



OFFICE OF
HEALTH POLICY



OFFICE OF THE SECRETARY
PATIENT-CENTERED OUTCOMES
RESEARCH TRUST FUND

PROJECT REPORT

Building Data Capacity for Patient-Centered Outcomes Research (PCOR) for COVID-19

Highlights of Seven OS-PCORTF Funded Multiagency Data Initiatives

June 2021

Department of Health and Human Services (HHS)

Office of the Assistant Secretary for Planning and Evaluation

The Assistant Secretary for Planning and Evaluation (ASPE) advises the Secretary of the U.S. Department of Health and Human Services (HHS) on policy development in health, disability, human services, data, and science; and provides advice and analysis on economic policy. ASPE leads special initiatives; coordinates the Department's evaluation, research, and demonstration activities; and manages cross-Department planning activities such as strategic planning, legislative planning, and review of regulations. Integral to this role, ASPE conducts research and evaluation studies; develops policy analyses; and estimates the cost and benefits of policy alternatives under consideration by the Department or Congress.

The Office of Health Policy

The Office of Health Policy (HP) provides a cross-cutting policy perspective that bridges Departmental programs, public and private sector activities, and the research community, in order to develop, analyze, coordinate, and provide leadership on health policy issues for the Secretary. HP carries out this mission by conducting policy, economic and budget analyses, assisting in the development and review of regulations, assisting in the development and formulation of budgets and legislation, and assisting in survey design efforts, as well as conducting and coordinating research, evaluation, and information dissemination on issues relating to health policy.

For additional information about this subject, you can visit the HP home page at <https://aspe.hhs.gov/office-health-policy-hp> or contact the ASPE Project Officers, at: HHS/ASPE/HP

Room 4447D, H.H. Humphrey Building
200 Independence Avenue, S.W.
Washington, D.C., 20201

Scott.Smith@hhs.gov

Marcos.Trevino@hhs.gov

Authors

Violanda Grigorescu, MD, MPSH

Scott R. Smith, Ph.D.

Marcos Trevino, MPH

Division of Health Care Quality and Outcomes, Office of Health Policy, Assistant Secretary for Planning and Evaluation, U.S. Department of Health and Human Services; Washington, DC

Neala R. Campbell, MA

The MITRE Corporation; McLean, VA

Acknowledgements

This report was prepared under contract 75FCMC18D0047-75P00120F80242. We would like to thank all Fiscal Year 2021 (FY21) Office of the Secretary's Patient-Centered Outcomes Research Trust Fund (OS-PCORTF) projects leads and co-leads for their contributions to this report.

Executive Summary

As the coronavirus disease 2019 (COVID-19) pandemic enters its second year, many important questions remain about the health outcomes associated with treatment, prevention, and care of people with COVID-19. Why do some patients have infection but no symptoms, while some otherwise healthy people die or have symptoms that persist for months? Why are some population groups at higher risk of COVID infection, hospitalization, and death than other population groups? Which therapies are most effective for different groups of patients? What is the effectiveness of telehealth modalities on patient outcomes during the pandemic? Answering these questions requires a robust data infrastructure to understand health outcomes, effectiveness of interventions, and strategies for addressing inequities.

In 2021, the Office of the Secretary's Patient-Centered Outcomes Research Trust Fund (OS-PCORTF) funded seven new projects to strengthen the data infrastructure for conducting patient-centered outcomes research (PCOR) on COVID-19, during and after the pandemic.

In general, these projects either expand the use of existing data or link datasets across Federal and State sources. Some projects advance modeling using [artificial intelligence \(AI\)](#) and [machine learning \(ML\)](#) to identify subpopulations at risk for poorer outcomes, sequelae, and the impacts of care settings and community characteristics. Other projects gather critical data on the effectiveness of repurposed treatments and vaccines for COVID-19.

The project selection for funding was based on addressing at least one of four thematic priority areas:

- Leveraging Health Data and Methods
- Social and Medical Risk Factors
- Therapeutics and Vaccines
- Technology and the Pandemic

This report describes the new seven projects within the context of the four priority areas and their added value to the evidence base on risk factors for COVID-19 morbidity and mortality; appropriate treatments, interventions, and prevention approaches; and how effectiveness of these approaches varies across populations. In addition, the report introduces themes identified through conversations with these projects as opportunities for their cross collaboration and with other partners outside the federal government. That is aligned with OS-PCORTF role "to provide for coordination of relevant HSS programs to build data capacity for comparative effectiveness clinical research."

Table of Contents

1	INTRODUCTION	1
2	OS-PCORTF COVID-19 PROJECTS.....	4
	2.1 Leveraging Health Data and Methods.....	4
	2.1.1 Overview.....	4
	2.1.2 COVID-19 Projects: Current and Future Contributions.....	5
	2.2 Social and Medical Risk Factors.....	8
	2.2.1 Overview.....	8
	2.3 Therapeutics and Vaccines.....	11
	2.3.1 Overview.....	11
	2.3.2 Research Questions Related to PCOR and COVID-19	11
	2.3.3 COVID-19 Projects: Current and Future Contributions.....	12
	2.4 Technology and the Pandemic	13
	2.4.1 Overview.....	13
	2.4.2 COVID-19 Projects: Current and Future Contributions.....	14
3	COVID-19 OS-PCORTF PROJECTS THEMATIC ANALYSIS	16
4	CONCLUSION	17
	ACRONYMS.....	18
	GLOSSARY.....	20
	APPENDIX A: COVID-19 STEERING COMMITTEE MEMBER LIST	22
	APPENDIX B: FY21 OS-PCORTF PROJECT DESCRIPTIONS.....	23
	APPENDIX C: SEMI-STRUCTURED DISCUSSION QUESTIONS.....	26
	APPENDIX D: TAGS USED IN QUALITATIVE ANALYSIS OF SEMI-STRUCTURED DISCUSSIONS.....	27
	ENDNOTES	29

List Of Tables

Table 1. FY21 OS-PCORTF Initial Projects Goals.....	3
Table 2. Added Value from New Methods, by Project	7
Table 3. Social and Medical Risk Factors by Project	10
Table 4. Anticipated Impact of Projects on Therapeutics and Vaccines Research	13
Table 5. Anticipated Impact of Projects Examining Telehealth	15
Table 6. Description of AI Techniques by Project	15

1 Introduction

The coronavirus disease 2019 (COVID-19) pandemic has challenged the U.S. healthcare system in many ways. The pandemic has disproportionately impacted the elderly, people with disabilities, and people of color.¹ While COVID-19 has largely spared children and adolescents, a small subset of this group has experienced severe outcomes.² An important policy priority is building an evidence base to identify risk factors for COVID-19 morbidity and mortality, as well as understanding which treatments, interventions, and prevention approaches are effective for which populations.

At the same time, COVID-19 has forced system-level changes that accelerated innovation and learning across the healthcare system. The Office of the Secretary of Health and Human Services (HHS) has supported these changes through public policies such as new pathways for reimbursement, new diagnosis and procedure codes, and program changes to facilitate the rapid adoption of telehealth. Other investments include developing and rapidly manufacturing millions of vaccine doses, and employing novel methods, such as adaptive clinical trials, to rapidly test existing medications for potential repurposing to treat COVID-19.^{3,4,5}

Despite these efforts, in a single year (2020), COVID-19 became the third leading cause of death in the U.S. More than half a million Americans died of COVID-19, with death rates for indigenous Americans, Hispanics, and Blacks more than twice as high as for whites.⁶ These statistics point to the critical need for high-quality evidence to understand how COVID-19 affects health outcomes, which interventions can improve patient outcomes, and what strategies may help to address inequities across different population groups.

Strengthening data infrastructure for research is needed to create and advance that kind of evidence and inform many of the critical questions being asked by healthcare providers in the field about how to treat their patients and save lives. In addition, patients face questions when making healthcare choices, and understanding patient preferences and values of care is a key facet of studying the effectiveness of prevention and treatment options in patient-centered outcomes research (PCOR).⁷

PCOR is research designed to provide evidence that informs the choices facing patients and their caregivers when making health care decisions.⁸ PCOR is supported by the Patient-Centered Outcomes Research Trust Fund (PCORTF). A portion of the PCORTF is managed by the Office of the Assistant Secretary for Planning and Evaluation (ASPE) on behalf of HHS. The authority for this spending specifies, *“The Secretary shall provide for the coordination of relevant Federal health programs to build data capacity for comparative clinical effectiveness research, including the development and use of clinical registries and health outcomes research networks, in order to develop and maintain a comprehensive, interoperable data network to collect, link, and analyze data on outcomes and effectiveness from multiple sources including electronic health records [EHRs].”*⁹

