What would it take for health systems to better care for high-need, high-cost (HNHC) patients? Countries across the world are contending with growth in health-care expenditures, and a disproportionate share of spending is concentrated on small groups of patients with high healthcare utilization. In the United States, 1% of the population accounts for 21% of healthcare expenditures and the costliest 5% account for more than half of all health spending.

HNHC patients have traditionally been identified simply as those with high healthcare utilization. But recent work, including a report from the National Academy of Medicine, aims to stratify patients into subgroups, recognizing their varied circumstances and needs. These needs often extend beyond those caused by their medical conditions and include behavioral, functional, and social challenges.

“Precision medicine,” customizing diagnostics and treatments based on a patient’s genetic profile, is generally thought of as essential for the future of personalized care. But for HNHC patients today, “precision health” may be just as important: developing care models based on patients’ unique medical and sociocultural needs and matching those models with the patients most likely to benefit. Creating meaningful taxonomies and translating them into clinical and financial gains will require comprehensive databases with information on medical claims, clinical factors (including genomics), and social determinants of health.

Some health systems have begun collecting and analyzing data on HNHC patients, but current efforts are limited in important ways. First, they do not always include data from multiple hospitals and health systems. Healthcare utilization is often distributed across several systems, particularly in urban settings, and many HNHC patients are not identified by individual hospital analyses in the absence of robust data-sharing arrangements. For example, a study of HNHC patients in Maryland found that only 30% visited 1 hospital, whereas 38% visited 3 or more hospitals. Yet hospitals in more competitive markets are less likely to engage in health information exchange.

Second, many analyses are limited by payer type, often with a focus on Medicare claims. But differently insured patients have...
different drivers of healthcare utilization. Healthcare spending on Medicare patients is driven largely by end-organ sequelae of multiple chronic conditions, whereas mental health disorders predominate in Medicaid patients. Catastrophic injuries and specialty drugs are important drivers of high costs in commercially insured populations.

Third, most analyses rely primarily on claims data. Using claims data alone may underestimate the prevalence and severity of chronic disease and mental illness. It also fails to capture important clinical factors, such as functional status, or genomic information, which is increasingly recognized as contributing to variability in drug response, dosing, and adverse effects and may allow for safer and more targeted medication choices. Clopidogrel (Plavix), for example, requires biotransformation in the liver, and CYP2C19 homozygosity is associated with reduced response. Finally, claims-based analyses do not incorporate social determinants of health, which are widely understood to be important, sometimes modifiable, contributors to healthcare costs.

Addressing these concerns will require large structured databases that include inputs from multiple health systems and incorporate claims, clinical, and social information. Such databases would allow better prediction of the low- or medium-utilizer patients who will progress to become high utilizers, which high utilizers will remain high utilizers, which care delivery models will be effective for particular subgroups, and how to allocate limited health system resources in targeted ways to improve outcomes and curb costs.

One such effort is under way at the New York City Clinical Data Research Network (NYC-CDRN). Funded by the Patient-Centered Outcomes Research Institute (PCORI), the NYC-CDRN brings together 22 institutions—health systems, medical colleges, research groups, and government agencies—to share data, develop common tools, and collaborate on research. The network includes comprehensive longitudinal data on millions of patients and uses open-source code to develop high-utilizer “phenotypes” and focus interventions. Similar data-sharing initiatives are being developed in 12 other regions around the country, although they do not currently focus on HNHC patients.

These data-sharing networks allow for both a more complete identification of HNHC patients and a deeper understanding of how clinical and social factors interact at the population level to produce unique challenges for particular subgroups. For example, how does lack of transportation or unstable housing differentially influence care, costs, and outcomes for older, frailer patients compared with disabled patients or mentally ill patients? Transportation may be critical for HNHC patients from particular zip codes or with functional limitations but less relevant for a young patient with sickle cell disease. Integrated behavioral health services may be the most important intervention for patients with coexisting mental illness, but they are unlikely to affect utilization for patients with advanced illness and robust social support.

Consider a HNHC patient with multiple admissions for heart failure exacerbations (Table). How might merging claims, clinical, and social data in a multisystem registry help detect the patient’s needs and customize care? Integrated claims data would allow us to correctly identify the patient, especially if they are seeking care at multiple hospitals, and provide an estimate of disease severity based on International Classification of Diseases, Tenth Revision codes; number of comorbidities; and frequency of admissions. Clinical data would enhance our ability to risk stratify the patient using, for example, brain natriuretic peptide levels, recent weight fluctuations, and frequency of diuretic dose changes. Genomic information may ultimately inform warfarin or β-blocker dosing. Finally, social data would provide additional avenues for intervention: Does the patient...

