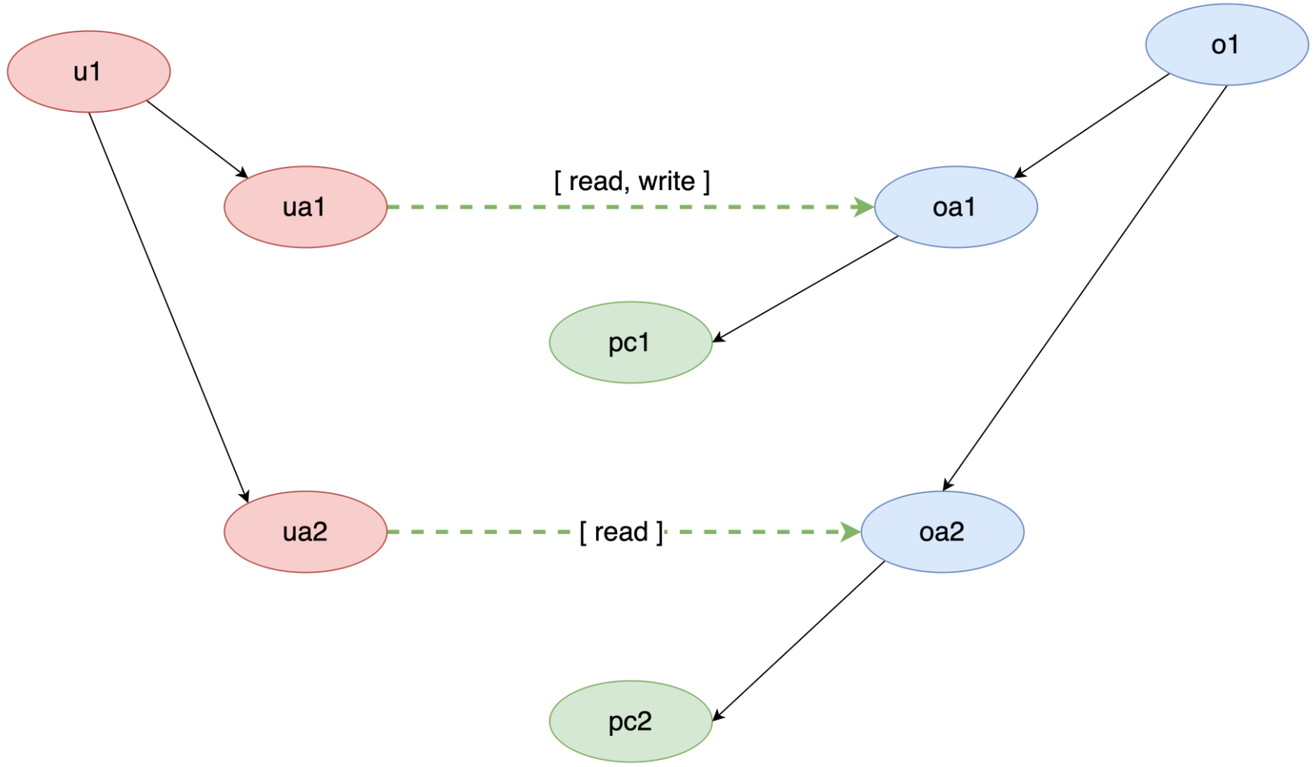
- Uses **Next Generation Access Control (NGAC)** to **translate** SQL statements into permitted SQL.
- DBMS and application agnostic.
- Enforce policies at a granularity not typically available to DBMSs.
- NGAC policy review allows access state to be cached, speeding up translation at query time.
- Performance: ~3 seconds per 1 million rows when caching.



Next Generation Access Control (NGAC)

ANSI/INCITS ABAC Standard

- **Privileges** are represented as a Directed Acyclic Graph (DAG).
  - Based on user attribute assignments, resource (objects) attribute assignments, and other relations specified in terms of those attributes.
- **Prohibitions** supplement policy graph and deny permissions for users on objects.
- **Obligations** dynamically change access state in response to policy and external events.



User **u1** has **[read]** access rights on **o1**

Access Control Policies

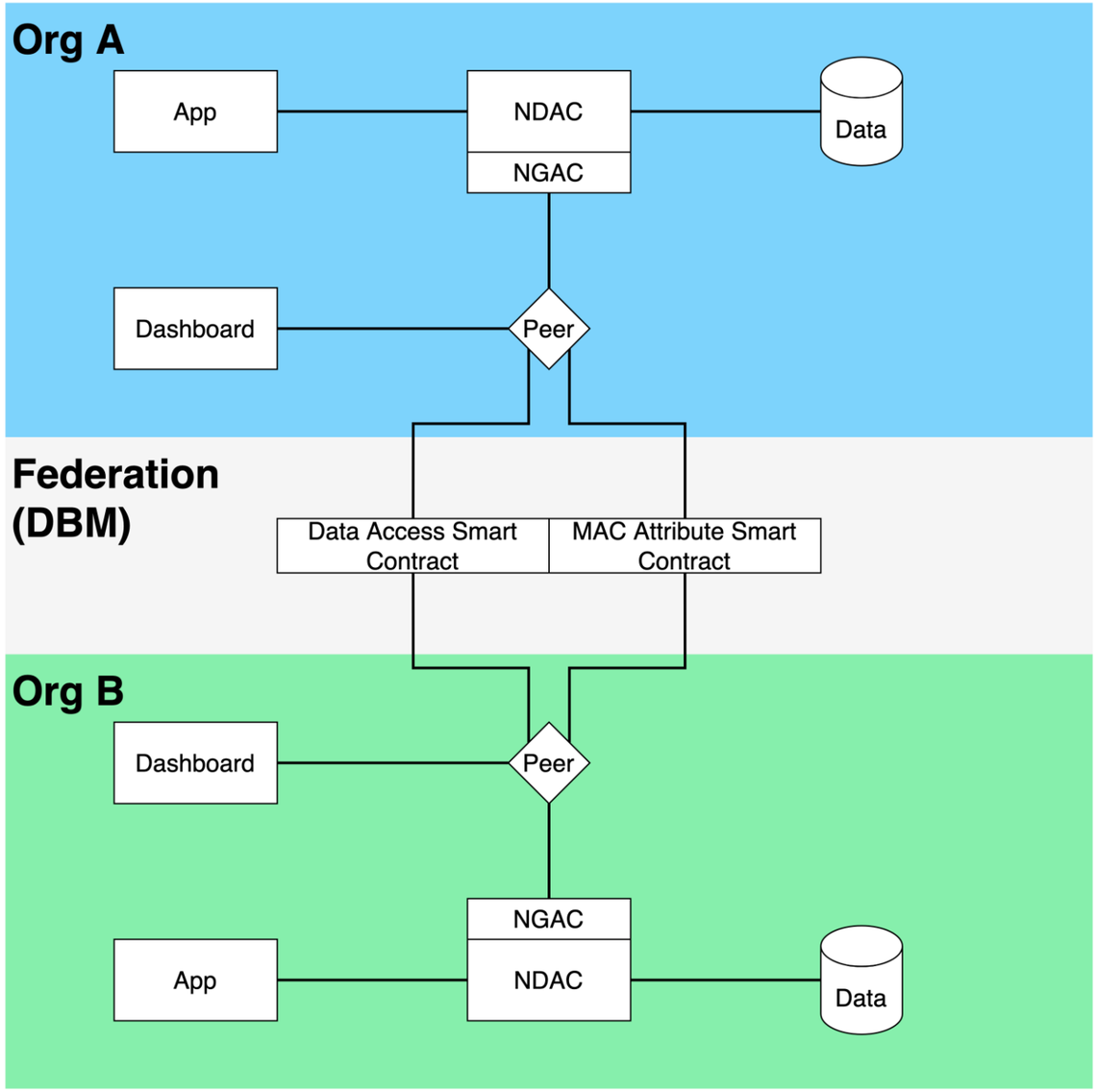
- **Discretionary Access Control (DAC)** - Provide organizations with capabilities to grant or prohibit DBM users access to their data sets.

- **Mandatory Access Control (MAC)** - Policies impose non-discretionary rules on users when accessing resources for the enforcement of mandated policies or regulation based on roles or other types of attributes.

- NGAC enforces policy combinations.

SFDS Architecture (2 Orgs)

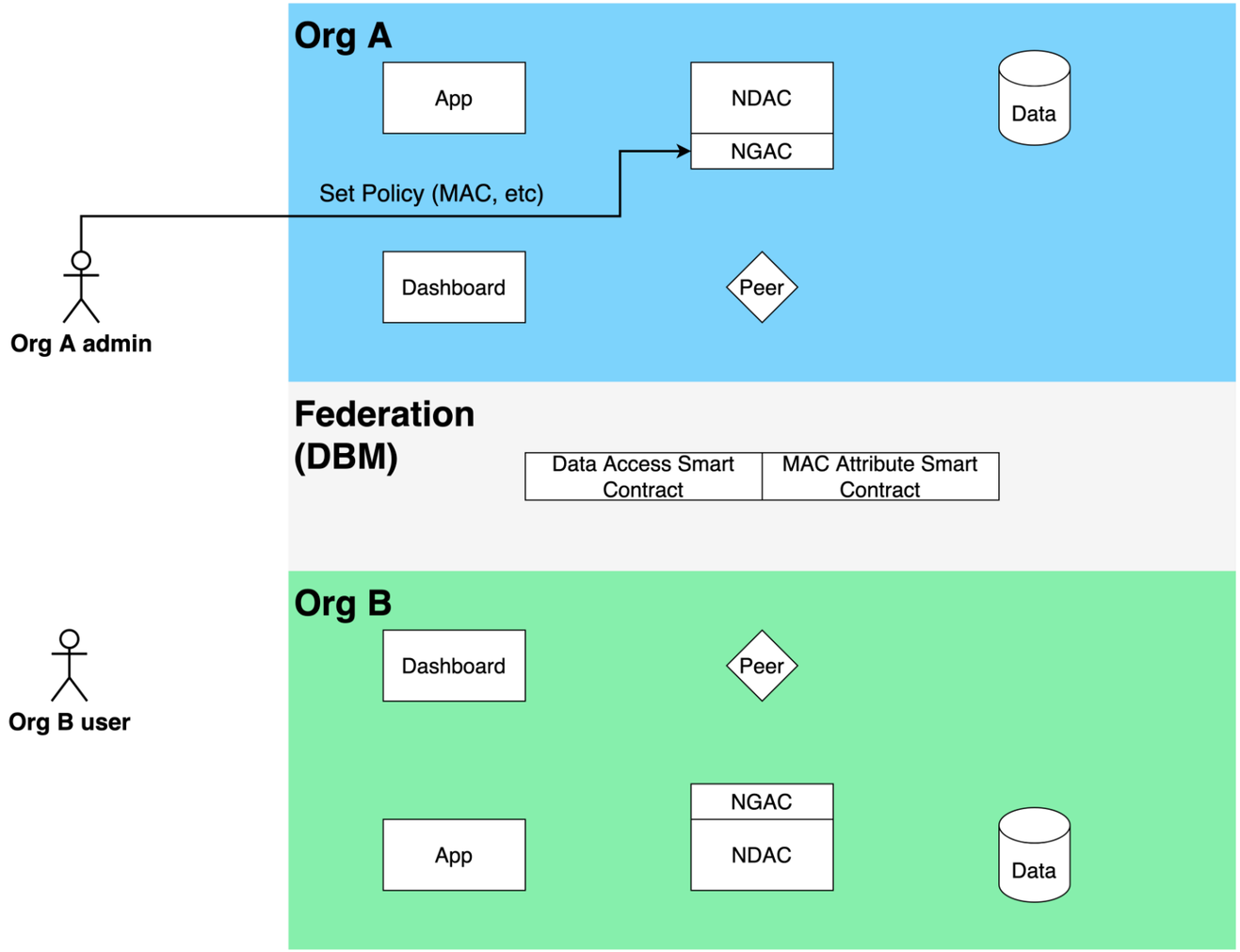
- Data access application stack (**App**, **Data**)
- An **NDAC** server with **NGAC** policy to protect the data
- **Dashboard** to request/grant access to local resources and **Peer** node to interact with DBM smart contracts