Every year, HHS agencies are invited to respond to an Office of the Secretary’s Patient-Centered Outcomes Research Trust Fund (OS-PCORTF) solicitation for proposals on topics identified as priorities at that time. In light of the COVID-19 pandemic, ASPE convened a Steering Committee in the summer of 2020 that consisted of key experts selected by the Leadership Council to discuss and identify funding priorities for the fiscal year 2021 (FY21) OS-PCORTF solicitation. Members of the Steering Committee are listed in [Appendix A](#). In addition to asking all members to complete a questionnaire at the beginning of the planning process, three meetings were held to discuss agencies’ priorities and data gaps. After careful deliberation, the Steering Committee identified four priority areas for PCOR data on COVID-19, which ASPE further described in a solicitation for proposals from across HHS agencies.¹⁰ The four thematic priority areas were:

1. **Leveraging Health Data and Methods:** Data standardization, harmonization, interoperability, and integration; data linkages across settings and time to create patient cohorts for longitudinal studies; and exploration of new computational tools and methods to improve data processing.
2. **Social and Medical Risk Factors:** Integration of data on social determinants of health (SDOH) and community services in health systems and community services, behavioral and mental health as well as including different populations at higher health risk such as those living in institutional settings or having multiple chronic conditions (MCC) or pregnant and postpartum women.
3. **Therapeutics and Vaccines:** Post-market evaluation of new treatments and vaccines, including risks and benefits in large populations of patients as well as equity across patient subgroups defined by differences in SDOH.
4. **Technology and the Pandemic:** Effectiveness of telehealth services for a range of health services, including mental health and opioids, and infrastructure to permit data science methods such as [artificial intelligence \(AI\)](#) and [machine learning \(ML\)](#).

Through a well-established and transparent review process, seven project proposals were selected for funding in the Fiscal Year (FY) 2021. These new funded projects were not designed to solely answer all critical data gaps and research questions related to COVID. Their contribution is rather focused on strengthening data infrastructure for PCOR and use cases that in turn will inform data strategies and collaborations across health care system, as well as policy decisions. In addition, projects teams and research partners engaged in the implementation process have the expertise to add their contribution in designing and conducting patient-centered outcomes research studies that will fill in the knowledge gaps.

This report is structured in sections describing these seven funded projects, cutting across all four thematic priority areas. It also discusses their expected contribution to data infrastructure that supports advancing evidence created through PCOR during and post pandemic. Projects' goals are presented in Table 1 below and brief descriptions can be found in [Appendix B](#). Finally, the report discusses opportunities for collaboration among projects and other partners as well as future steps in addressing remaining gaps.

Table 1. FY21 OS-PCORTF Initial Projects Goals

Project Title	Lead Agency*	Goal(s)
CURE ID: Aggregating and Analyzing COVID-19 Treatments from EHRs and Registries Globally	FDA	<p>Identify promising COVID-19 treatments through finding new uses for approved therapeutics (“repurposed” drugs). The project expands the Food and Drug Administration (FDA) and the National Center for Advancing Translational Sciences’ (NCATS) existing CURE ID platform to enable collection and automated extraction of data from two additional sources: (1) EHRs and (2) clinical disease registries.</p> <p>Data provided by the program will:</p> <ul style="list-style-type: none"> • Give healthcare providers real-time insight into the use of existing drugs through the implementation of visualization tools within CURE ID, helping to indicate those that are promising versus those that may be unhelpful or even harmful. • Offer more advanced analytics serving the greater scientific community to guide new drug development and lead optimization based on clinical experience.
Using Machine Learning Techniques to Enable HIE Data Sharing to Support COVID-19 Focused PCOR	ONC	<p>Enhance data infrastructure and data availability for PCOR by:</p> <ul style="list-style-type: none"> • advancing implementation and use of the United States Core Data for Interoperability (USCDI) standard and the HL7® Bulk FHIR® (Fast Healthcare Interoperability Resources) API among Health Information Exchanges (HIEs); and • demonstrating the utility of these implementations by applying a privacy-preserving, ML split learning technique to develop predictive models related to patient-centered outcomes (PCO) and COVID-19 across three participating HIEs. <p>An important component of this project is the task of sharing the knowledge that is generated with HIE stakeholders and PCOR researchers. By doing so, other HIEs and research organizations can observe the benefits of these standards and methods, leading to potentially wider adoption of these techniques and participation in PCOR.</p>
Understanding COVID-19 Risks, Course, and Outcomes in the Context of MCC Through e-Care Plan Development	AHRQ NIH/NDDK	<p>Add COVID-19-relevant data elements to standards-based, comprehensive electronic care (e-Care) plan tools. These tools are designed to facilitate collection, aggregation and sharing of critical person-centered data across research and care settings for adults with Multiple Chronic Conditions (MCC). The data collected will inform risk and protective factors, natural history, and near- and long-term outcomes in COVID-19 survivors, especially in high-risk MCC populations. The project will also include the development of a caregiver-facing e-Care plan application, which will complement the existing clinician- and patient-facing applications.</p>
A National COVID-19 Longitudinal Research Database Linked to CMS Data	NIH/NCATS	<p>Evaluate the U.S. health system’s response and enable analyses of patient care utilization patterns for COVID-19 over time, by creating a longitudinal research dataset. The new Federal COVID-19 Longitudinal PCOR Database will be built by linking and triangulating three sources of data: Provider characteristics and AHRQ’s U.S. compendium of health systems; EHR clinical data in National Institutes of Health’s (NIH’s) National COVID-19 Cohort Collaborative (N3C) dataset; and Centers for Medicare and Medicaid Services (CMS) claims data on Medicare fee-for-service (FFS), Medicare Advantage (MA), and Medicaid populations.</p>

Project Title	Lead Agency*	Goal(s)
Multistate EMS and Medicaid Dataset (MEMD): A Linked Dataset for PCOR	ASPE	Create a publicly available dataset containing linked Medicaid and Emergency Medical Services (EMS) records in a sample of States, giving researchers the ability to longitudinally analyze outcomes of Medicaid beneficiaries who engage with EMS. Research conducted by the project team will focus on behavioral health emergencies (prioritizing opioid overdose) before and during the COVID-19 pandemic. Research using MEMD can analyze the use of EMS in relation to other healthcare services to inform decisions about clinical protocols, Medicaid coverage of EMS, and the organization of health services.
Dataset on Intellectual and Developmental Disabilities: Linking Data to Enhance Person-Centered Outcomes Research	ASPE	<p>Create a dataset that links the National Core Indicators (NCI) In-Person Survey, Support Intensity Scale, Medicaid claims, and other relevant data sources for four to six states.</p> <p>The linked dataset will be used for research examining relationships between PCOs and assessed Home- and Community-Based Services (HCBS) needs, service utilization, and expenditures for services. Research findings can:</p> <ul style="list-style-type: none"> • Inform program-level quality improvement initiatives and the design and implementation of value-based payment arrangements and alternative payment models for the intellectual or developmental disabilities (ID/DD) population. • Support State and Federal policymaking and provider decision-making.
Building Infrastructure and Evidence for COVID-19 Related Research, Using Integrated Data from NCHS Data Linkage Program	CDC/NCHS	<p>Expand the utility and accessibility of NCHS linked data by developing publicly available synthetic linked data products that protect privacy while integrating the following sources of data:</p> <ul style="list-style-type: none"> • NCHS survey data on social determinants (such as race and ethnicity, healthcare access, education, income, and urbanicity) and other health-related data (e.g., health conditions, self-reported vaccination status, source of healthcare coverage) • Administrative data (e.g., healthcare utilization and spending, including vaccination uptake, from CMS) • Receipt of housing assistance from the Department of Housing and Urban Development (HUD) • Cause of death data from National Death Index (NDI). <p>The linked datasets can be used to evaluate healthcare delivery, utilization, and patient outcomes.</p>

* Full names on page 18 (Acronyms)

2 OS-PCORTF COVID-19 PROJECTS

2.1 Leveraging Health Data and Methods

2.1.1 Overview

Sources of health information on patients and populations are growing daily. They include clinical and administrative data from EHRs and medical billing claims, patient-reported information found in registries for evaluating patient outcomes and software application programs (i.e., “health apps”), and

new sources such as data from wearable devices and social media. Taken together, these rich sources of data have the potential to improve data capacity for PCOR. But the frequent lack of standardization and harmonization of data complicates their application to both PCOR and public health surveillance. Leveraging standards, increasing interoperability across disparate systems, and applying new computational tools and methods will enable researchers to harmonize data across disparate systems. Comprehensive, national datasets are not available that follow people across time, across geographies, and across care settings, which has led researchers to link disparate data sets to enable certain research projects.

The current data infrastructure for research and analysis has left many stakeholders wanting more in terms of the real-time ability to access high-quality data, to integrate it with existing sources, and to seamlessly use it to make timely decisions that can be monitored for effectiveness over time. In addition, leveraging health data and methods to improve data integrity for scientific evidence-based decision-making requires cross-agency collaboration and aligns with the spirit of the Executive Order on Restoring Trust in Government Through Scientific Integrity and Evidence-Based Policymaking.¹¹

A critical consideration for PCOR on COVID-19 is the novelty of the disease itself. Examples of the high-priority research questions around how best to leverage health data and methods include:

- What are the key data standards needed to document and follow COVID-19 risk factors, natural history, and sequelae over time?
- What data are currently missing from EHRs to understand COVID-19 patient outcomes? What sources of data can be used to fill these gaps?
- What policy changes may be needed and how can they improve data infrastructure for exploring patient outcomes throughout the continuum of care (e.g., longitudinal studies)?
- What data sources need to be linked to enable population level analyses?

