| PROJECT INFORMATION | May 2020 |

Developing Methods to Link Patient Records across Data Sets that Preserve Patient Privacy

| **Principal investigator** | **Organization** |
| Kevin Haynes, PharmD, MSCE | HealthCore |

What was the project about?

Data from healthcare systems, patients and communities, and health plans can support health research. Two types of data sources are

- **Patient-powered research networks, or PPRNs.** In PPRNs, patients, families, caregivers, and community members share health data with the network. They work closely with researchers to plan and conduct research.

- **Health plan research networks, or HPRNs.** In HPRNs, networks of health plans have access to health claims data from members for research.

By linking patient records across PPRNs and HPRNs, researchers may be able to do more robust research. To link records, researchers use computer programs to connect the records of people in a PPRN with their claims data in an HPRN. Current methods to link records require use of personal information, such as names and dates of birth. But patients may not want to share this information.

In this project, the research team developed methods for linking data from PPRNs and HPRNs without using patients' personal information.

What did the research team do?

The research team developed new methods that protected patients' privacy. The team linked data from four PPRNs and an HPRN. The team confirmed the diagnosis of seven health problems in PPRN data with claims data in the HPRN. The team then compared patients who had linked data with patients without linked data.

Next the research team talked to nine patients from seven PPRNs. The team asked patients their views about taking part in research being done by HPRNs.

What were the results?

The methods linked data from 4,487 of the 21,616 PPRN patients with claims data in the HPRN.

For 50 to 75 percent of patients with linked records, claims data confirmed diagnoses that patients shared in PPRNs. The longer patients were enrolled in their health plan, the more likely claims data could confirm diagnoses in linked data.

Compared with patients without linked data, patients with linked data were younger, more likely to be women, and less likely to have several health problems.

In interviews, patients noted barriers to taking part in HPRN research:

- Changes in health plan enrollment

- Trust in health plans

- Fear of data breaches or reduced insurance benefits

What were the limits of the project?

Errors in claims data could affect this study's findings. The nine interviews may not represent all views on taking part in HPRN research.

Future research could compare the methods in this study versus methods that use personal information with proper permissions from patients.

How can people use the results?

To maintain privacy, researchers can use the methods to link data between PPRNs and HPRNs without using patients' personal information.

To learn more about this project, visit www.pcori.org/Haynes320.