1. Set initial NGAC policy

Scenario: Org B user accesses Org A data

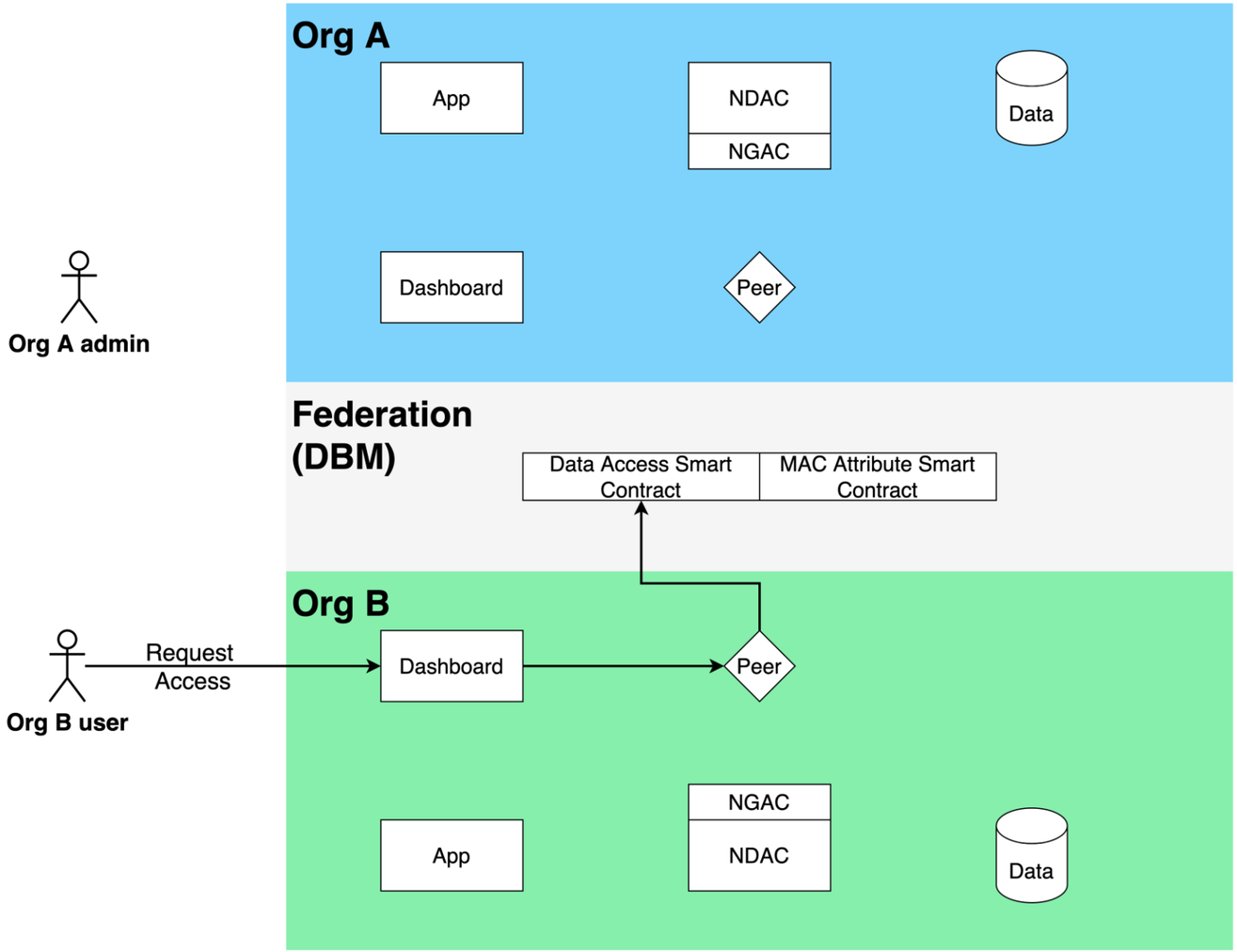
- Org A admin sets initial NGAC policy using, but not limited to, the attributes defined in the standard catalog



2. Request access

Scenario: Org B user accesses Org A data

- Org B user requests access to Org A's data using the Org B dashboard which invokes the Data Access Smart Contract



3. Set DAC policy and onboard user

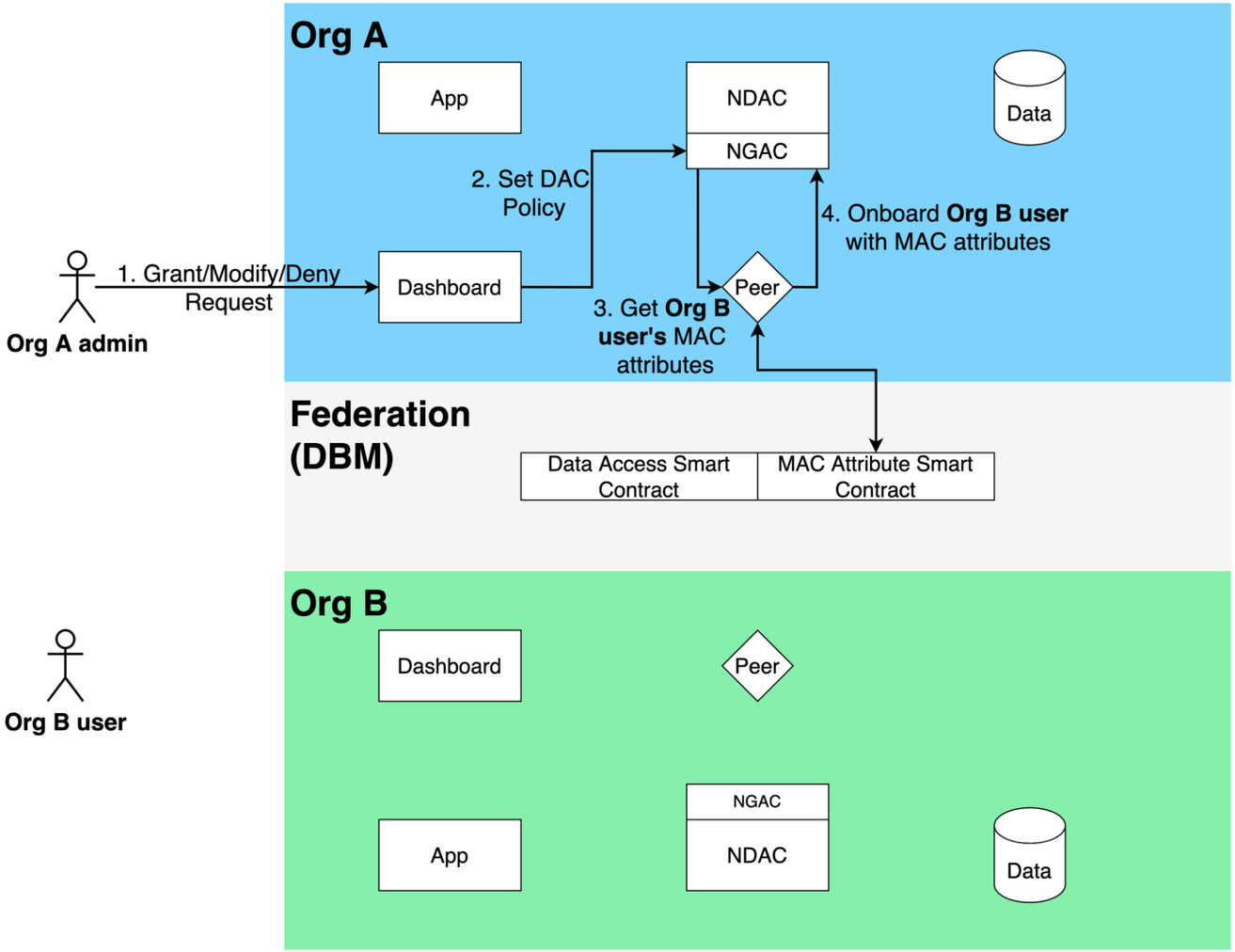
Scenario: Org B user accesses Org A data

1. Org A admin grants/modifies the request in Org A's dashboard.

2. Org A admin sends the required **DAC** policy to NGAC.

3. NGAC retrieves the MAC attributes from the MAC Attribute Smart Contract.

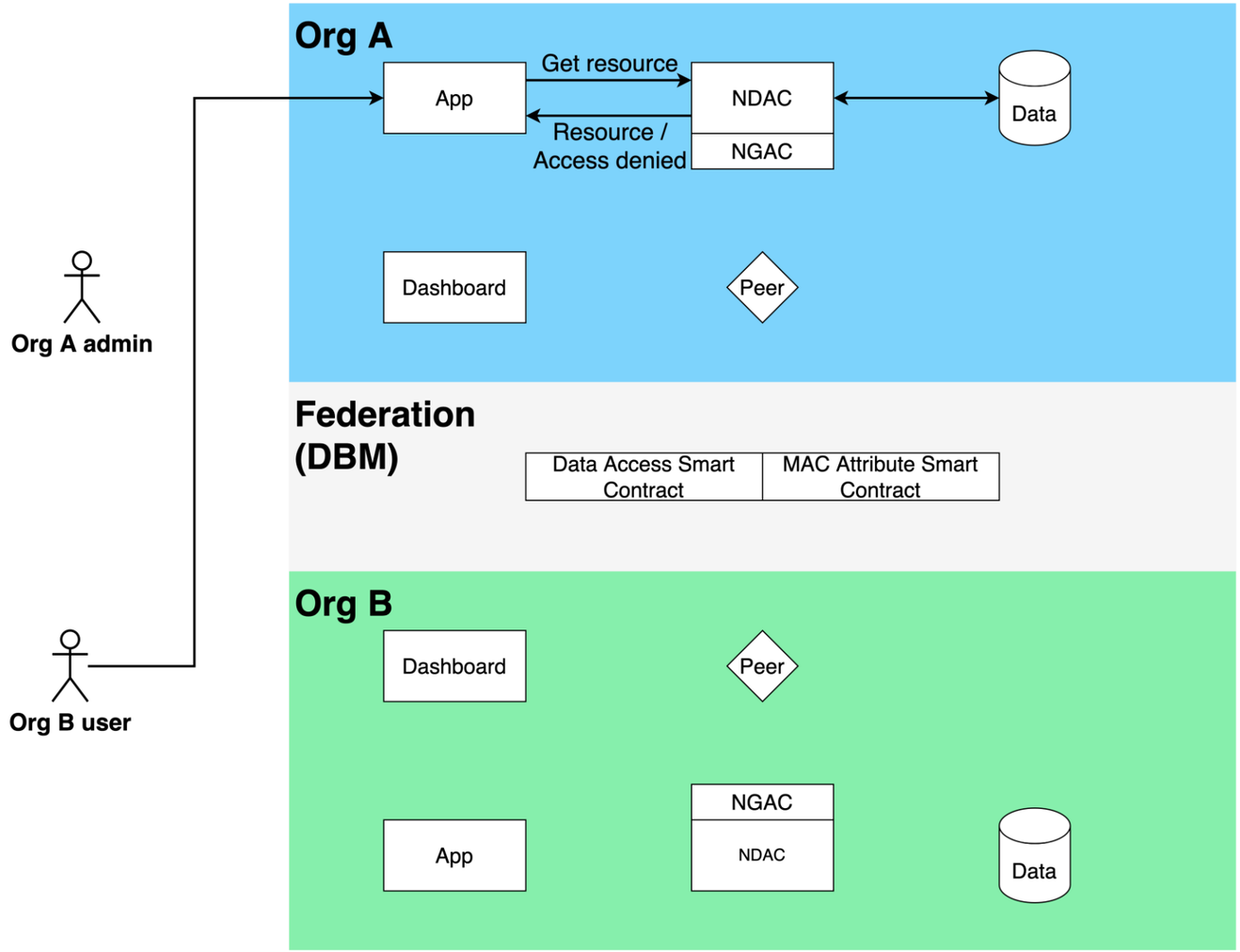
4. NGAC assigns the user to the matching **MAC** attributes in the local policy graph.



4. Access Data

Scenario: Org B user accesses Org A data

- Org B user accesses Org A's data the same as a regular Org A user
- Organizations have control over the authentication of external users that access their data



SFDS value in performing clinical research

Analytics with Altitude

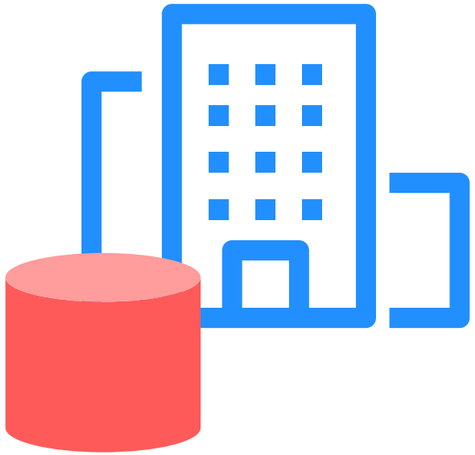
2023 HDAA Conference

Denver, Colorado

Children's Hospital Colorado

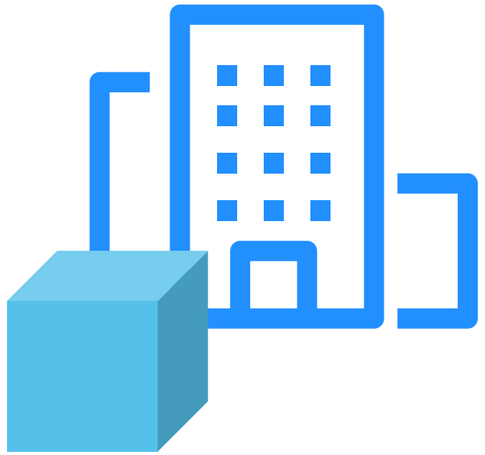
HDAA
HEALTHCARE DATA & ANALYTICS ASSOCIATION

Clinical Data Sharing Challenge



Institution A

Has a very large dataset with **few restrictions**.