2.1.2 COVID-19 Projects: Current and Future Contributions

Each of the seven FY21 OS-PCORTF COVID projects leverages health data from disparate systems using innovative methods to enhance the overall data infrastructure needed to support PCOR.

- The projects link datasets such as: 1) Medicare and Medicaid claims, 2) patient data from EHRs and registries or EHRs and claims, and 3) databases created for research purposes (e.g., N3C and the AHRQ U.S. Compendium of Health Systems). These linkages will bring together data on COVID-19, SDOH, and other clinical and patient-reported outcomes data. These linked datasets will improve data infrastructure for PCOR studies by permitting a broader range of research questions to be studied and by adding data on specific populations and conditions that would otherwise not be available to researchers.
- The projects apply standardized, innovative approaches that leverage modern web-based technologies used in other industries, to securely package and deliver data from various systems. These technologies improve upon existing data exchange standards in that they are internationally recognized, easily implemented, and scalable and have support from government and industry stakeholders. The projects are using technologies such as the Observational Medical Outcomes Partnerships ([OMOP](#)) Common Data Model (CDM) to standardize data (e.g., encounters, patients, providers, diagnoses, drugs, and procedures) across different institutions. The projects are relying on open-source standards to support

interoperability, such as [FHIR](#)[®] and Substitutable Medical Apps, Reusable Technologies ([SMART](#)) on [FHIR](#)[®].

- The projects will employ new data science methods such as automation, [AI](#), and [ML](#) to analyze data found in EHRs, registries, and other health databases. With an overwhelming volume of data accruing across HHS research projects and partners, these tools can expedite processing and analysis. [AI](#) and [ML](#) algorithms can also help to identify patterns within and across disparate datasets to elicit insights that may have not been apparent through manual processing.
- The projects will explore and leverage commonalities that exist—but are often difficult to draw out due to inconsistencies in the data from State Medicaid programs. The projects aim to achieve this by validating proof of concepts, implemented in a few States, to achieve the most effective linked datasets to support PCOR and later potentially expand these to other States.

Table 2 highlights the added value of new methods for data infrastructure by project.

Table 2. Added Value from New Methods, by Project

Project Title (Lead Agency)	New Methods
CURE ID: Aggregating and Analyzing COVID-19 Treatments from EHRs and Registries Globally (FDA)	Enable automated extraction and manual data collection from EHRs and registries into the CURE ID case report form (CRF) .
Using Machine Learning Techniques to Enable HIE Data Sharing to Support COVID-19 Focused PCOR (ONC)	Develop federated ML models that can analyze the standardized Bulk FHIR® API data, leveraging the USCDI standards, in a privacy-preserving manner.
Understanding COVID-19 Risks, Course, and Outcomes in the Context of MCC Through e-Care Plan Development (AHRQ, NIDDK)	Expand open-source, standards-based patient- and clinician-facing SMART on FHIR® e-Care plan applications and associated implementation guide (IG) by adding COVID-19-relevant data elements and developing a complementary caregiver-facing application.
A National COVID-19 Longitudinal Research Database Linked to CMS Data (NIH)	Link EHRs from NIH's N3C platform with CMS claims to create a longitudinal, de-identified dataset housed on a secure, centralized, cloud-based platform, and shared with researchers inside and outside of government for advanced analytics. Link databases created for research such as N3C and the AHRQ U.S. Compendium of Health Systems
Multistate EMS and Medicaid Dataset (MEMD): A Linked Dataset for PCOR (ASPE)	Link EMS and Medicaid data to enable a longitudinal analysis of health outcomes across care settings. Create a longitudinal dataset for up to 3-5 States representing diverse characteristics (i.e., geography, Medicaid expansion status).
Dataset on Intellectual and Developmental Disabilities: Linking Data to Enhance Person-Centered Outcomes Research (ASPE)	Link survey and administrative data to create a comprehensive dataset that permits PCOR studies of persons with ID/DD, an underserved and understudied population.
Building Infrastructure and Evidence for COVID-19 Related Research, Using Integrated Data from NCHS Data Linkage Program (CDC/NCHS)	Integrate social determinants and other health data with NCHS data systems by creating publicly available synthetic datasets (based on the existing and new linked data sources), develop a validation mechanism for the synthetic data, and develop a public data dashboard.

To ensure ongoing progress, it is imperative to leverage the lessons learned during COVID-19 to address data infrastructure needs that will be relevant to PCOR during and after COVID-19. The FY21 OS-PCORTF project portfolio showcases a number of use cases to validate their output aimed to support PCOR and leverage health data in innovative, scalable ways. Their contribution to a data infrastructure that supports PCOR and expeditiously informs public health emergencies depends on their successful implementation.

The complexity of each project can only benefit from leveraging knowledge and expertise through collaboration across all projects and with other partners to ensure consolidation of efforts towards the same goal of improving the data infrastructure for PCOR. For continuity, future projects should build upon the foundational successes and leverage the lessons learned from the FY21 portfolio. Furthermore, the contribution of these projects to improving data quality at the point of collection remains a key

priority given the importance of consistent information across the healthcare system to allow longitudinal PCOR in a timely manner, inform decision-makers, and ensure equitable care to all patients.

2.2 Social and Medical Risk Factors

2.2.1 Overview

The COVID-19 pandemic has exacerbated vast disparities experienced by racial and ethnic minorities and other vulnerable populations across the country.¹² These disparities were well documented prior to the pandemic by researchers and clinicians studying outcomes and implementing evidence-based policy. However, COVID-19 highlights the critical need for higher-quality and better-connected data to protect, prevent, and treat those most impacted.

Research suggests that about 20 percent of modifiable health outcomes can be attributed to clinical care; the remaining 80 percent of health outcomes are estimated to be attributed to social determinants of health (SDOH).¹³ SDOH are “the conditions in which we are born, live, learn, work, play, worship, and age.”¹⁴ Addressing disparities in patient outcomes for COVID-19 including addressing SDOH factors is a priority for HHS.

Healthy People 2020 defines health equity as the “attainment of the highest level of health for all people...” through “valuing everyone equally.”¹⁵ Measuring SDOH is one method of developing and targeting interventions for vulnerable populations to help address health inequities. The Executive Order “Ensuring an Equitable Pandemic Response and Recovery” identified some of the data challenges that hinder the pandemic response including “a lack of complete data, disaggregated by race and ethnicity, on COVID-19 infection, hospitalization, and mortality rates.”¹⁶ In addition, it calls for HHS to establish a COVID-19 Health Equity Task Force responsible for “identifying data sources, proxies, or indices that would enable development of short-term targets for pandemic-related actions for such communities and populations” that will “develop, in collaboration with the heads of relevant agencies, a set of longer-term recommendations to address these data shortfalls and other foundational data challenges, including those relating to data intersectionality, that must be tackled in order to better prepare and respond to future pandemics.”¹⁷ This task force was recently created and more information can be found on the Office of Minority Health website (<https://www.minorityhealth.hhs.gov/omh/browse.aspx?lvl=3&lvlid=118>).

Improved data will allow answer key PCOR questions faced by individuals with social and medical risk factors and those providing care such as:

- Do vulnerable, low-income patients have better outcomes in safety-net, non-profit hospitals compared with similar patients who sought care from for-profit hospitals while taking into account population differences?
- Has the increased proportion of treatment in place for behavioral health emergencies during the COVID-19 pandemic been associated with changes to rates of follow-up care?
- What is the impact of COVID-19 on health behaviors and self-reported physical and mental health and how do these impacts vary across populations?
- With the goal of prioritizing and resourcing actionable best practices, what is the comparative effectiveness of different types and intensities of four basic preventive interventions—screening, isolation, contact tracing, and personal protective practices—in reducing rates of COVID-19 among individuals with disabilities who reside in group care settings?

2.2.2 COVID-19 Projects: Current and Future Contributions

All seven OS-PCORTF projects aim to build data capacity to understand the impact of COVID-19 including social and medical risk factors. Risk factors are addressed in each project, such as either race and ethnicity and/or housing, education, and geography. A common theme across all projects is leveraging existing data sources from Federal agencies such as the [CDC Social Vulnerability Index \(SVI\)](#), HUD Data, CMS claims data, and linking to EHRs, including from [HIEs](#), EMS, and Patient-Reported Outcome data to build capability to understand disparities and effectiveness of interventions.

