



Partners equalize rights & raise efficiency with local data systems

Why a team in India always keeps data in country



data sovereignty | de-identification

Duke | GLOBAL HEALTH
INSTITUTE



BACKGROUND

Manoj Mohanan is an applied microeconomist, focusing on health and development economics, with a background in medicine and public health. He's been collaborating with partners in India to study health care and health outcomes for over a decade. Current projects involve primary quality of care data and secondary mental health data from the largest household panel being conducted anywhere in the world. Almost every study Manoj and his partners conduct involves individually-identifiable health information (IIHI).



AT A GLANCE

equalize
data rights



simplify
oversight

streamline
data sharing



Researcher

Manoj Mohanan

Field

Health and development
economics

Country

India

Key approaches

Data systems in-country
De-identification



THE CHALLENGE

Since Manoj began this work in 2009, hosting global health data at Duke has become increasingly challenging due to an ever-evolving landscape of research ethics, security, law, and policy. Challenges are most strongly felt by those outside the institution. External research partners rely on colleagues for access to Duke systems. Limitations may be placed on how they engage with sensitive data like IIHI – for example, prohibitions against moving data off Duke-owned resources. This produces inequity around how much and how easily different members of a research team can access and use data. It can be especially frustrating when there is a disconnect between how data is handled in the US versus the local study context. For example, data considered sensitive under HIPAA may not be regulated the same way in other countries.

“They would like to see a system in the future where all data is collected, owned, and stored in India.”

Differing regulations are one reason India’s government and local IRBs increasingly favor a data sovereignty approach. Data sovereignty refers to the practice of keeping data within the borders of the country where it was generated. “The government wants to use ... principles which basically say that the data cannot leave India,” said Manoj. “They would like to see a system in the future where all data is collected, owned, and stored in India.” However, the academic institutions Manoj partners with aren’t usually equipped to manage research data with IIHI. “They are not as flexible with their infrastructures or IT needs if we need [data] to be maintained in a specific manner,” said Manoj. And at Duke and overseas institutions alike, there’s risk that changes like a co-PI switching jobs will disrupt data access in ways difficult to resolve from afar.



SOLUTION



Manoj's solution was to engage an in-country research organization to collect, manage, and store study IIHI. "We work with academic partners in India for the research," said Manoj. "Together we write manuscripts and look at the analysis. The data is collected by a third party [research organization], de-identified, and shared directly with the Duke team, and then we manage all the data analysis with colleagues as we would for any de-identified data project."

This was an attractive solution for multiple reasons. Local research organizations either already have the infrastructure and expertise to manage sensitive data like IIHI, or the incentive to invest in it. They store identifiable data needed for auditing and oversight in-country, providing compliance with the local IRBs and government. Finally, they eliminate the problem of data access disruptions occurring as a result of changes at either partner institution. Academic partners based in India can go directly to the research organization for their data access and use needs. "We decided that this strategy was the best because it guarantees that over the life of the project... the research organization holds on to identifying information in a systematic manner that will not compromise data security," said Manoj. He noted this arrangement meant slightly higher costs around data, but added "I've come to realize in the world of data collection, you get what you pay for."

"In all of our projects, we have no identifiable information that comes to Duke at any time."



RESULTS

Manoj and his partners in India have implemented this data strategy on all their projects. They find it promotes equitable data rights among all the researchers involved in the partnership. “We have the same level of authority, same level of access to data, and influence on the project because [of] the underlying structure. The research team [Duke and non-Duke academic partners] develops the project, designs the data collection, and has all the technical input” said Manoj. “The kind of issues you raised around equity and access to data and control, I think in many ways are worse when you have a subcontracting relationship where I as the lead PI am subcontracting with another academic institution and I’m telling them what to do. It creates this hierarchy.” He cited India’s stringent reporting requirements around foreign money received by not-for-profits as another reason his academic partners like using local research organizations for data needs. “It reduces the amount of headaches for them.

Keeping IIHI in-country instead of at Duke reduces the time spent on study oversight

and data management. “It was easier for me to write IRB applications and get permissions and approvals,” said Manoj. Contracts like data use and data transfer agreements (DUAs, DTAs) are either simplified or unnecessary with de-identified data, and it’s much easier to share analytic data files with external partners via Duke

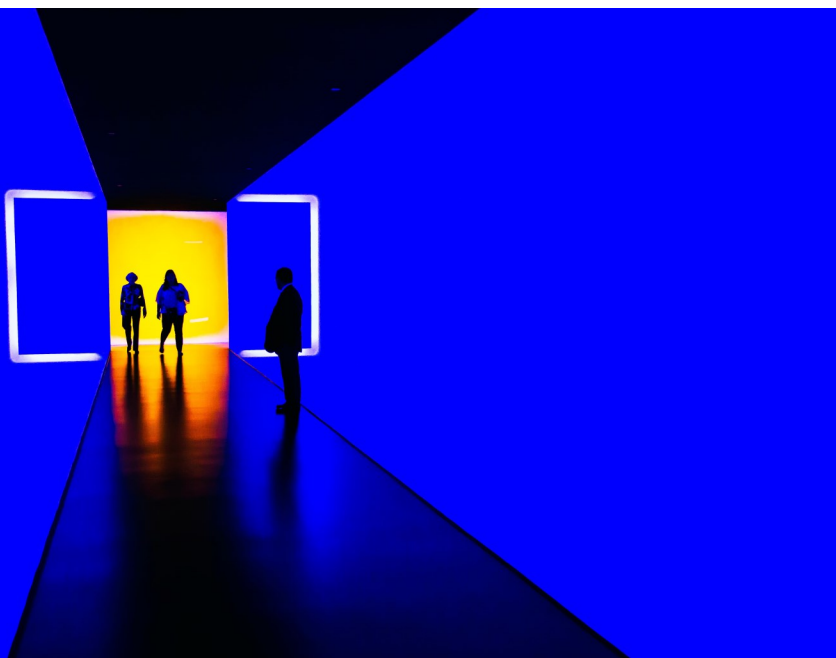
“We have the same level of authority, same level of access to data, and influence on the project because [of] the underlying structure.”

systems like Box since no PHI is leaving the institution. Another positive benefit is that de-identified data is ready to be shared outside the study. “We tend to make all data publicly available, usually in appendices to papers, online with the journals, or [in repositories like] Dataverse,” Manoj said.



WHAT'S NEXT

Manoj and his partners aren't planning on changing their strategy anytime soon. Keeping identifiable data in-country and de-identifying data for analysis early on have helped their collaborations run smoothly. "I don't see why that cannot become the standard default," said Manoj, "rather than getting into this business of Duke owning the data and restricting who gets to use data."



TRY IT OUT



Limit identifier collection

Omit direct and indirect identifiers from data whenever possible. Separate study administrative records to keep identifiers like contact information out of data. Avoid free-text fields, which often contain indirect identifiers, when you can.

Keep data in-country

Local research organizations aren't the only option. Many academic institutions have research data infrastructure. Assess how partner institutions could support local data management and storage.

Follow local developments around data

India isn't the only country prioritizing data sovereignty. Work with local partners to plan ahead for navigating anticipated changes in data requirements.