Institution B

Has a **more restrictive posture** with sharing due to the nature of some of the population served.

Institution C

Has significant **research expertise** and resources.

The Participants

- Technologists
- Administrators
- Researchers
- Clinicians
- Subject Matter Experts
- Governance Bodies
- Assistants/Support
- Students

Likely to Vary Across Institutions

Some May Serve Multiple Roles

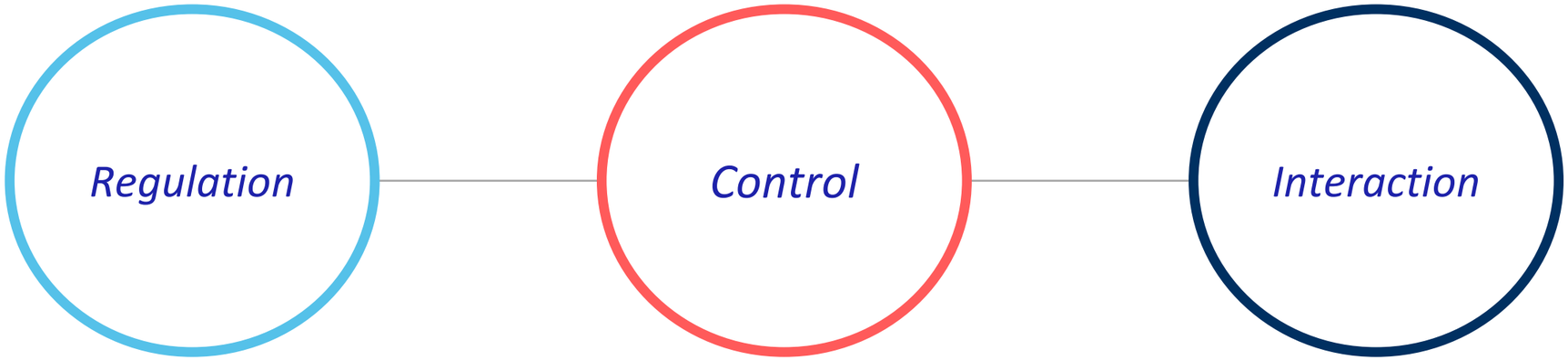
Participants May Change Over Time

Barriers to Data Sharing

- Potential misuse or misinterpretation of data
- Lack of resources
- Loss of control
- Socio-cultural factors and ethical and legal barriers

Devriendt T, Borry P, Shabani M (2021)
Factors that influence data sharing through data sharing platforms: A qualitative study on the views and experiences of cohort holders and platform developers.

Data Sharing Questions



What are we required to do?

Are we sure it is protected?

How do we provide access?

Demonstration

Analytics with Altitude

2023 HDAA Conference

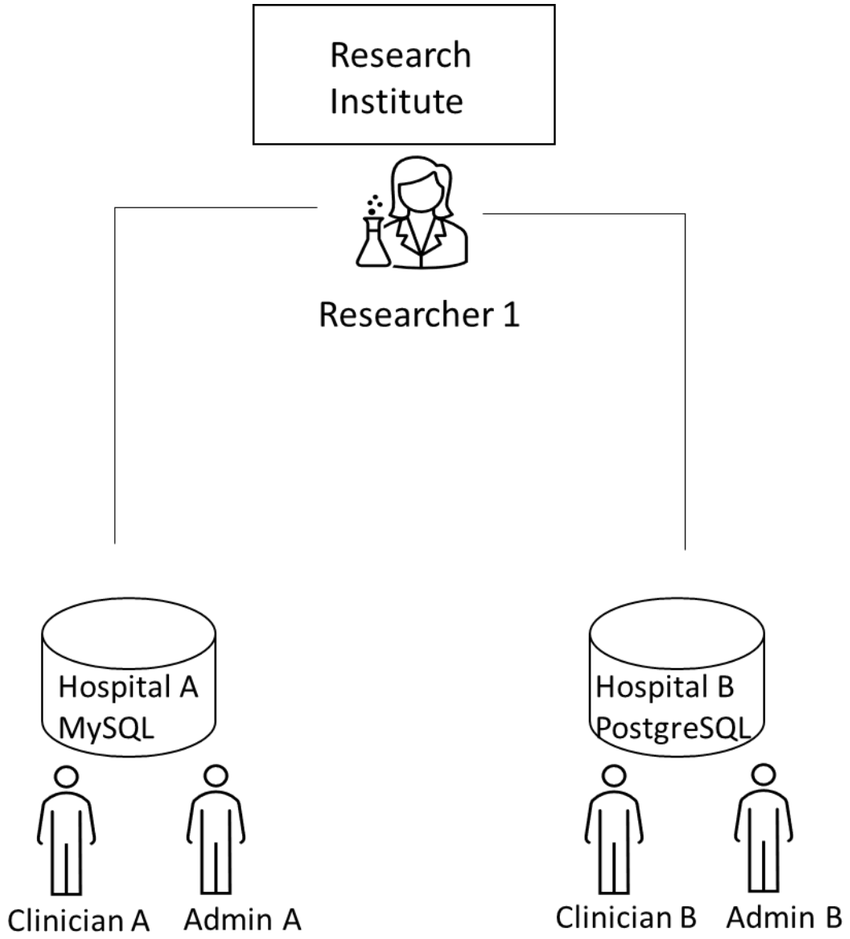
Denver, Colorado

Children's Hospital Colorado

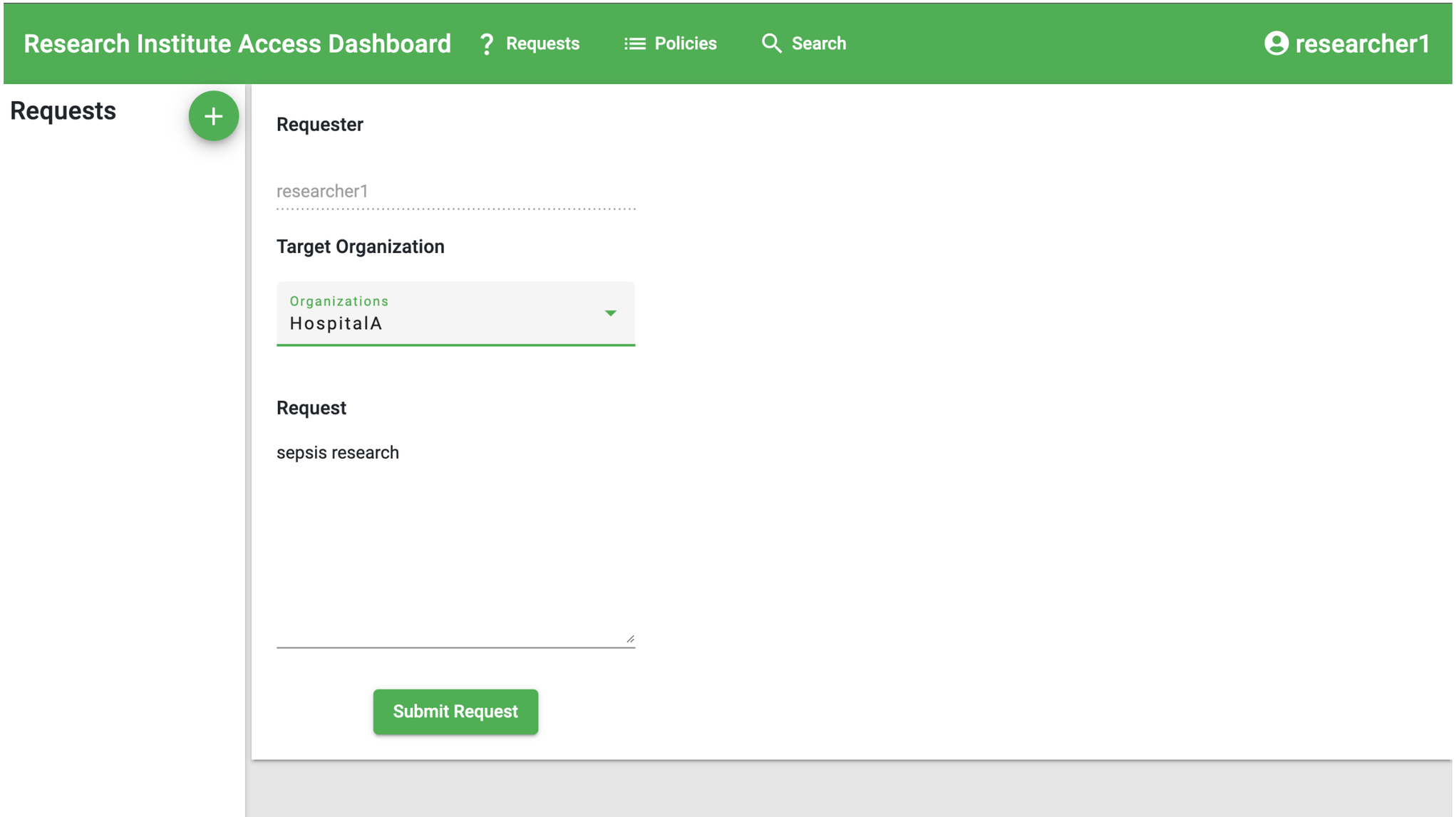
HDAA
HEALTHCARE DATA & ANALYTICS ASSOCIATION

Use Case Scenario

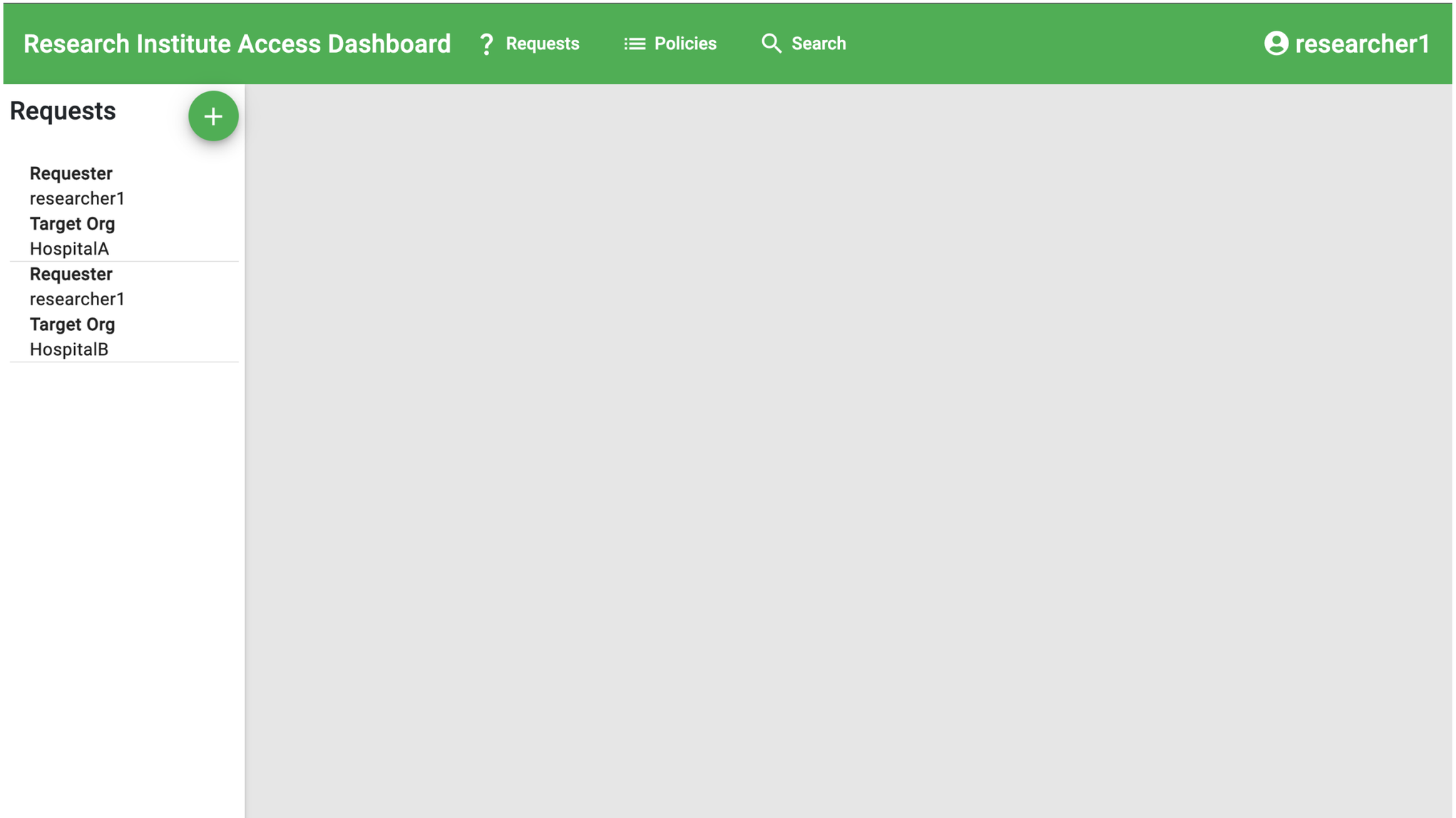
- **Researcher 1** wants access to clinical data at **Hospital A** and **Hospital B** to perform an analysis of the variables and indicators of patients at risk of **sepsis**.
- Hospital A and Hospital B restrict access to patient **PII** and data related to **mental health** (MAC policy).
- Hospital B considers substance abuse related to mental health, while Hospital A does not.
- Both hospitals will create similar DAC policies which grant Researcher 1 access to **patient identity** and **financial information** for only sepsis patients.