For instance, the “CURE ID: Aggregating and Analyzing COVID-19 Treatments from EHRs and Registries Globally” project and the “Building Infrastructure and Evidence for COVID-19 Related Research, Using Integrated Data from NCHS Data Linkage Program” project are expanding the capacity to explore repurposed therapies and analyze outcomes for different treatment regimens on vulnerable populations. The “Using Machine Learning Techniques to Enable HIE to Support COVID-19-Focused PCOR,” “A National COVID-19 Longitudinal Research Database Linked to CMS Data”, “Dataset on Intellectual and Developmental Disabilities: Linking Data to Enhance Patient-Centered Outcomes”, and “Multistate EMS and Medicaid Dataset (MEMD): A Linked Dataset for Patient-Centered Outcomes” projects are creating linkages that can help researchers investigate the effects of the pandemic on their social patterns, including housing and subsequent clinical outcomes for people with ID/DD and the relationship between community-level social characteristics and health outcomes among those experiencing behavioral health emergencies. “Understanding COVID-19 Trajectory and Outcomes in the Context of MCC Through e-Care Plan Development” focuses on improving management of people with MCC that are at high risk for COVID-19 infections and complications.

Table 3 highlights the social and medical risk factors addressed by the seven projects.

Table 3. Social and Medical Risk Factors by Project

Project Title (Lead Agency)	Social and Medical Risk Factors
Building Infrastructure and Evidence for COVID-19 Related Research, Using Integrated Data from NCHS Data Linkage Program (CDC/NCHS)	Create publicly available synthetic data integrating SDOH (healthcare access, education, income, housing, and urbanicity) to enable researchers to assess associations between individual risk factors, health behaviors, and SDOH and pre-COVID-19 respiratory related healthcare seeking behavior and treatments, patterns of care, and outcomes.
Using Machine Learning Techniques to Enable HIE to Support COVID-19-Focused PCOR (ONC)	Develop PCO models with special emphasis on the inclusion of social determinant risk factors, behavioral/mental health factors, and other factors unique to people with chronic conditions.
A National COVID-19 Longitudinal Research Database Linked to CMS Data (NIH)	Evaluate disparities by looking at community characteristics using geocoding approaches and information on air quality and housing, and capture care and outcomes of vulnerable populations such as patients on dialysis, nursing home residents, and low-income individuals. The project will also examine functional health status, frailty, and mental health disorders.
CURE ID: Aggregating and Analyzing COVID-19 Treatments from EHRs and Registries Globally (FDA)	Analyze outcomes for different repurposed treatment regimens and treatment options to determine their impact on subpopulations, particularly vulnerable populations such as pregnant women and neonates, and subpopulations defined by geography, health insurance status, income, race/ethnicity, and medical conditions.
Dataset on Intellectual and Developmental Disabilities: Linking Data to Enhance Person-Centered Outcomes Research (ASPE)	Create a publicly available dataset linking survey and administrative data to enable the study of relationships between PCO and long-term and services and supports for individuals with ID/DD.
Multistate EMS and Medicaid Dataset (MEMD): A Linked Dataset for Patient-Centered Outcomes (ASPE)	Assess outcomes of Medicaid beneficiaries who engage with EMS that include data on community conditions (e.g., poverty, race/ethnicity) to identify variations in care delivery related to SDOH and understand outcomes of individuals experiencing behavioral health emergencies (e.g., opioid overdose) before and during the pandemic.
Understanding COVID-19 Trajectory and Outcomes in the Context of MCC Through e-Care Plan Development (AHRQ/NIDDK)	Expand standards-based e-Care plan tools to facilitate aggregation and sharing of person-centered-planning data, including health concerns, SDOH, complex constellations of COVID 19 sequelae, and behavioral health concerns, across disparate systems for people living with MCC, their caregivers, and their care teams.

These COVID-19 projects and potential future projects building on the lessons learned that are important for data standardization and use for PCOR may inform initiatives like the [Gravity Project](#) that has the potential to improve risk factor collection for vulnerable populations through their work on expanding the categories of SDOH standards for which it has balloted and created [FHIR®](#) implementation guides.

Additionally, many platforms have been developed using data from Federal agencies that can be used to identify disparities in COVID-19 infections, hospitalizations, and deaths, including CDC's [COVID-NET³³](#), NIH's [N3C Data Enclave](#), the CMS Office of Minority Health [Mapping Medicare Disparities Tool](#), and AHRQ's [PCOR SDOH Platform](#). There is an opportunity for projects to collaborate with these organizations to create visualization, analysis, and public use datasets and disseminate them to researchers and the public.

2.3 Therapeutics and Vaccines

2.3.1 Overview

Development of prophylactic and therapeutic countermeasures and vaccines rapidly became a national priority to reduce morbidity and mortality associated with COVID-19. Prophylactic medicines, including vaccines, are typically given to healthy or exposed individuals to prevent disease; therapeutic medicines are given to infected individuals to treat disease.

Evaluation of the efficacy of therapeutics and vaccines requires access to quality data to support rapid research in the context of a novel and highly contagious pandemic pathogen. In addition, the data needs and research questions evolved during each phase of the pandemic, ranging from early questions surrounding repurposed medicines, to later questions surrounding vaccine prioritization and efficacy, and persistent questions around “[Long COVID-19](#)” and other long-term patient outcomes. Improving the data collection and infrastructure is also outlined in the National Strategy for the COVID-19 Response and Pandemic Preparedness.¹²

Identifying patterns in disease progression and evolution, safety and efficacy of therapeutics, and the demographic disparities of COVID-19’s impact can help researchers understand which treatments work for specific subpopulations of patients and use this information to drive patient care that produces better health outcomes.¹⁸ Researchers have traditionally relied on time- and resource-intensive randomized controlled trials (RCT) to study the efficacy of treatment interventions. Although these studies are the gold standard for eliminating potential sources of bias, these studies take time for protocol development, institutional review board (IRB) approval, and patient recruitment that increase the costs and timeline of conducting clinical research.

Real-world evidence (RWE) derived from real-world data (RWD) can provide more timely insights to inform early clinical decision-making and can be validated later through RCTs. RWE is clinical evidence regarding the usage, or potential benefits or risks, of a medical product derived from analysis of sources other than traditional clinical trials.¹⁹ The FDA further defines RWD as data relating to patient health status and/or the delivery of healthcare routinely collected from a variety of sources. Various study designs such as pragmatic randomized trials and observational studies can be used on RWD to generate RWE.²⁰ It is important to note that RWD is only as good as the quality, accessibility, and frequency with which it is updated. Advanced analytic methods and techniques such as [AI](#) and [ML](#) can bring RWE to bear to address dynamic public health emergencies now and enhance outcomes research in the future, thus representing a clear transition from surveillance to research. [Risk prediction models](#) that use [ML](#) techniques could be applied to pandemic response to better predict patient outcomes, optimize patient triage, and improve allocation of medical resources.²¹

Research questions related to PCOR and COVID-19 can be grouped in different categories and examples include:

- **Repurposed therapeutic and prophylactic medicines**
 - Which prescription drugs are also safe and effective for new indications associated with treating COVID-19 and other infectious diseases?
- **Individual COVID-19 risk factors**
 - How do an individuals’ prior/current medications affect therapeutic outcomes?

- Are COVID-19 patients more likely to receive early combination therapies for COVID-19 in hospitals affiliated with academic institutions and large health systems, compared with community hospitals?
- **Vaccine efficacy, prioritization, deployment**
 - How do COVID-19 vaccination coverage trends differ by key population subgroups (e.g., race/ethnicity), compared to prior observed trends of vaccination coverage for influenza and pneumonia?
 - What are long-term health outcomes among those vaccinated for COVID-19 or unvaccinated?
- **Long COVID-19**
 - What are the long-term sequelae of those who have been infected with COVID-19? Do these long-term sequelae differ for those with prolonged COVID-19 infections?
 - What is the natural history of COVID-19 infection?
 - What treatments or interventions can improve health outcomes in patients with [Long COVID-19](#) infections?

2.3.2 COVID-19 Projects: Current and Future Contributions

Four of the seven proposed projects will provide capacity for PCOR related to therapeutics and vaccines. Table 4 explains the anticipated impacts each of these projects will have on PCOR related to therapeutics and vaccines. These projects will create data infrastructure based on EHR, registry, and [HIE](#) data to identify potential safe and effective therapeutics for COVID-19. “Using Machine Learning Techniques to Enable [HIE](#) Data Sharing to Support COVID-19 Focused PCOR” will enable researchers to access longitudinal records to discover trends in treatments across multiple health settings. The projects titled “CURE ID: Aggregating and Analyzing COVID-19 Treatments from EHRs and Registries Globally” and creating “A National COVID-19 Longitudinal Research Database Linked to CMS Data ” will build capacity to assess outcomes and treatment patterns for COVID-19 patients in different healthcare settings (system-affiliated versus non-system practices) and whether prior medications may affect therapeutic outcomes. In addition, newly created [synthetic](#) public-use NCHS-linked data resources developed in the “Building Evidence for COVID-19 Related Research, Using Integrated Data from NCHS Data Linkage Program” project may also be used to compare outcomes of different treatment regimens and vaccination uptake for different subpopulations based on prior linkages with Medicare and mortality data.