---

### TABLE. Using Integrated Claims, Clinical, and Social Data to Stratify High-Need, High-Cost Patients

<table>
<thead>
<tr>
<th>Patient/Disease</th>
<th>Claims</th>
<th>Clinical/Genomics</th>
<th>Social</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heart failure</td>
<td>• Frequency of admissions&lt;br&gt;• Diastolic vs systolic&lt;br&gt;• Number of cardiopulmonary comorbidities</td>
<td>• Weight fluctuations&lt;br&gt;• Frequency of diuretic dose changes&lt;br&gt;• Brain natriuretic peptide&lt;br&gt;• Genetic polymorphisms informing β-blocker or warfarin dosing</td>
<td>• Availability of air conditioner&lt;br&gt;• Residence in food desert&lt;br&gt;• Regional air pollution severity</td>
</tr>
<tr>
<td>Lung cancer</td>
<td>• Staging&lt;br&gt;• Procedures (eg, frequency of thoracentesis)&lt;br&gt;• Supplemental oxygen requirement&lt;br&gt;• Opioid use</td>
<td>• Ambulatory desaturation&lt;br&gt;• Pain scores&lt;br&gt;• Palliative care consultation&lt;br&gt;• Albumin level and degree of cachexia&lt;br&gt;• Tumor genetics (eg, anaplastic lymphoma kinase positivity)</td>
<td>• Religious preference&lt;br&gt;• Nutritional support&lt;br&gt;• Health literacy&lt;br&gt;• Educational level</td>
</tr>
<tr>
<td>Parkinson disease</td>
<td>• With or without cognitive impairment&lt;br&gt;• Depression&lt;br&gt;• Neurogenic bladder&lt;br&gt;• Polypharmacy</td>
<td>• Extent of functional limitation&lt;br&gt;• Behavioral health specialist involvement&lt;br&gt;• Classes of resistant organisms</td>
<td>• Lack of transportation&lt;br&gt;• Rural residence&lt;br&gt;• Social isolation</td>
</tr>
</tbody>
</table>
have an air conditioner? Does the patient live in a neighborhood with limited access to healthy food?

Widespread application of integrated data networks does not yet exist, but recent efforts offer a sense of how multidimensional assessments of patient needs could focus care delivery. Parkland Health and Hospital System in Texas, for example, has developed a risk model to predict the likelihood of readmission for patients with heart failure based on 29 clinical, social, behavioral, and utilization factors. The safety-net health system concentrates an intensive set of interventions—counseling, monitoring, specialty follow-up care, pharmacy, and care management services—on patients identified as highest risk, allowing more efficient use of limited resources. The program has resulted in a 26% relative reduction in the odds of readmission.

Using a grant from the Center for Medicare & Medicaid Innovation, Denver Health, a vertically integrated health system in Colorado, developed a predictive algorithm to divide patients into 4 risk tiers and personalize care based on their medical and social needs. The Department of Veterans Affairs has developed hot-spotter algorithms using population-level data to identify homeless veterans at risk for high healthcare utilization and then target services such as housing, food assistance, care management, and substance use support. Comprehensive data networks would enrich these efforts and allow health systems to tailor and scale programs for patients with similar profiles.

Moving toward precision health for HNHC patients will require greater public investment in data-sharing infrastructure. The NYC-CDRN, for example, is supported by PCORI, and HHS could allocate more funding to promote interoperability of electronic health records and facilitate research across data-sharing networks. Payment models that create incentives for institutions to share data may help allay health systems’ financial and competitive concerns. Dedicated training of a multidisciplinary informatics workforce, including health services researchers, computer scientists, and systems engineers, will also be necessary to compile, organize, and clean data.

Serious challenges remain. Data integration and transparency raise privacy concerns, and health systems must work to relieve patient fears about data security. Managing expansive data networks is expensive and will likely require additional or reallocated funding. It is also not clear who should “control” such data networks and how competitors should share potential costs or savings. Finally, the extent to which care models based on integrated data networks can improve health outcomes will require rigorous and prolonged study.

Nonetheless, better health, especially for HNHC patients, will likely require more precise population health. It is increasingly clear that HNHC patients are not a homogenous group, but acting on this insight requires integrated multifaceted data networks that we have not traditionally had. Investing in these networks would allow us to better anticipate the needs of patients, conceptualize care models to meet those needs, and put targeted interventions into action.

**Author Affiliations:** Department of Healthcare Policy & Research (DK, RK), and Department of Medicine (DK), Weill Cornell Medical College, New York, NY.

**Source of Funding:** None.

**Author Disclosures:** The authors report no relationship or financial interest with any entity that would pose a conflict of interest with the subject matter of this article.

**Authorship Information:** Concept and design (DK, RK); acquisition of data (RK); analysis and interpretation of data (RK); drafting of the manuscript (DK, RK); critical revision of the manuscript for important intellectual content (DK, RK); statistical analysis (RK); provision of patients or study materials (RK); obtaining funding (RK); administrative, technical, or logistic support (RK); and supervision (RK).

**Address Correspondence to:** Dhruv Khullar, MD, MPP, Department of Healthcare Policy & Research, Department of Medicine, Weill Cornell Medical College, 402 E 67th St, New York, NY 10065. Email: khd9010@med.cornell.edu.

**REFERENCES**


Full text and PDF at [www.ajmc.com](http://www.ajmc.com)