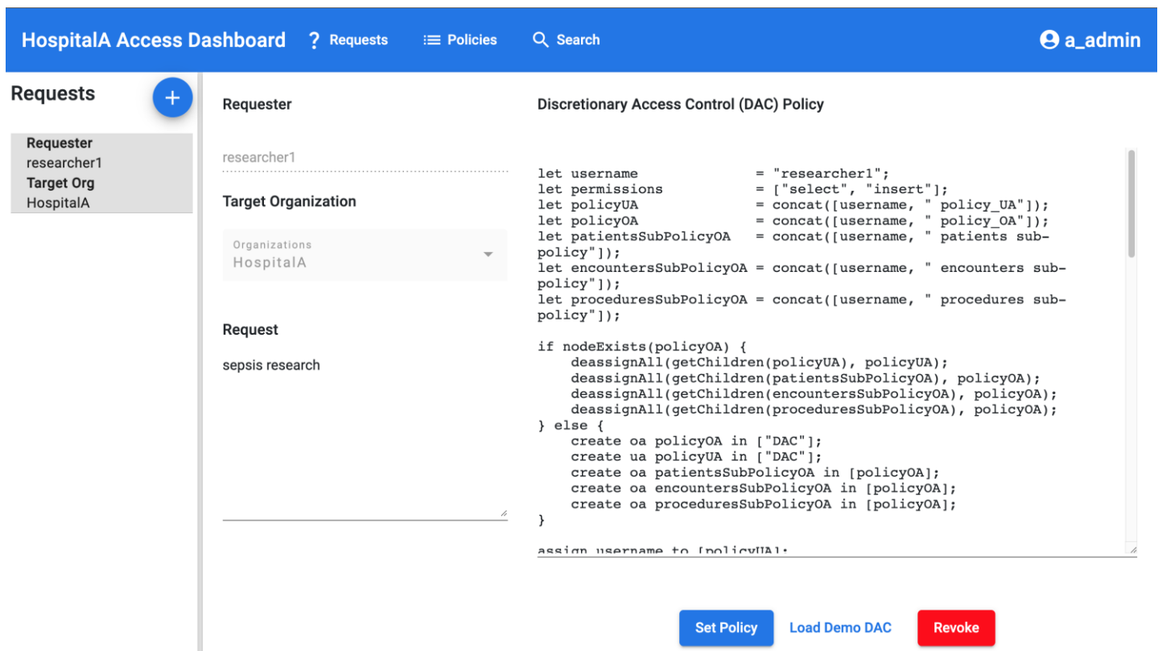
1. **Researcher1** requests access to data at **HospitalA** and **HospitalB**

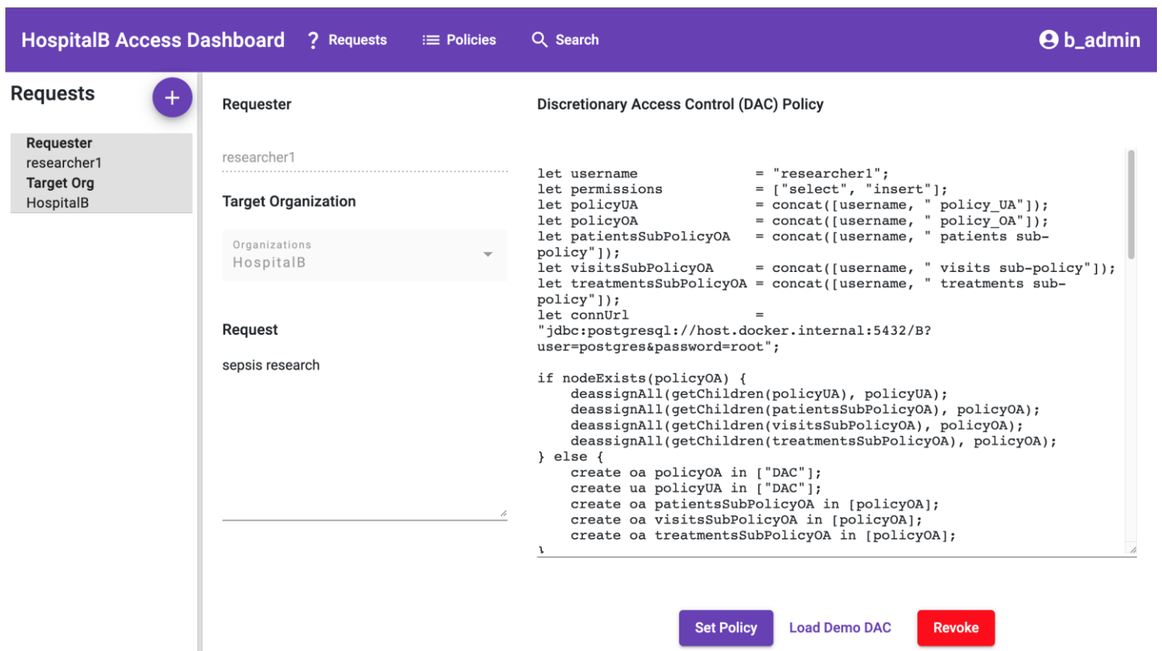


1. **Researcher1** requests access to data at **HospitalA** and **HospitalB** (cont)

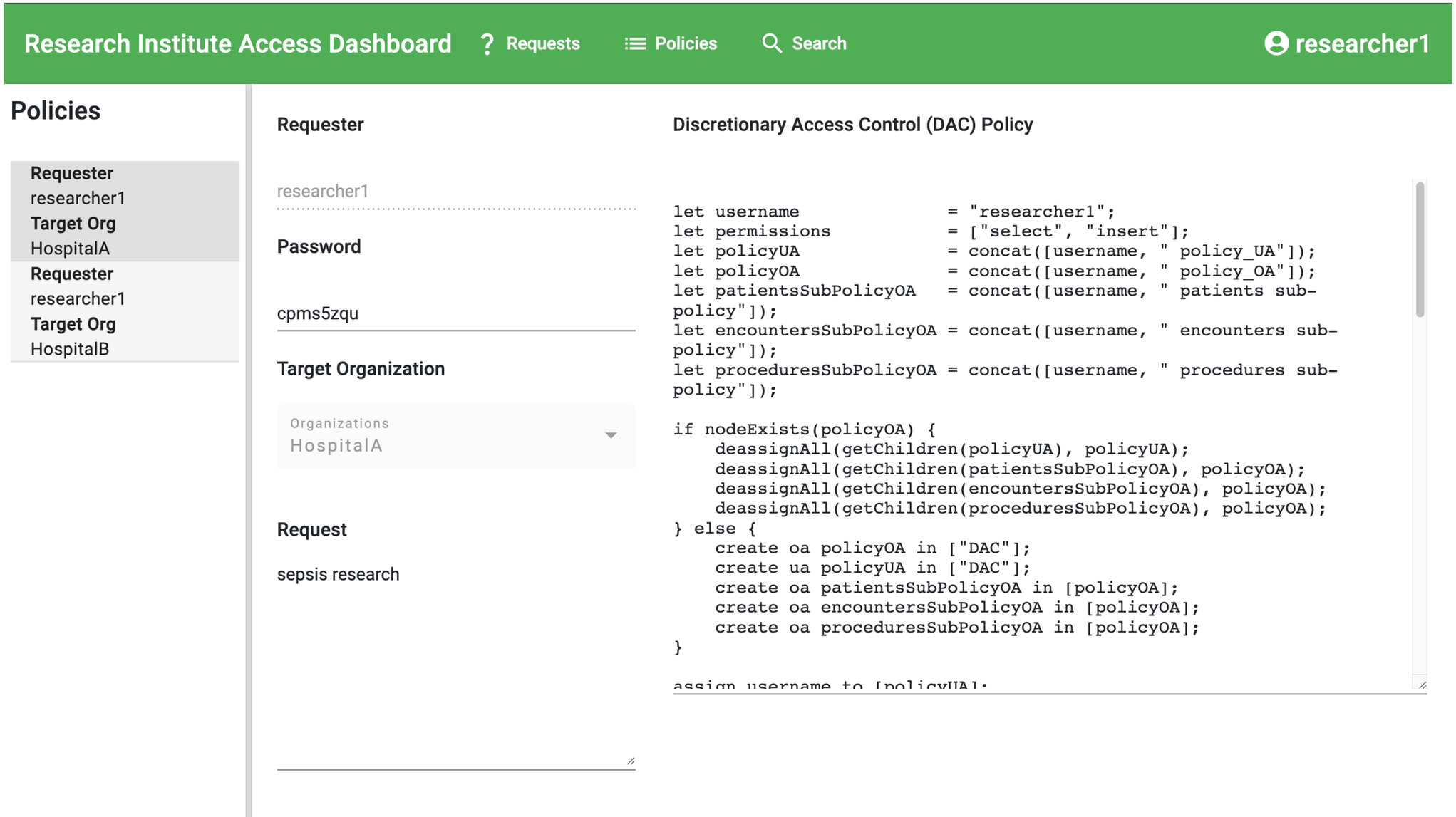


2. **HospitalA** and **HospitalB admins** set **DAC** policies for **Researcher1**