Table 4. Anticipated Impact of Projects on Therapeutics and Vaccines Research

Project Title (Lead Agency)	Anticipated Impact
CURE ID: Aggregating and Analyzing COVID-19 Treatments from EHRs and Registries Globally (FDA)	Enable aggregation of EHR data for evaluation of efficacy of repurposed drugs on COVID-19.
Using Machine Learning Techniques to Enable HIE Data Sharing to Support COVID-19 Focused PCOR (ONC)	Upgrade HIE infrastructure to enable application of AI/ML techniques to predict and understand COVID-19 outcomes and evaluate COVID-19 treatments.
A National COVID-19 Longitudinal Research Database Linked to CMS Data (NIH)	Enable assessment of longitudinal COVID-19 outcomes to support clinical trial design and evaluation of therapeutics and vaccines, including interactions between medicines.
Building Evidence for COVID-19 Related Research, Using Integrated Data from NCHS Data Linkage Program	Create new data linkages and provide access to synthetic data that could enable vaccine and therapeutic efficacy research.

The FY21 project portfolio includes collaborations with Federal and State agencies and the private sector that will allow for quicker access to datasets in the future. Future iterations of the OS-PCORTF portfolio may benefit the priority area of therapeutics and vaccines by leveraging the existing FY21 PCORTF COVID-19 project outputs.

Additionally, novel datasets, including genomic data, may enhance PCOR to better understand evolving genetic variants of a virus or other infectious agent, or to better understand the impact of an individual's genome on the benefit or harm from a particular therapeutic.

Furthermore, COVID-19 initially emerged and spread in countries outside of the U.S., so developing expanded capacity to access PCOR data from other countries like CUREID, one of these seven projects, will help to identify therapeutics and vaccines as quickly as possible—particularly in a pandemic or public health emergency.

2.4 Technology and the Pandemic

2.4.1 Overview

The COVID-19 pandemic has highlighted the need for technological approaches and solutions in healthcare provision and research in ways that were not apparent previously. In particular, the pandemic has illustrated the need for expanded and improved telehealth and AI methods and techniques.

Telehealth

COVID-19 disrupted healthcare delivery as many patients changed their care-seeking patterns or delayed elective procedures, while providers adjusted their practices. Although the provision of healthcare services through electronic means (e.g., emails, phone calls, real-time video connections) has increased over time as new technologies emerged, the response to the pandemic facilitated broad adoption of telehealth among patients and providers who had not previously used it.

The transition to telehealth from in-person visits has allowed continued care for patients, especially in the outpatient setting. However, certain medical and surgical procedures have been postponed, which can have further implications on patient outcomes and the healthcare system.²² It will be important to monitor patient outcomes as these services are deferred to help make decisions about the continued

use of telehealth services and resuming procedures that require in-person provision of healthcare services.

Outcomes research will be essential in determining the access, utilization, cost and quality of certain telehealth visits (e.g., video and telephone) and their impacts during COVID-19 on the healthcare system, as well as answering other key research questions as discussed below.

Artificial Intelligence

[AI](#) techniques and approaches, such as [ML](#) and [NLP](#), will be necessary to uncover PCOR insights, especially when looking for associations or patterns in large datasets. [AI](#) also has the potential to help in many areas, including diagnosing COVID-19 with radiographic chest imaging, predicting mortality risk of COVID-19, and as mentioned previously, identifying vaccines and treatments for COVID-19.²³ Although essential for PCOR, using AI requires considerable data infrastructure and processing power to draw insights in real time on large datasets.

Overall, changes in how patients receive care during COVID pandemic due to implementation of telehealth have raised previously unaddressed questions and decision points for patients and their caregivers that can be answered using [AI](#) techniques. Examples include:

- Were there differences in health-related outcomes (e.g., length of hospital stays and severity of COVID-19 related illness) by key population subgroups (e.g., race/ethnicity) and the accessibility and utilization of health care (such as the use of telehealth during the COVID-19 pandemic)?
- How do subgroups differ in their ability to substitute telehealth for in-person services (e.g., by geography, income, broadband access or type of health care coverage)?
- Models developed through [split-learning techniques](#) may be able to answer the following questions:
 - What is the probability that individuals will contract COVID-19?
 - What is the probability that an individual who tested positive for COVID-19 will develop an adverse outcome?
 - What is the probability that a person with certain medical risk factors will be admitted to the hospital because of COVID-19?

2.4.2 COVID-19 Projects: Current and Future Contributions

Two of the projects are integrating telehealth data to compare the effectiveness and impact on patient outcomes compared to in-person services. Telehealth data is being captured in “Building Infrastructure and Evidence for COVID-19 Related Research, Using Integrated Data from NCHS Data Linkage Program” from patient surveys such as the National Health Interview Survey (NHIS) and from claims data in “A National COVID-19 Longitudinal Research Database Linked to CMS Data.” Table 5 describes how the projects will incorporate insights in telehealth delivery for PCOR.

Table 5. Anticipated Impact of Projects Examining Telehealth

Project Title (Lead Agency)	Anticipated Impact
A National COVID-19 Longitudinal Research Database Linked to CMS Data (NIH)	Allow researchers to conduct longitudinal studies to determine how the healthcare system is operating during the COVID-19 pandemic. For example, researchers can evaluate if receipt of primary care services via telehealth prior to a COVID-19 diagnosis led to early treatment and better COVID-19 outcomes.
Building Infrastructure and Evidence for COVID-19 Related Research, Using Integrated Data from NCHS Data Linkage Program (CDC/NCHS)	Create the infrastructure to develop synthetic, publicly available dataset(s) using existing and future linkages of NCHS survey data and administrative records that can be used to evaluate healthcare delivery, utilization including telehealth (when it becomes available), and health outcomes.

Additionally, four of the seven projects will utilize various [AI](#) techniques to build data infrastructure and disseminate data to researchers to provide insights into COVID-19 treatment, patient outcomes, and healthcare delivery. Employing [ML](#), [NLP](#), and other [AI](#) techniques to build linked datasets and enhance data platforms will allow innovative and complex research on patient outcomes and improve the interactions among researchers. For example, "Using Machine Learning Techniques to Enable HIE Data Sharing to Support COVID-19 Focused PCOR" will allow researchers to access and use data to create predictive models to answer key research questions related to PCOR.

Table 6 describes how the proposed projects plan to leverage [AI](#) techniques to improve data capacity.

Table 6. Description of AI Techniques by Project

Project Title (Lead Agency)	AI Techniques
CURE ID: Aggregating and Analyzing COVID-19 Treatments from EHRs and Registries Globally (FDA)	Allow researchers to use ML and AI technologies for automated extraction of structured data with the potential to use more advanced AI methods to analyze free text captured in EHRs and registries.
Using Machine Learning Techniques to Enable HIE Data Sharing to Support COVID-19 Focused PCOR (ONC)	Develop a privacy-preserving federated ML model using split-learning methods for HIE data at participating HIE sites.
A National COVID-19 Longitudinal Research Database Linked to CMS Data (NIH)	Using ML and NLP to process data in N3C .
Building Infrastructure and Evidence for COVID-19 Related Research, Using Integrated Data from NCHS Data Linkage Program (CDC/NCHS)	Using ML and NLP techniques to build a publicly available synthetic dataset .

Future iterations of the OS-PCORTF portfolio may benefit from integrating additional telehealth data such as data from wearable devices. Wearable devices can be used for the entire continuum of care for COVID-19: prevention, diagnosis, treatment, and rehabilitation. For instance, symptom-monitoring devices can assess at-risk populations for vital signs such as heart rate, respiratory rate, body temperature, and blood oxygen saturation. In addition, populations diagnosed with COVID-19 can benefit from wearable technologies. The pandemic and the shift to telehealth accelerated the adoption of these wearable devices, which can be leveraged for real-time data collection and use.^{24,25}

Data from FDA-approved systems linked with EHRs and claims data can be used to answer key research questions for PCOR and provide longitudinal data for analysis. Continuing to obtain telehealth data from claims and linking to other data sources may permit research on how telehealth services affect patient outcomes for individuals with behavioral health conditions such as depression. Future PCOR projects will also benefit from continued integration of [AI](#) techniques; for instance, by facilitating extraction of unstructured clinical notes from EHRs and analysis of data from other health data sources.

3 COVID-19 OS-PCORTF Projects Thematic Analysis

To foster collegial environment for knowledge sharing and structured input seeking from an expert panel, the OS-PCORTF will implement and coordinate the OS-PCORTF COVID-19 Collaborative. In preparation for that, each of the seven projects teams was invited to participate in one-hour semi-structured discussion in order to identify themes spanning across all projects as opportunities for cross-cutting collaboration, and project technical support needs.

Projects teams had to answer eight open-ended questions regarding their project's contributions to the OS-PCORTF portfolio, potential impact on COVID-19 PCOR and future pandemics, and perceived challenges. The list of questions for these discussions can be found in [Appendix C](#).