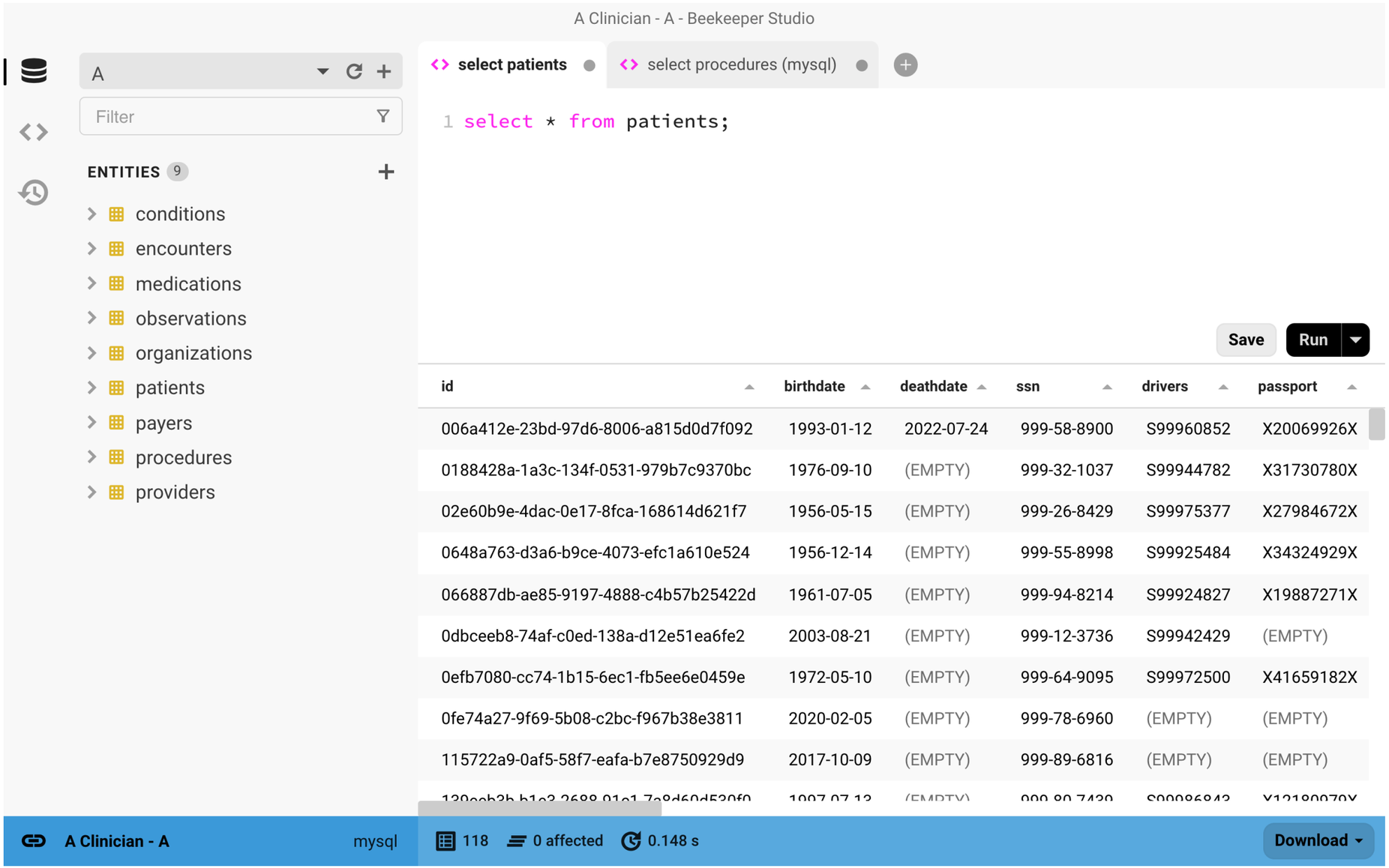




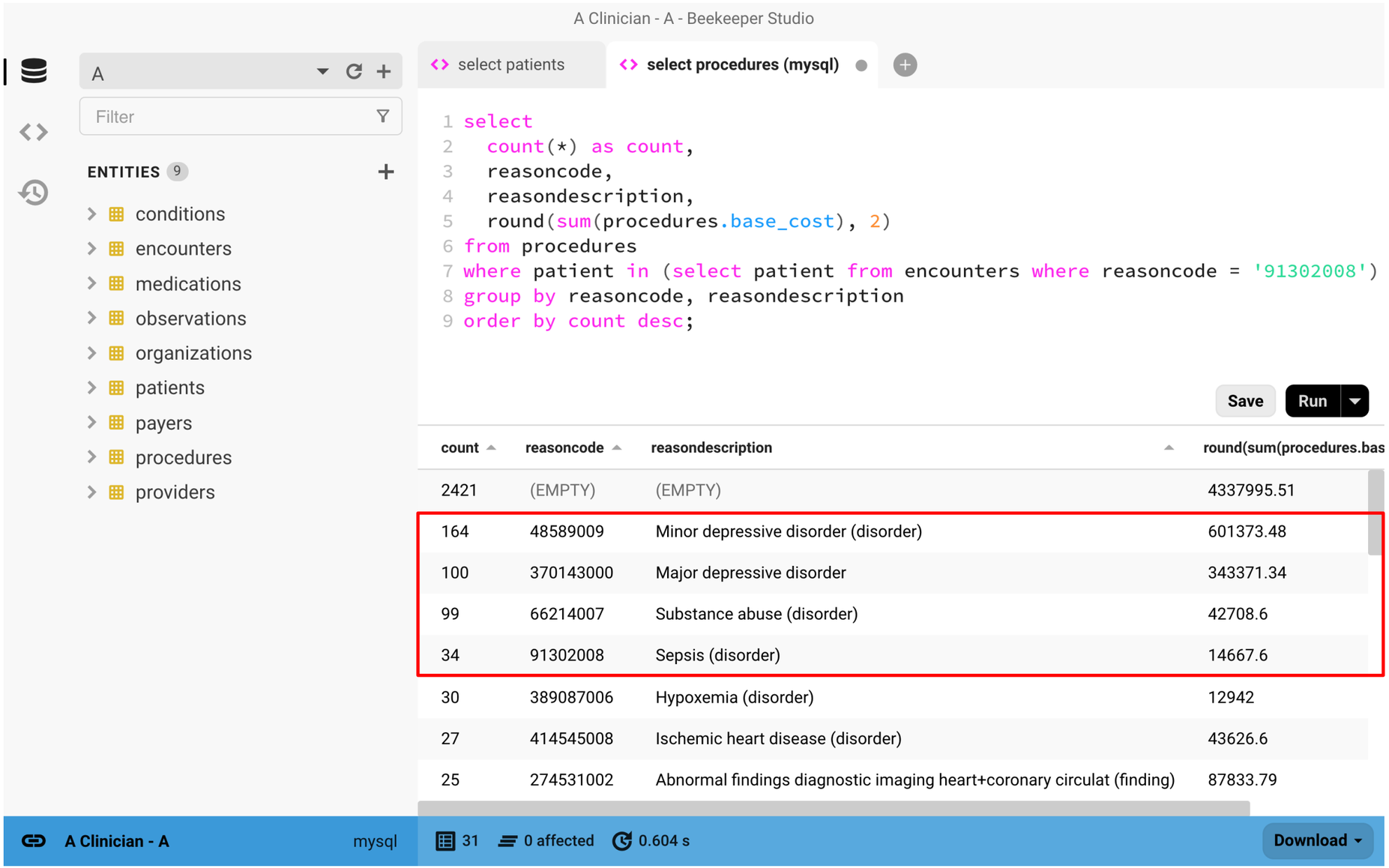
3. **Researcher1** reviews **DAC** policies and obtains authentication **tokens**



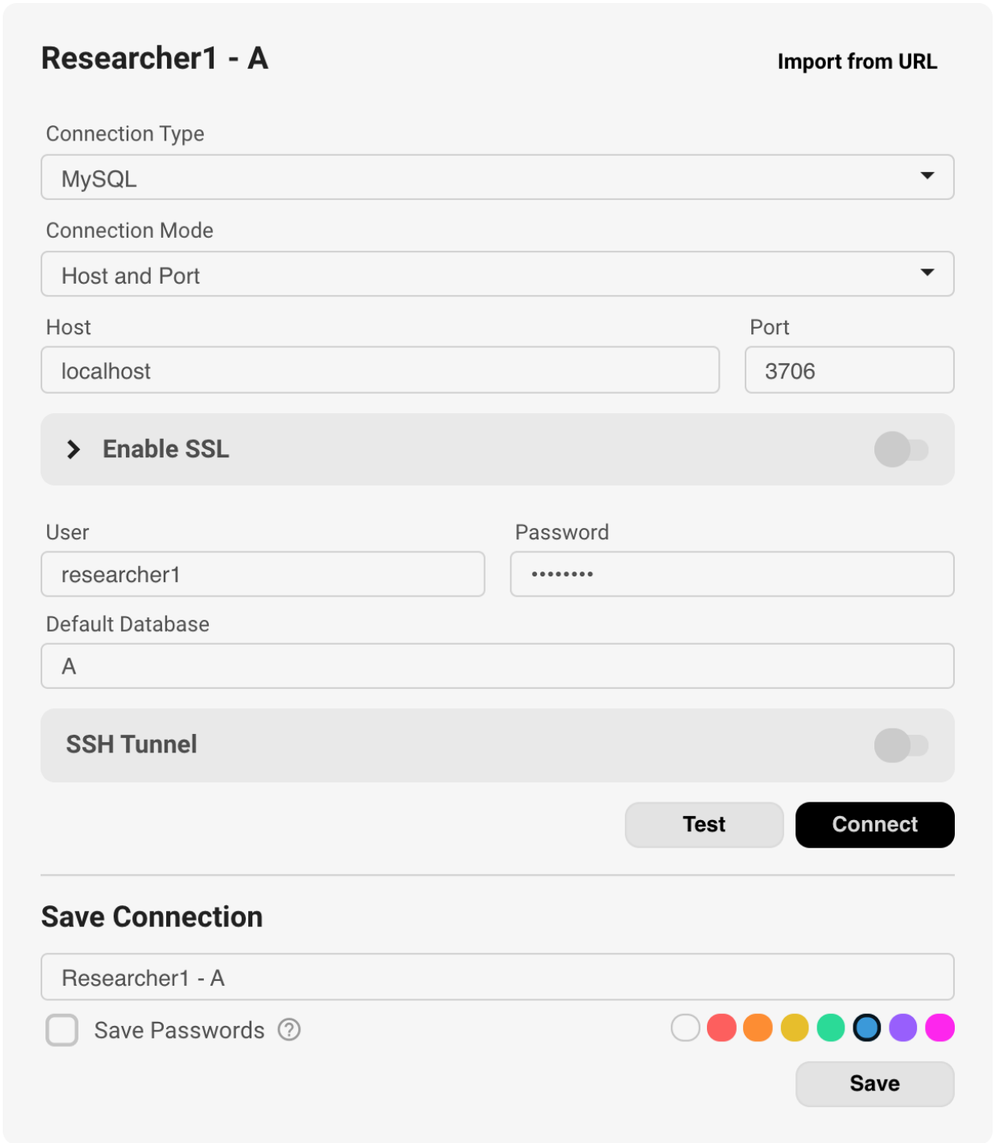
4. **A_clinician** selects from **patients** table at **HospitalA**



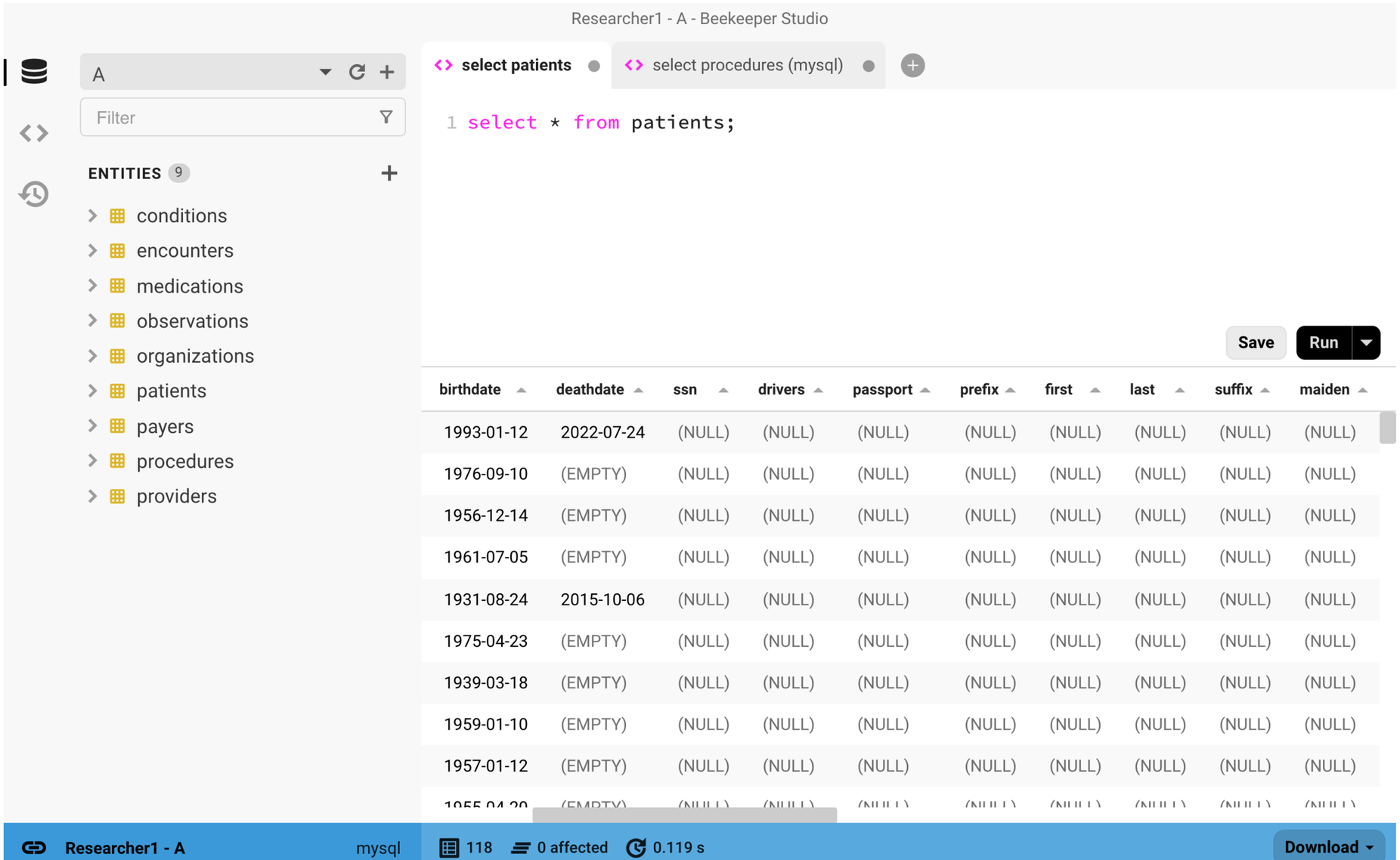
5. **A_clinician** selects from **procedures** table at **HospitalA**



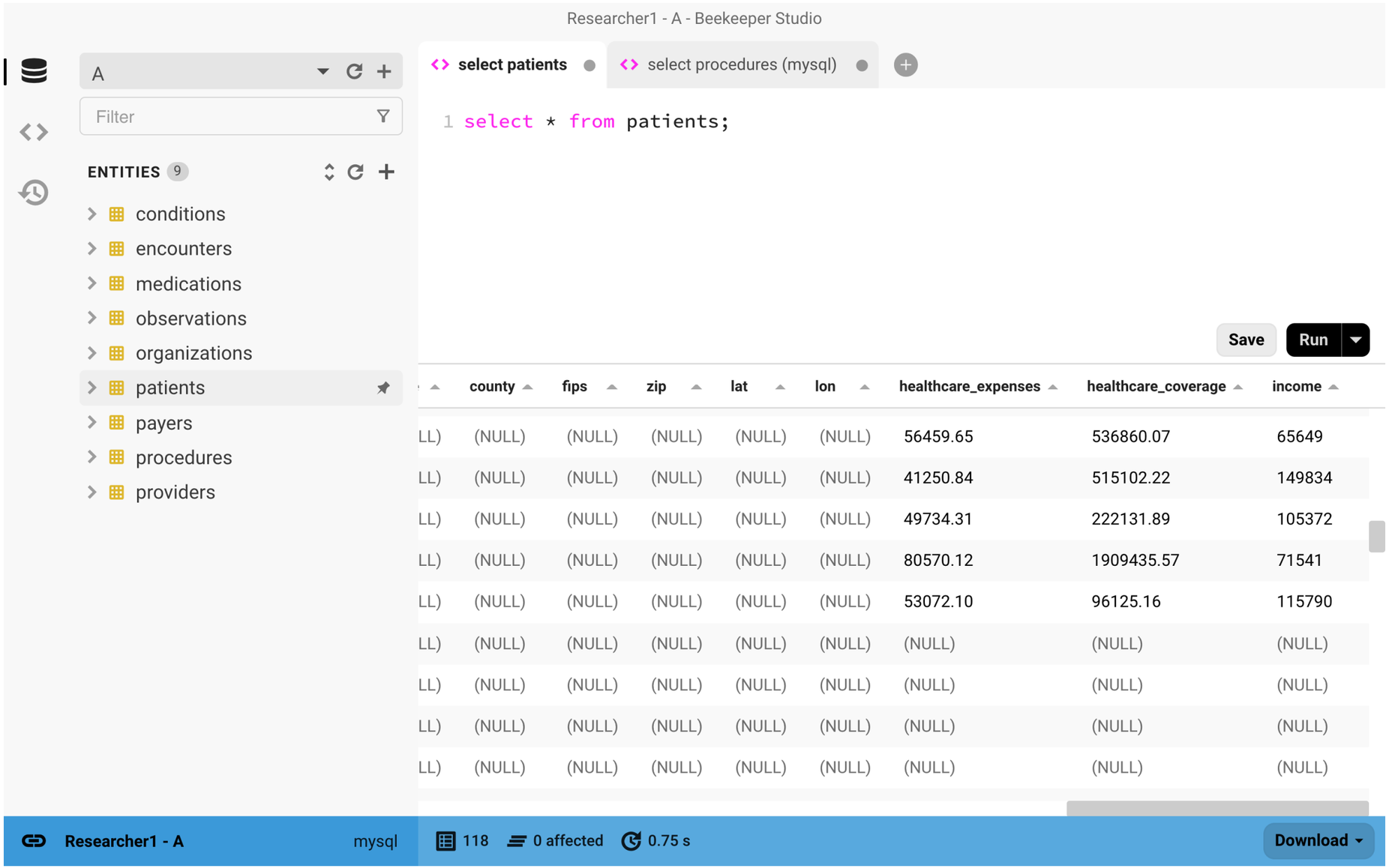
6. **Researcher1** accesses **HospitalA's** database



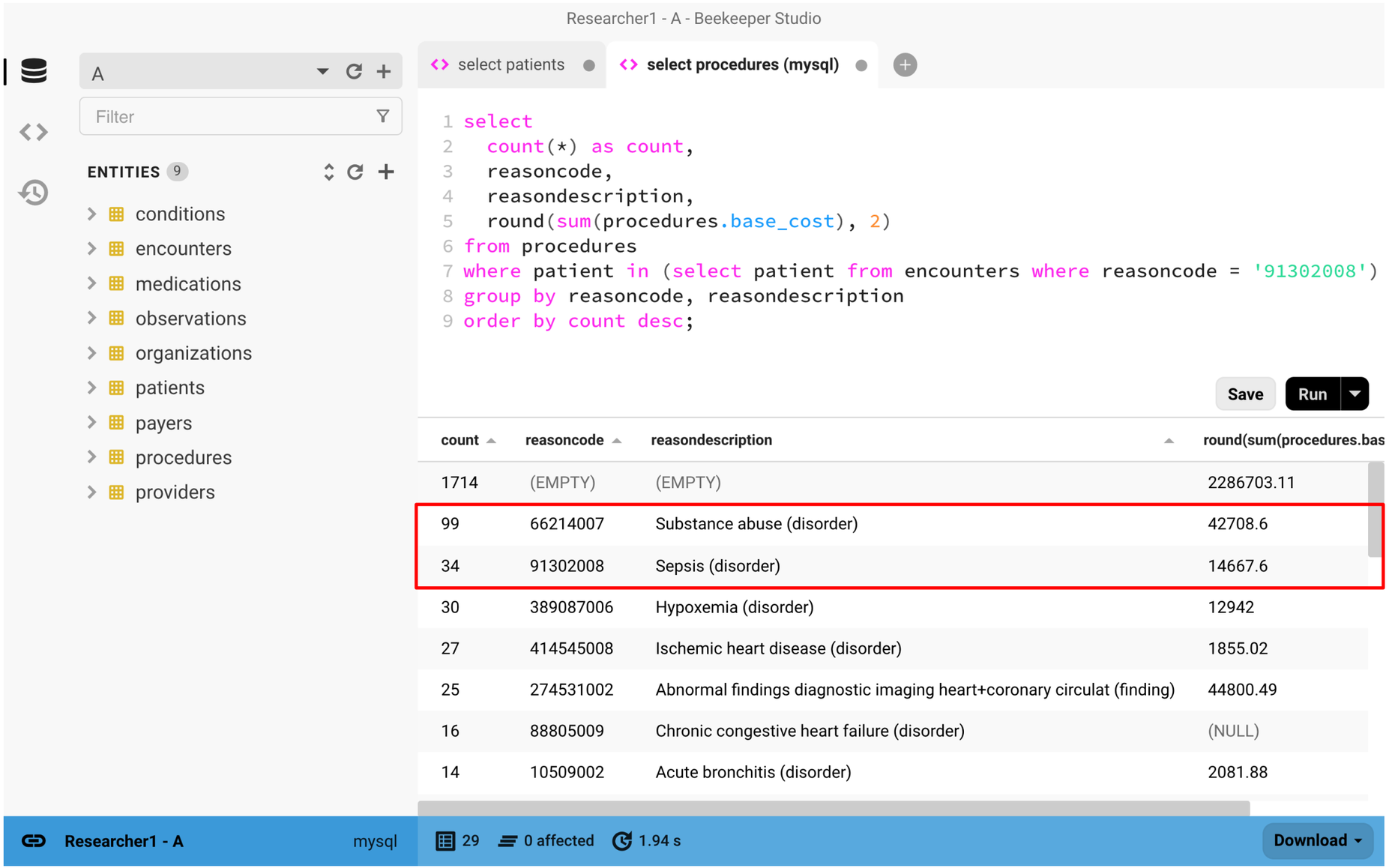
7. **Researcher1** selects from **patients** table at **HospitalA**



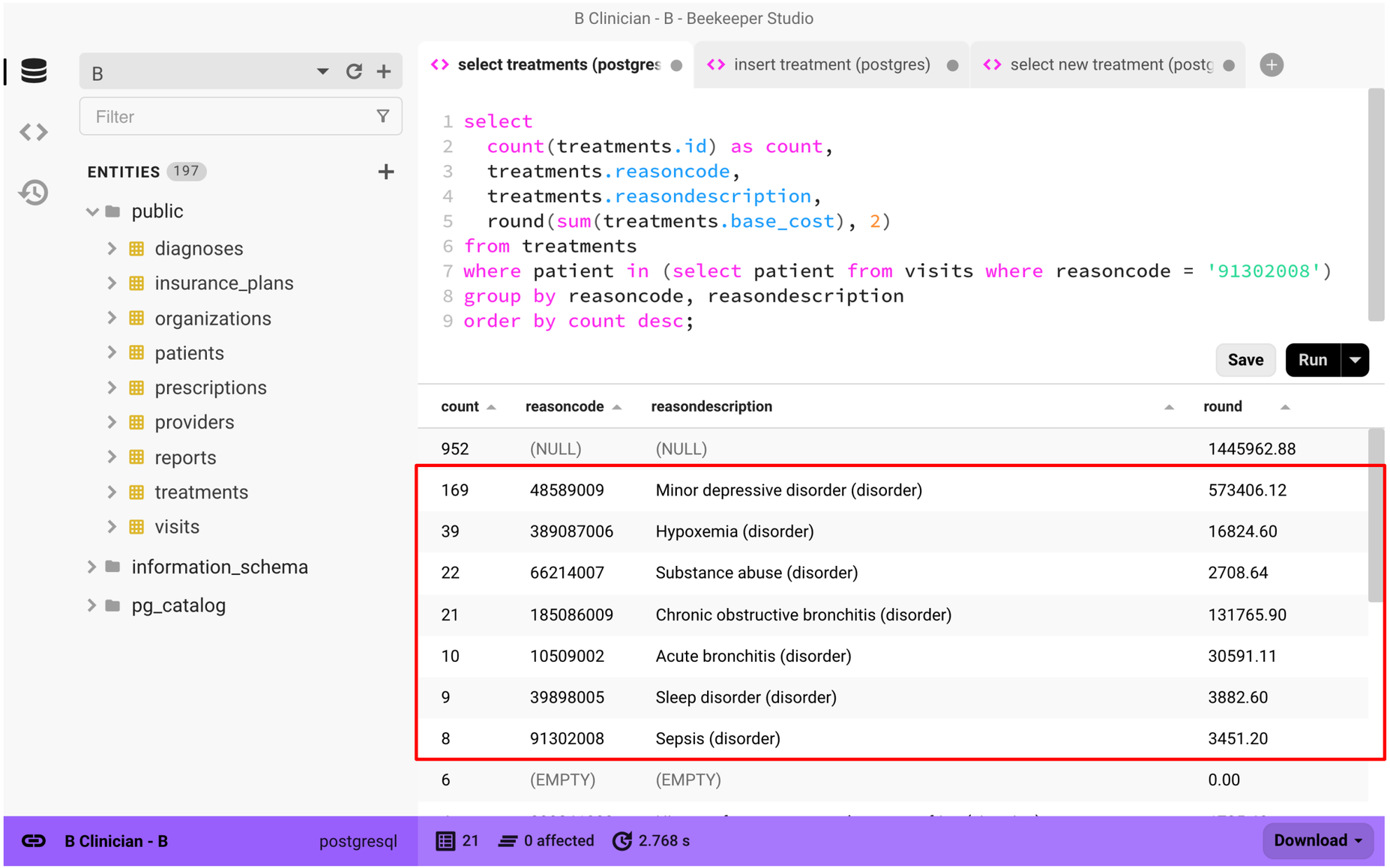
7. **Researcher1** selects from **patients** table at **HospitalA** (cont)



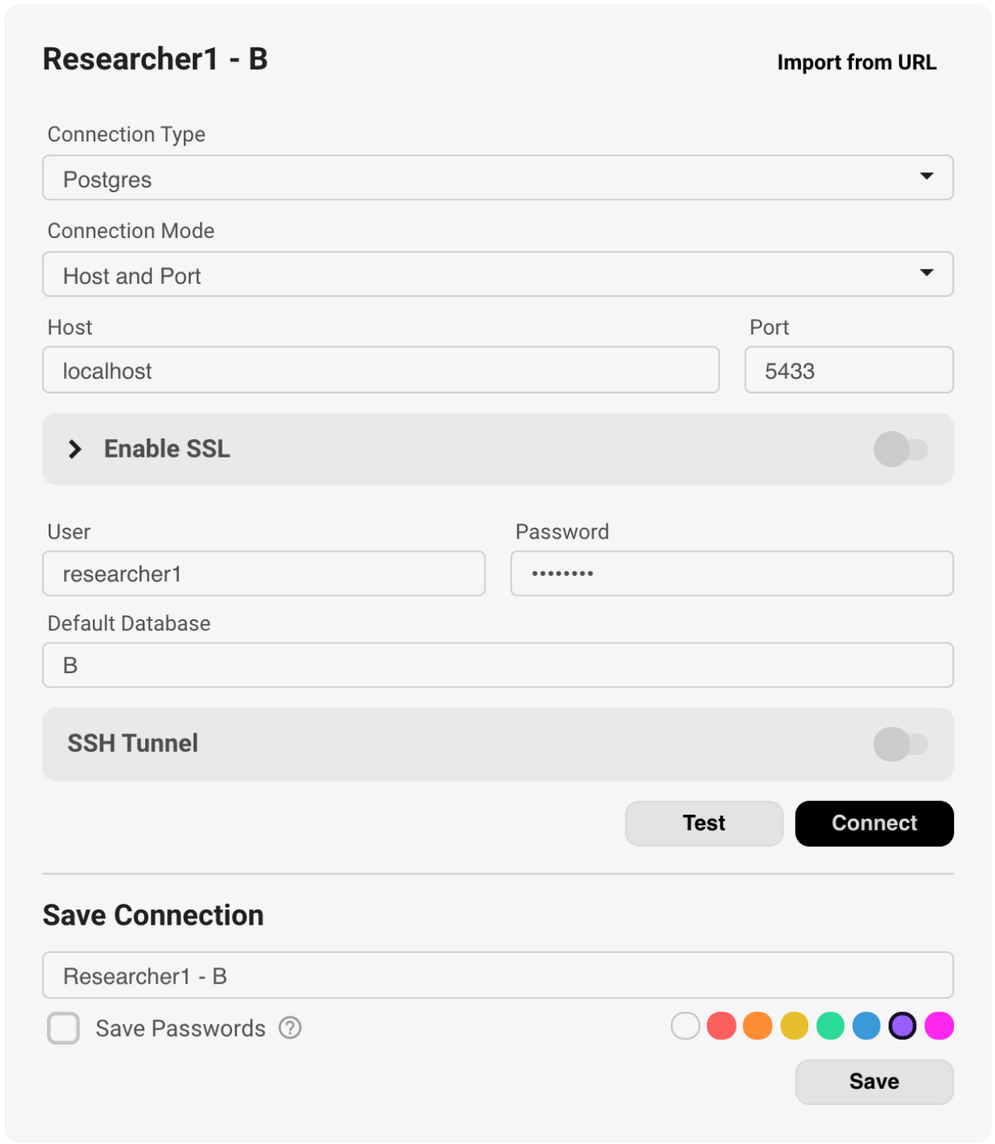
8. **Researcher1** selects from **procedures** table at **HospitalA**