The discussions were conducted by a contractor who used a contractor-developed proprietary data collection and qualitative analysis tool to analyze the findings. The tool allows for tracking and relating core research elements for purposes of establishing traceability and ensuring that findings are based upon available evidence. It ensures traceability of all research responses, observations, conclusions, themes, and recommendations back to an original source.

Discussion responses were tagged in the tool using keywords ([Appendix D](#)) to allow for sorting and grouping of questions and responses. Responses were distilled into observations, observations were grouped together to arrive at conclusions, and conclusions were further grouped together to arrive at themes. The first three themes identified are related to:

1. Data governance, privacy and sharing to allow longitudinal follow-up of patients across the continuum of care and disseminating for research.
2. Data standards development and implementation for timely and consistent data collection
3. Federal and private partnership as collaboration within a learning healthcare system

These findings will be discussed with the projects teams and their input will be sought for a final selection of themes for collaboration at the first meeting of the OS-PCORTF COVID-19 Collaborative that will also mark its launch.

4 Conclusion

The challenges posed by the rapidly evolving COVID-19 pandemic included the timely integration of lessons learned from the field into research, policy, and practice. The differences among individual patient symptoms and disease severity highlight the critical nature of being able to quickly capture, aggregate, update, and analyze high-quality data for PCOR and surveillance. Leveraging and linking RWD, such as EHR and SDOH data, remain key for research and assessment of public health strategies and interventions, as well as discovery and evaluation of vaccines, tests, and therapeutics designed to prevent, diagnose, and treat COVID-19 and its post-infection sequelae.

The FY21 OS-PCORTF project portfolio includes seven innovative projects that will address some of the key gaps and challenges in each of the four priority areas identified to support patient and provider decision-making around COVID-19. This portfolio is an example of Federal investment in building the foundation of “a comprehensive, interoperable data network to collect, link and analyze data on outcomes and effectiveness from multiple sources including [EHRs]”⁹ for PCOR. Also, this year’s suite of COVID-19 projects enables coordination and collaboration across projects from various Federal health programs and private entities to advance the overall goal of improving PCOR data infrastructure and support research to protect and prepare for future pandemics.

However, much work remains. These seven projects only recently embarked on their implementation journey and unknown challenges may be encountered along the way. Positive synergy resulting from group and collegial discussions and decisions through the OS-PCORTF collaborative will help overcome future and sometimes inherent hurdles by generating more ideas and creative solutions, and increased acceptance of the decision by group members and opportunity for the expression of diverse opinions.

In conclusion, the FY21 OS-PCORTF portfolio of COVID-19 projects and the new collaborative provide a foundation for future projects toward strengthening and improving the data infrastructure for PCOR and decision-making during and post pandemic. In parallel, a new strategic plan is currently in development in response to OS-PCORTF re-authorization for another 10 years. The timing of this reauthorization as the COVID-19 pandemic gripped the nation and the world presents an opportunity for the OS-PCORTF to leverage the portfolio over the next decade to bolster our nation’s capacity for PCOR and improve health outcomes during both steady state and future pandemics.

Acronyms

Term	Definition
AHRQ	Agency for Healthcare Research and Quality
AI	artificial intelligence
API	application programming interface
ASPE	Office of the Assistant Secretary for Planning and Evaluation
CDC	Centers for Disease Control and Prevention
CDER	Center for Drug Evaluation and Research
CDM	Common Data Model
CDRH	Center for Devices and Radiological Health
CMS	Centers for Medicare and Medicaid Services
COVID-19	coronavirus disease 2019
CPAP	continuous positive airway pressure
CRF	case report form
CSELS	Center for Surveillance, Epidemiology, and Laboratory Services
DIDD	Dataset on Intellectual and Developmental Disabilities
e-Care	electronic care
EHR	electronic health record
EMS	Emergency Medical Services
FDA	Food and Drug Administration
FFS	fee-for-service
FHIR®	Fast Healthcare Interoperability Resources
FY21	fiscal year 2021
HCBS	Home- and Community-Based Services
HHS	U.S. Department of Health and Human Services
HIE	Health Information Exchange
HP	Office of Health Policy
HUD	Department of Housing and Urban Development
ICD	International Classification of Diseases
ID/DD	intellectual or developmental disabilities
IRB	institutional review board
MA	Medicare Advantage
MCC	multiple chronic conditions
MEMD	Multistate EMS and Medicaid Dataset
ML	machine learning
NCATS	National Center for Advancing Translational Sciences
NCCDPHP	National Center for Chronic Disease Prevention and Health Promotion
NCHS	National Center for Health Statistics

Term	Definition
NCI	National Core Indicators
NHIS	National Health Interview Survey
NIH	National Institutes of Health
NLP	natural language processing
N3C	National COVID Cohort Collaborative
OASH	Office of the Assistant Secretary for Health
OEDA	Office of Enterprise Data and Analytics
OMOP	Observational Medical Outcomes Partnerships
ONC	Office of the National Coordinator for Health Information Technology
OS-PCORTF	Office of the Secretary’s Patient-Centered Outcomes Research Trust Fund
PCO	patient-centered outcomes
PCOR	patient-centered outcomes research
PCORTF	Patient-Centered Outcomes Research Trust Fund
RCT	randomized controlled trials
RWD	real-world data
RWE	real-world evidence
SDOH	social determinants of health
SMART	Substitutable Medical Apps, Reusable Technologies
SVI	Social Vulnerability Index
T-MSIS	Transformed Medicaid Statistical Information System
USCDI	United States Core Data for Interoperability
VA	Veterans Affairs

Glossary

Term	Definition
Agency for Healthcare Research and Quality PCOR SDOH Platform	A database on SDOH, developed with funding from the PCORTF, that links data from existing Federal datasets to be used in PCOR research, inform existing approaches for emerging health issues, and improve health outcomes. The platform is in the beta phase and the data contains variables related to five SDOH domains: social context (age, race/ethnicity, veteran status), economic context (e.g., income, unemployment rate), education, physical infrastructure (crime, housing, transportation), and healthcare context (e.g., health insurance). ²⁶
Application programming interface	Set of definitions and protocols for building and integrating application software. ²⁷
Artificial Intelligence	The theory and development of computer systems able to perform tasks that normally require human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages. ²⁸
Automated extraction	Ability to automatically pull data from various sources, including structured and semi-structured data. ²⁹
Case report form	A document used in clinical trial research to collect patient data. ³⁰
Centers for Disease Control and Prevention COVID-NET	COVID-NET is CDC’s source for data on hospitalization rates associated with COVID-19 and important clinical information on COVID-19-associated hospitalizations, including age group, sex, race/ethnicity, and underlying health conditions. ³¹
Centers for Disease Control and Prevention Social Vulnerability Index	An interactive tool and dataset that assess the negative impacts of external stressors such as national disasters on communities and on human health. The CDC SVI uses 15 U.S. census variables, including poverty, lack of vehicle access, crowded housing, and others. The SVI can help communities determine how much supplies, emergency personnel, and emergency shelters are needed before, during, and after a disaster. ³²
Centers for Medicare and Medicaid Services Office Minority Health Mapping Medicare Disparities Tool	The Centers for Medicare & Medicaid Services (CMS) Office of Minority Health (OMH) developed the Mapping Medicare Disparities Tool to understand geographic disparities in chronic disease for beneficiaries. The tool helps to identify disparities between racial and ethnic and age groups, and sexes, at the national, State/territory, or county level and understand differences in health outcomes, utilizations, spending, and quality measures. The users of the tool include government agencies, policymakers, hospitals, researchers, community-based organizations, health providers, quality improvement organizations, and the general public. Users can utilize mapping tools for visualizations and comparison and can access the database with files from CMS’s Chronic Conditions Data Warehouse, the CMS Research Identifiable Files, and the CMS Hospital Compare database. ³³
Electronic Care (e-Care)	A global term for the automation of aspects of the healthcare delivery process across administrative, clinical, and departmental boundaries in the healthcare delivery system. ³⁴
Fast Healthcare Interoperability Resources (FHIR®)	Standard for exchanging healthcare information electronically built on previous HL7 standards. ³⁵