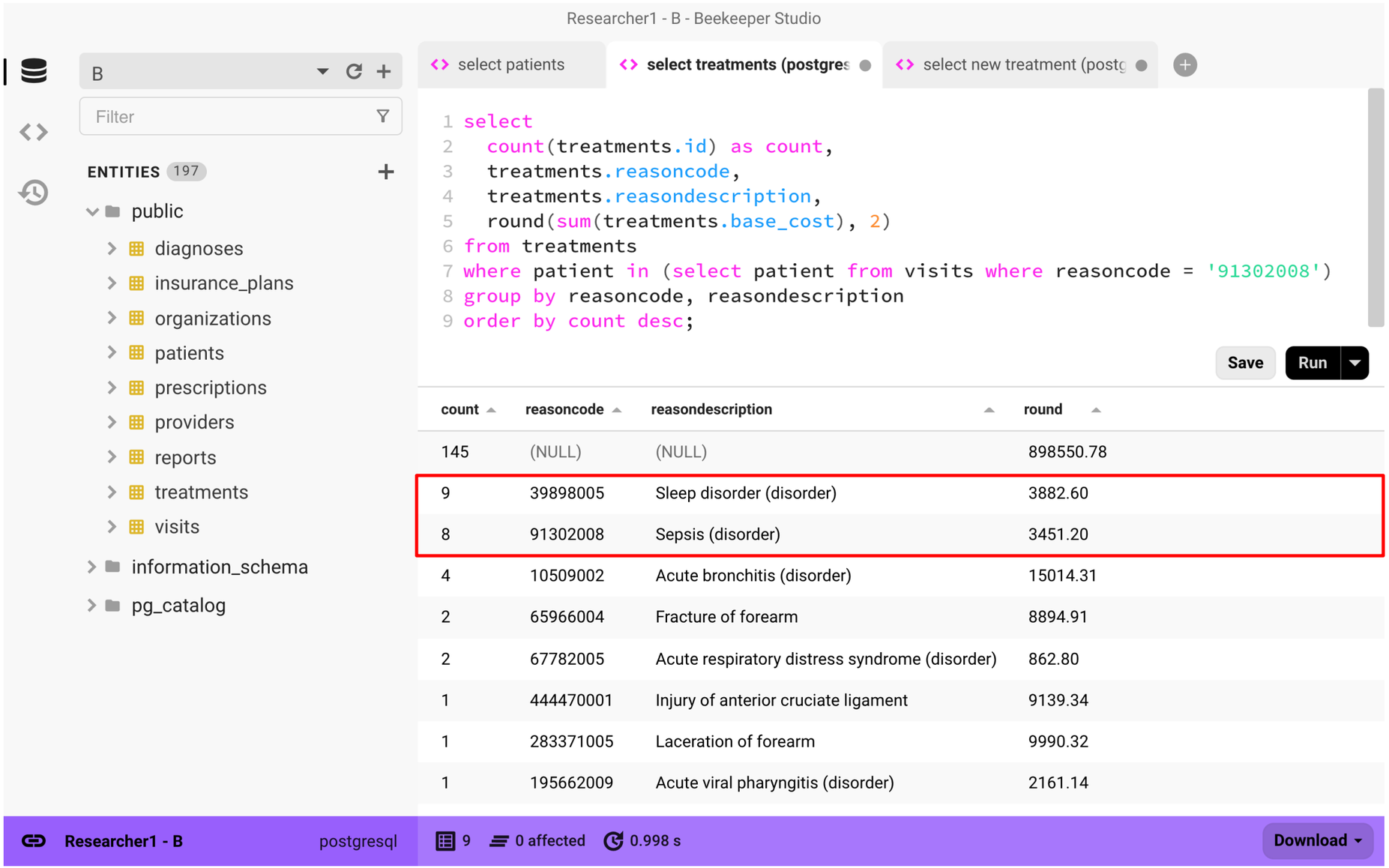
9. **B_clinician** selects from **treatments** table at **HospitalB**



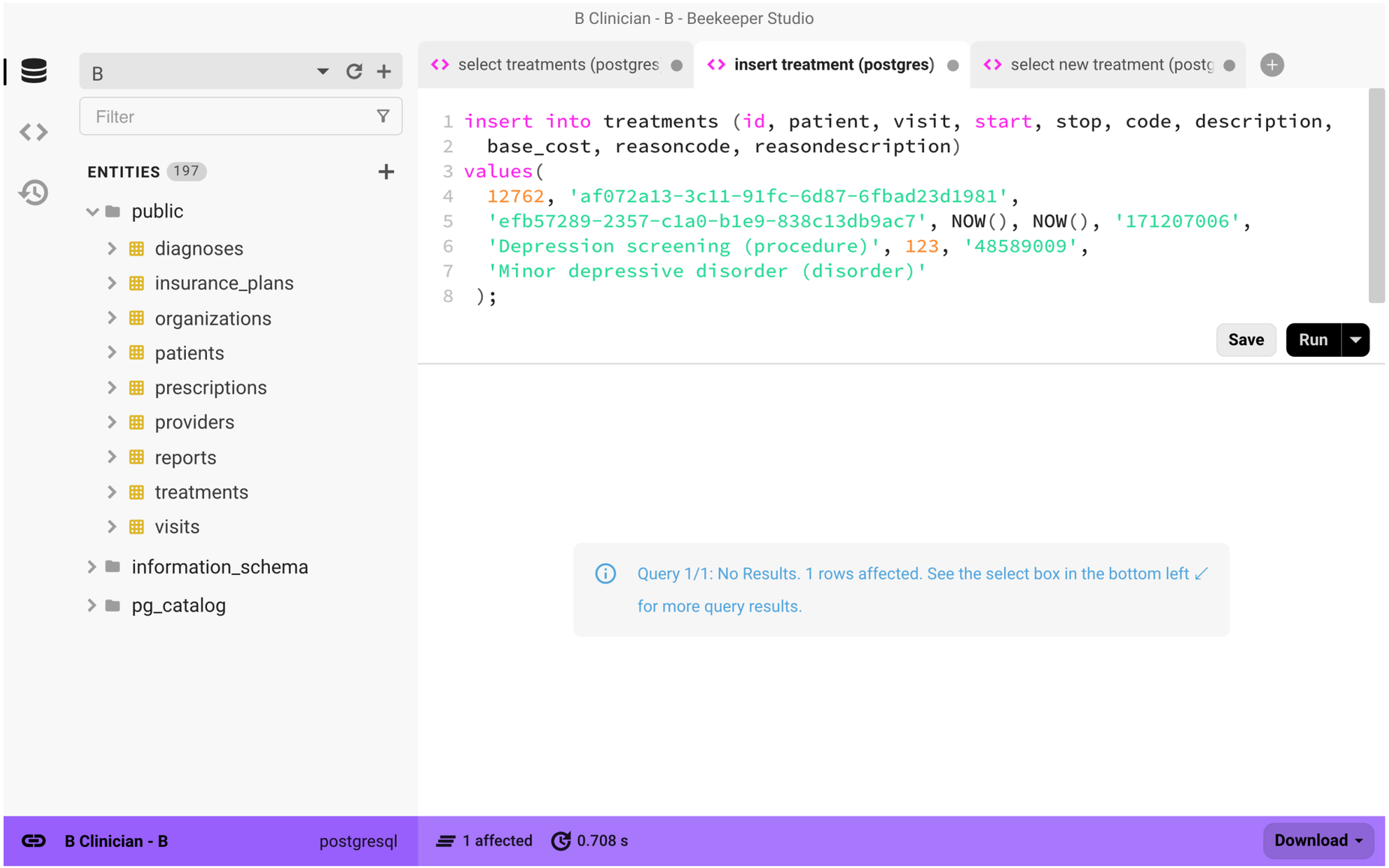
10. **Researcher1** accesses **HospitalB's** database



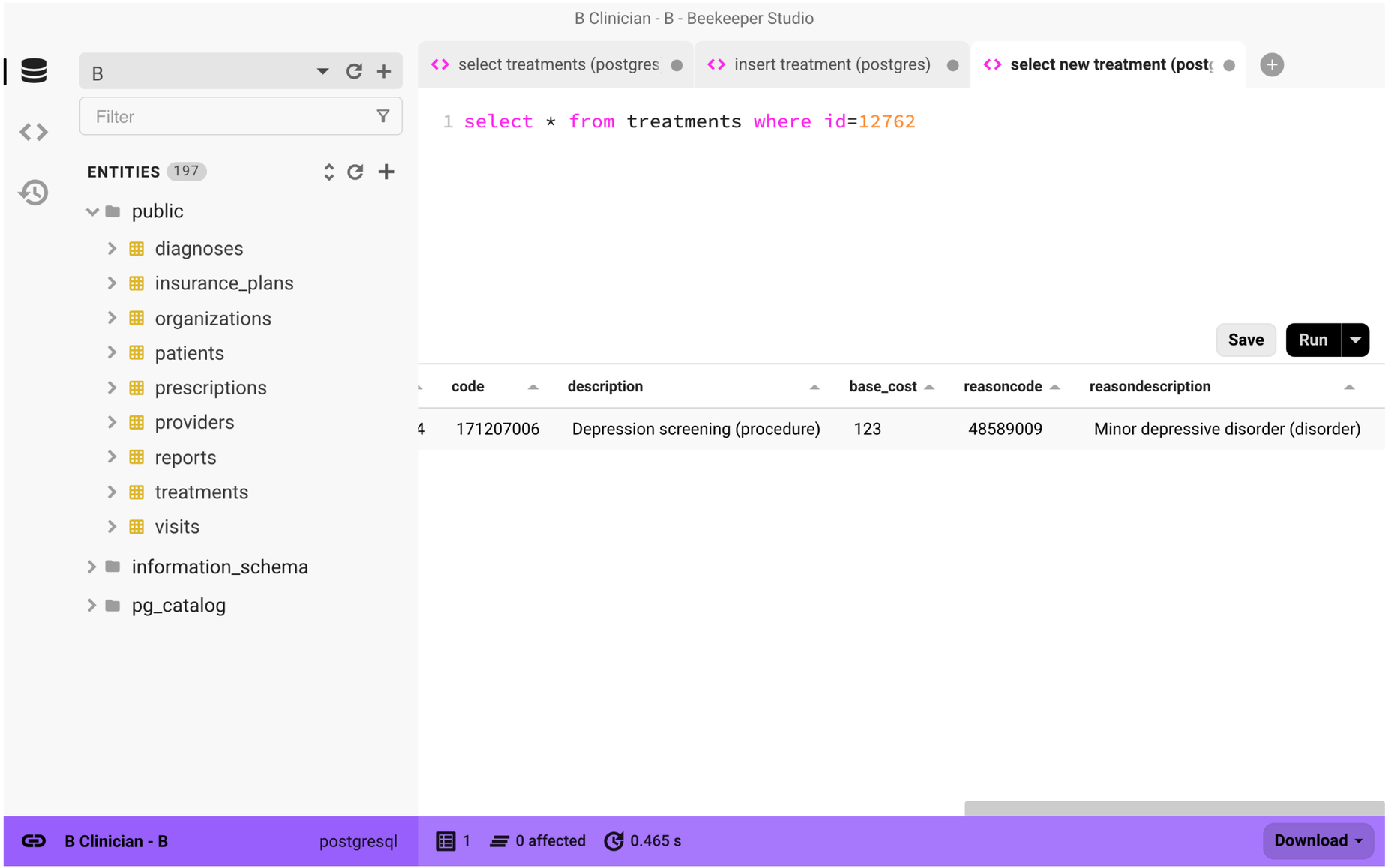
11. **Researcher1** selects from **treatments** table at **HospitalB**



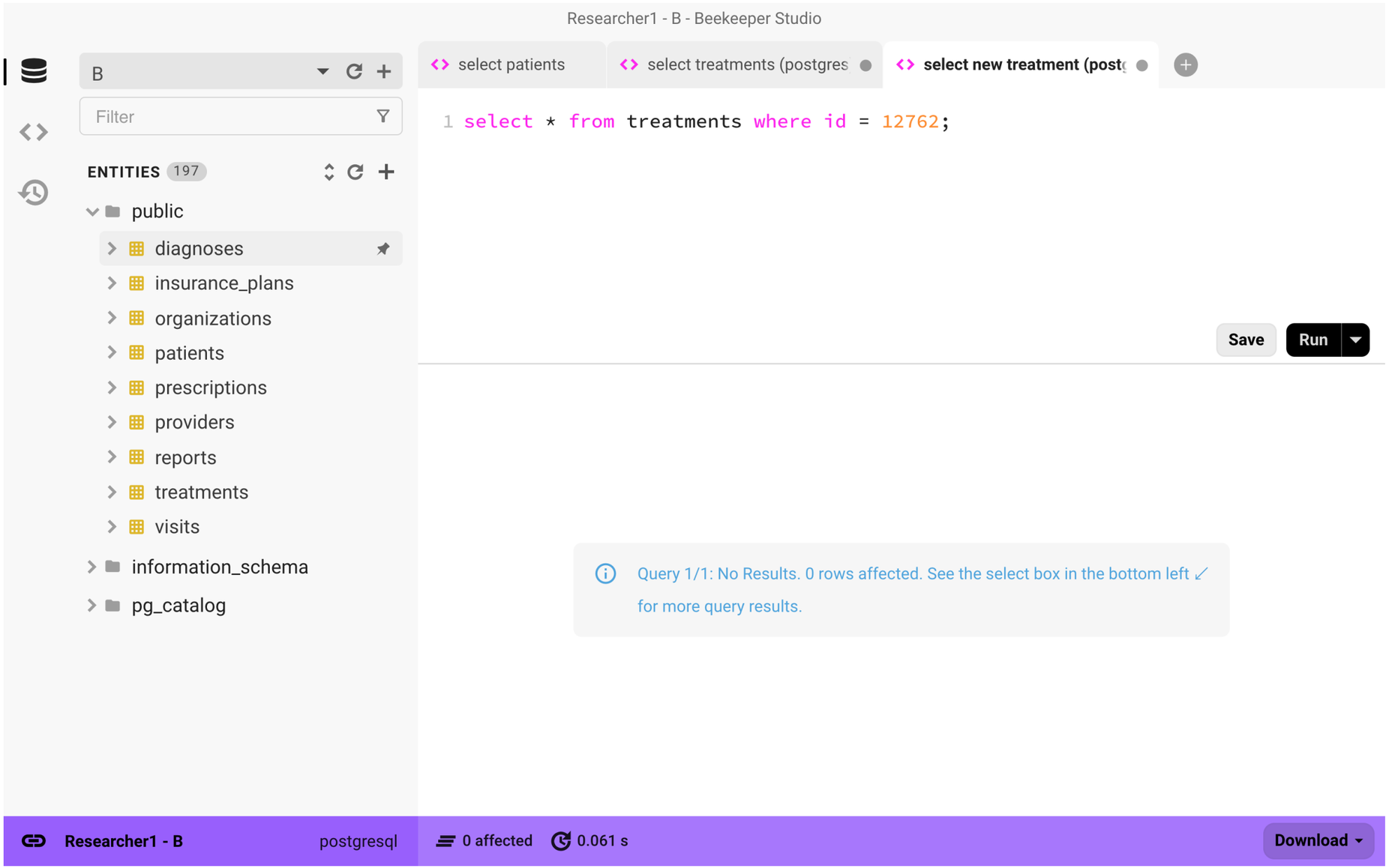
12. **B_clinician** inserts new **mental health treatment** row



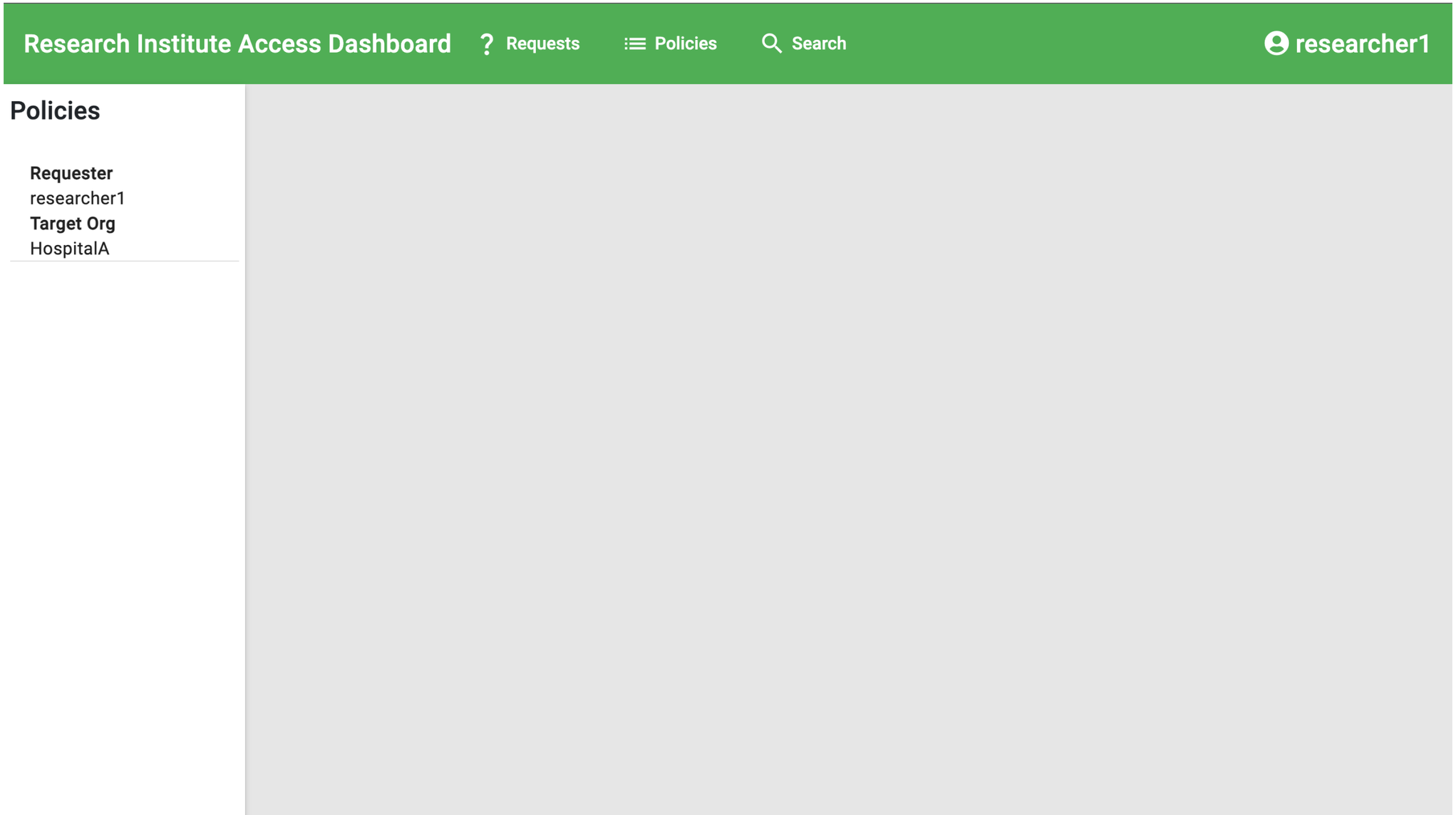
12. **B_clinician** inserts new **mental health treatment** record (cont)



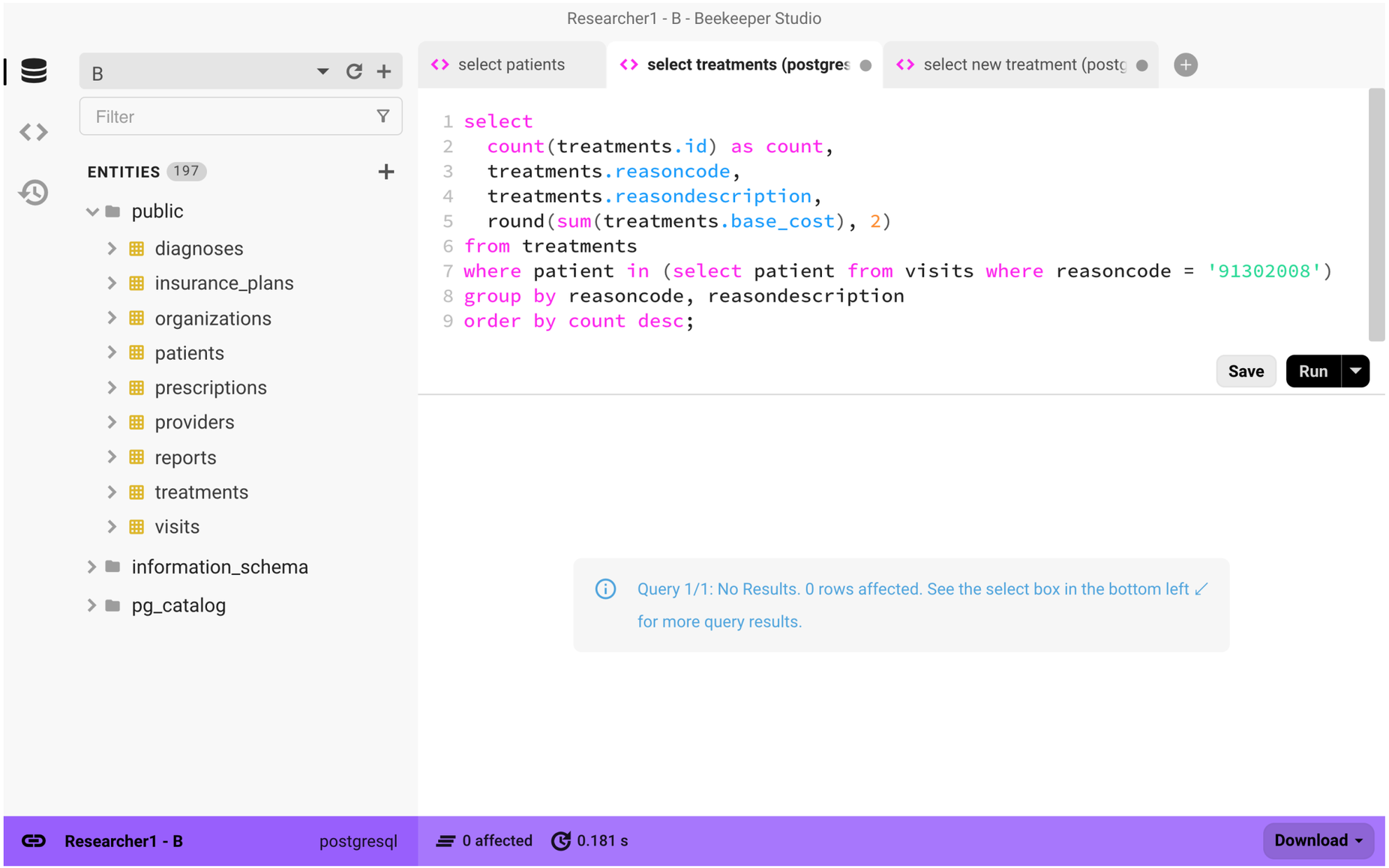
13. **Researcher1** selects new row



14. **HospitalB admin revokes** Researcher1’s access



15. **Researcher1** selects from **treatments** table at **HospitalB**



Conclusion and Summary

- The ability to share database resources among collaborating organizations is highly desirable.
- However, challenges persist regarding interoperability in the exchange of resources between organizations and the preservation of local access policies.
- SFDS provides a generic data sharing infrastructure that effectively and securely achieves data sharing objectives.
- Transparent to the otherwise normal business operations of participating organizations.
- This ease of deployment, granularity of control and efficiency make this new infrastructure solution practical for meeting the data sharing and protection objectives of the clinical research community.

Future Work

- The current focus of SFDS is to allow sharing of database resources. Included in our plans is to extend the infrastructure to allow controlled access to non-structured data such as files.

- We would like to organize a full-scale pilot study involving institutions that house medical information and/or conduct clinical research.

- Our goal is to transition SFDS from a research project to operational use.

Thank You

Joshua Roberts and David Ferraiolo

Acknowledgement: D. Chris Compton and Joanna DeFranco

Analytics with Altitude

2023 HDAA Conference

Denver, Colorado

Children's Hospital Colorado

HDAA
HEALTHCARE DATA & ANALYTICS ASSOCIATION