Term	Definition
Federated Machine Learning	A machine learning technique that allows for a device to download and compute a model with localized data. The device then sends the locally trained model to a central server for aggregation. ³⁶
Gravity Project	An effort from the University of California San Francisco Social Interventions Research & Evaluation Network (SIREN) to develop consensus-driven data and interoperability SDOH standards to exchange patient-level data documented in screening, assessment/diagnosis, goal setting, and treatment/interventions. ³⁷
Health Information Exchange	Provides the ability to electronically share clinical information among disparate healthcare information systems to facilitate access to and retrieval of clinical data. ³⁸
Health Level 7 (HL7) Standards	Standards that define how electronic health information is packaged and communicated from one party to another, setting the language, structure, and data types required for seamless integration between systems. ³⁹
Homegrown	Software, tools or applications developed by IT teams to address the needs of their organizations.
Long COVID-19	Not recovering for several weeks or months following the start of symptoms that were suggested of COVID-19, whether an individual tested positive or not. ⁴⁰
Machine learning	A form of artificial intelligence that uses and develops computer systems that are able to learn and adapt without following explicit instructions, by using algorithms and statistical models to analyze and draw inferences from patterns in data. ⁴¹
Natural language processing	A form of artificial intelligence that helps computers understand, interpret, and manipulate human language, ⁴² such as speech and text.
National Institutes of Health National COVID Cohort Collaborative (N3C) Data Enclave	NIH's secure data platform with clinical data with specific domain groups focused on key topics. The N3C is making data available for the clinical and research community to use for studying COVID-19 and for identifying potential treatments as the pandemic continues to evolve. The N3C strives to coordinate and harmonize needed data derived from electronic health records to support efforts in understanding how best to direct research efforts and care of COVID-19 patients. . ⁴³
Observational Medical Outcomes Partnership	Informs the appropriate use of observational healthcare databases for studying the effects (risks and benefits) of medical products. ⁴⁴
Risk Prediction model	A mathematical equation that uses patient risk factor data to estimate the probability of a patient experiencing a healthcare outcome. ⁴⁵
Split Learning Techniques	Allows various cooperating groups to train and infer from machine learning models without sharing raw data. ⁴⁶
Substitutable Medical Apps, Reusable Technologies (SMART)	Open, standards-based technology platform that enables innovators to create apps that seamlessly and securely run across the healthcare system. ⁴⁷
Synthetic data	Data that is generated by applying a sampling technique to real-world data or by creating simulations scenarios where models and process interact to create completely new data not directly taken from the real world. ⁴⁸
United States Core Data for Interoperability	A standardized set of health data classes and constituent data elements for nationwide, interoperable health information exchange. ⁴⁹

Appendix A: COVID-19 Steering Committee Member List

The COVID-19 Steering Committee Member List reflects members from FY2020 and their titles and affiliated offices during 2020.

- David Meyers, Deputy Director, AHRQ
- Mamatha Pancholi, Chief Data Officer, AHRQ
- Adi Gundlapalli, Chief Public Health Informatics Officer, CDC—Center for Surveillance, Epidemiology, and Laboratory Services (CSELS)
- Karen Hacker, Director, National Center for Chronic Disease Prevention and Health Promotion (NCCDPHP)-CDC
- Nic Schluterman, Statistician, CMS-Office of Enterprise Data and Analytics (OEDA)
- Amy Abernethy, Principal Deputy Commissioner, FDA—Office of the Commissioner (no longer an active member)
- Scott Gordon, Senior Health Informatics Officer, FDA—Center for Drug Evaluation and Research (CDER)
- Daniel Caños, Acting Director, Office of Clinical Evidence and Analysis, FDA—Center for Devices and Radiological Health (CDRH)
- Susan Gregurick, Associate Director for Data Science, NIH
- Meena Vythilingam, Director, HHS Center for Health Innovation, Office of the Assistant Secretary for Health (OASH)
- Teresa Zayas Cabán, Chief Scientist, ONC

Appendix B: FY21 OS-PCORTF Project Descriptions

CURE ID: Aggregating and Analyzing COVID-19 Treatments from EHRs and Registries Globally (FDA)
Existing Capacity: CURE ID, a website and mobile app developed by FDA and National Center for Advancing Translational Sciences (NCATS) to capture use of existing drugs to treat infectious diseases that lack approved therapies.
New Capabilities: Enable automated extraction and manual data collection from EHRs and registries into the CURE ID CRF.
Impact for COVID-19 PCOR: Identify signals for potential COVID-19 treatments that will be safe and effective from existing FDA-approved drugs and therapeutic agents.
Alignment to Priority Areas: Therapeutics & Vaccines, Technology & the Pandemic, Leveraging Health Data & Methods, Social & Medical Risk Factors

Using Machine Learning Techniques to Enable HIE Data Sharing to Support COVID-19 Focused PCOR (ONC)
Existing Capacity: State, regional, and local HIEs, which are growing as adoption of APIs and exchange of standardized data (e.g., FHIR®) expand.
New Capabilities: Develop federated ML models that can analyze the standardized Bulk FHIR® API data, leveraging USCDI standards, in a privacy-preserving manner.
Impact for COVID-19 PCOR: Use ML methods to analyze patients' EHR data at scale and create predictive models to identify trends and practices that can improve patient outcomes related to COVID-19.
Alignment to Priority Areas: Social & Medical Risk Factors, Technology & the Pandemic, Leveraging Health Data & Methods, Therapeutics & Vaccines

Understanding COVID-19 Risks, Course, and Outcomes in the Context of MCC Through e-Care Plan Development (AHRQ, NIDDK)
Existing Capacity: Patient- and clinician-facing e-Care plan applications and associated HL7 implementation guide (IG) under development to help health care teams better manage patients with MCC.
New Capabilities: Expand the open-source, standards-based (SMART on FHIR®) e-Care plan applications and IG by adding COVID-19-relevant data elements, and developing a caregiver-facing application.
Impact for COVID-19 PCOR: Expands e-Care plan applications and standards infrastructure by adding COVID-19 risk factors, disease course, and sequelae to facilitate research to better understand risks, natural history, and near- and long-term outcomes, especially in the high-risk MCC population.
Alignment to Priority Areas: Social & Medical Risk Factors, Leveraging Health Data & Methods

A National COVID-19 Longitudinal Research Database Linked to CMS Data (NIH)
<p>Existing Capacity: Three existing datasets:</p> <ol style="list-style-type: none"> 1. Provider characteristics and AHRQ’s U.S. compendium of health systems 2. EHR clinical data in NIH’s N3C dataset 3. CMS claims data on FFS, MA, and Medicaid populations
<p>New Capabilities: Link these three sources of data to create a longitudinal, de-identified dataset housed on a secure, centralized, cloud-based platform, and shared with the Federal and outside community for advanced analytics.</p>
<p>Impact for COVID-19 PCOR: Evaluate the U.S. health system’s response and enable comprehensive longitudinal patient care utilization patterns for COVID-19, using a Federal longitudinal COVID-19 research dataset linking and triangulating three sources of data.</p>
<p>Alignment to Priority Areas: Social & Medical Risk Factors, Therapeutics & Vaccines, Technology & the Pandemic, Leveraging Health Data & Methods</p>

Multistate EMS and Medicaid Dataset (MEMD): A Linked Dataset for PCOR (ASPE)
<p>Existing Capacity: Two existing datasets:</p> <ol style="list-style-type: none"> 1. State Medicaid data 2. EMS data captured at the point of care through electronic patient care reports
<p>New Capabilities: Link data to enable a longitudinal analysis of health outcomes. Create a longitudinal dataset for up to 5 States representing diverse characteristics (e.g., geography, Medicaid expansion status).</p>
<p>Impact for COVID-19 PCOR: Ability to analyze outcomes of Medicaid beneficiaries who use EMS before and during the COVID-19 pandemic, with a specific focus on behavioral health emergencies (prioritizing opioid overdose).</p>
<p>Alignment to Priority Areas: Social & Medical Risk Factors, Leveraging Health Data & Methods</p>

Dataset on Intellectual and Developmental Disabilities: Linking Data to Enhance Person-Centered Outcome Research (ASPE)
<p>Existing Capacity:</p> <ol style="list-style-type: none"> 1. NCI In Person Survey data (survey administered by States that uses nationally standardized indicators to assess the quality of life of persons with ID/DD; captures information on community inclusion, social participation, employment, safety, choice, and decision-making, as well as the COVID-19 pandemic experience through a supplement) 2. State Medicaid claims data 3. Supports Intensity Scale HCBS needs assessment data
<p>New Capabilities: This project will link the datasets for four to six States to create a comprehensive dataset for PCOR studies of individuals with ID/DD.</p>
<p>Impact for COVID-19 PCOR: Understand experiences and service use of the ID/DD population before and during the COVID-19 pandemic.</p>
<p>Alignment to Priority Areas: Social & Medical Risk Factors, Leveraging Health Data & Methods</p>

Building Infrastructure and Evidence for COVID-19 Related Research, Using Integrated Data from NCHS Data Linkage Program (CDC/NCHS)

Existing Capacity: Seven existing publicly available datasets:

1. NCHS surveys (National Health Interview Survey and National Hospital Care Survey)
2. CMS Medicare data
3. Federal housing assistance data from HUD
4. National Death Index
5. Medicare data
6. Transformed Medicaid Statistical Information System (T-MSIS)
7. Veterans Affairs (VA) data on healthcare utilization by Veterans

New Capabilities: Integrate social determinants and other health data with NCHS data systems by creating publicly available synthetic datasets (based on the linked data sources) and a public data dashboard.

Impact for COVID-19 PCOR: Assess the association between individual risk factors, health behaviors, SDOH and COVID-19-related healthcare treatments, patterns of care, and health outcomes, while protecting patients' privacy through use of synthetic data.

Alignment to Priority Areas: Leveraging Health Data & Methods, Social & Medical Risk Factors, Therapeutics & Vaccines, Technology & the Pandemic

Appendix C: Semi-Structured Discussion Questions

Question Numbers	Discussion Question
1	In your own words, briefly summarize your project and what you are trying to accomplish. What problem are you trying to solve?
1.2	What do you see as your proposal's impact on patient-centered outcomes research?
1.3	How could your project adapt and be applicable as the pandemic changes?
1.4	What is new in your approach and why do you think it will be successful? Alternate: How do you feel your project enhances, builds on existing capabilities, or creates new capabilities/practices?
2	Which OS-PCORTF functionalities will your project impact, and how?
2.1	<ul style="list-style-type: none"> • What COVID-19 OS-PCORTF priorities do you feel your project covers/impacts? (Provide list in discussion as reference)
2.2	What connections do you see with your proposal and existing PCOR projects and/or COVID-19 proposals?
3	What types of data will your project use (e.g., EHR, claims, patients' survey) and how will it be accessed? If your project will standardize the data, how will it be made interoperable (if it was not originally in this format)?
3.1	Please summarize how your data will be collected.
3.2	Describe how the data sources used by your project will be linked (i.e., integrated).
3.2.1	How do you foresee drawing COVID-19 insights?
3.3	Will the data elements be pulled/informed by data elements already present in the data sets used by the project or will they be completely new?
4	Please describe any new tools and/or technologies your project anticipates using, along with why they will be used.
5	How will you make the data available to other researchers?
5.1	How do you envision your project output being used by researchers and others?
5.2	How will you sustain/maintain the output of your project for future users?
6	Describe project team expertise that will be utilized to implement the project, along with plans to mitigate gaps in existing expertise. (Examples of expertise: data curation, standardization, and interoperability; computational tools; ML and NLP; subject matter expertise on type of data and/or the population and/or health-related topic)
7	Describe any issues you anticipate during the project's start-up, along with mitigation plans.
8	What is the ideal data infrastructure the country needs to understand COVID-19?
8.1	What gaps need to be filled to get to this ideal?

Appendix D: Tags Used in Qualitative Analysis of Semi-Structured Discussions

Tag	Definition
Agency collaborators	Current government agency partnerships
Areas for project support	Areas where the project may need external support to reach project goals and objectives
Data infrastructure	The hardware and software needed to promote data collection, sharing, and use
Data platforms	Specific types of technology being used to allow data to be governed, accessed, and used
Datasets	The specific collections of data being used for the project
Dissemination	How the project plans to deliver and/or make output(s) available to end users
End users	The group(s) that will use the output(s) from the project
Functionalities: collection of participant-provided information	Participants, including those in safety net organizations, will be able to participate more fully in clinical research by directly providing information (i.e., data points provided by the participant such as PROs).
Functionalities: linking of clinical and other data for research	Researchers will be able to follow patients across the care continuum over time, including those enrolled in clinical trials. Researchers will be able to capture the range of variables influencing health outcomes, and link clinical and other types of data (e.g., other clinical data, claims data, participant-provided information, and environmental data) required for research regardless of where the participant goes.
Functionalities: standardized collection of standardized clinical data	Researchers will be able to use standardized clinical data based on common data element standards across research projects and networks, thereby facilitating linkage and aggregation of data across data sources.
Functionalities: use of clinical data for research	Researchers will be able to utilize and analyze routinely collected clinical data for implementation of clinical studies (observational and interventional) including data relevant to assessing safety, efficacy, and adherence, as well as genetic data and Patient-Reported Outcomes (PROs).
Functionalities: use of enhanced publicly funded data systems for research	Researchers will be able to readily use, retrieve, link, and aggregate publicly funded data for research due to enhancements in publicly funded data systems.
Funding priorities: leveraging health data and methods	Standardization, harmonization, interoperability, and integration of data about the effects of the coronavirus disease 2019 (COVID-19) pandemic on patient health outcomes and the effectiveness of interventions for patients affected by COVID-19. Data linkages across settings and time to create patient cohorts for longitudinal studies. Applicants may propose linkages among two or more data sources such as existing registries, genomics data, EHRs, claims, vital records, and other program data at the individual level. Development and exploration of new computational tools and methods to improve data processing.
Funding priorities: social and medical risk factors	This includes projects focused on the integration of social and community services information into the healthcare system. Areas of integration may include, but are not limited to, behavioral health (e.g., smoking, opioids use), mental health (e.g., depression), and vulnerable populations (e.g., race and ethnicity, low income, nursing home populations, rural and inner-city geographic areas). The intent of the OS-PCORTF funding for projects to study patient outcomes related to COVID-19 is to support analysis using well-specified and interoperable definitions for data on such aspects as race/ethnicity, food insecurity, transportation, socioeconomic status, rurality, and other factors that may mediate or moderate clinical or other patient outcomes (survival, quality of life, length of hospitalization).

Tag	Definition
Funding priorities: technology and the pandemic	Telehealth: This includes projects assessing data infrastructure to study the effectiveness of telehealth services for a range of health services utilization, including but not limited to depression and behavioral issues such as smoking and opioids. Using AI for patient outcomes research including but not limited to ML and NLP, to accurately evaluate a clinical decision support rule through an EHR system.
Funding priorities: therapeutics and vaccines	This includes projects that focus on data infrastructure that would support evaluating treatment and vaccine benefits, risks, and effectiveness while being distributed to large populations; comparing side effects in patients from different population subgroups including social and medically vulnerable (e.g., elderly living in nursing homes, with MCC and disabilities); and assessing economic impact including healthcare services cost and payment models for interventions and vaccines.
Interoperability	Integration of systems and data standards to facilitate data exchange
Notable quote	A quotation that might be of use for the report
Opportunities for cross-cutting project collaboration	Opportunities to work with other projects and/or agencies to improve project outcomes
Outputs	The end product(s) of the project
Pain point	Areas of improvement, barriers, or gaps in the project plan
Positive sentiment	Positive attitudes or responses from the discussion participants
Private sector partnerships	Current private sector partnerships

Endnotes

¹ Pulrang, A. (2020, April 30). COVID-19 Teaches us a Lot About Differences in in the Disability Community. Forbes. <https://www.forbes.com/sites/andrewpulrang/2020/04/30/covid-19-teaches-us-a-lot-about-differences-in-the-disability-community/?sh=7eaae7707714>.

² Centers for Disease Control and Prevention. (2021, March 17). COVID-19 in Children & Teens. U.S. Department of Health and Human Services. <https://www.cdc.gov/coronavirus/2019-ncov/daily-life-coping/children/symptoms.html>.

³ State Telehealth Laws & Reimbursement Policies. (2020). Public Health Institute Center for Connected Health Policy. Retrieved from <https://www.cchpca.org/sites/default/files/2020-10/CCHP%2050%20STATE%20REPORT%20FALL%202020%20FINAL.pdf>.

⁴ Butler, R. (2020, December 1). More COVID-19 diagnosis and procedure codes: ICD-10-CM.PCS and MS-DRGs updated for January 1, 2021. 3M. <https://www.3mhisinsideangle.com/blog-post/more-covid-19-diagnosis-and-procedure-codes-icd-10-cm-pcs-and-ms-drgs-updated-for-january-1-2021/>.

⁵ Innovative Medicines Initiative. (n.d.). The WHO's COVID-19 clinical trial is "adaptive." What does that mean? <https://www.imi.europa.eu/news-events/newsroom/whos-covid-19-clinical-trial-adaptive-what-does-mean>.

⁶ Ahmad, F.B., Cisewski, J.A., Miniño, A., Anderson, R.N. (2021). Provisional Mortality Data—United States. *MMWR Morb Wkly Rep*, 70:519-522. <http://dx.doi.org/10.15585/mmwr.mm7014e1>.

⁷ Office of the Assistant Secretary for Planning and Evaluation. (n.d.). Patient-Centered Outcomes Research Trust Fund, U.S. Department of Health and Human Services. <https://aspe.hhs.gov/patient-centered-outcomes-research-trust-fund>.

⁸ Patient-Centered Outcomes Research Institute. (2013). Patient-Centered Outcomes Research. <https://www.pcori.org/research-results/about-our-research/patient-centered-outcomes-research>.

⁹ § 937(f) of the U.S. Public Health Service Act, 42 U.S.C. § 299b-37. (2010). <https://www.govinfo.gov/app/details/USCODE-2010-title42/USCODE-2010-title42-chap6A-subchapVII-partD-subpart2-sec299b-37/summary>.

¹⁰ Office of the Secretary Patient-Centered Outcomes Research Trust Fund. (2020). Expanding Data Capacity for Patient-Centered Outcomes Research Request for Proposals (RFP), U.S. Department of Health and Human Services.

¹¹ Executive Office of the President of the United States. (2021, January 27). Memorandum on Restoring Trust in Government Through Scientific Integrity and Evidence-Based Policymaking. <https://www.whitehouse.gov/briefing-room/presidential-actions/2021/01/27/memorandum-on-restoring-trust-in-government-through-scientific-integrity-and-evidence-based-policymaking/>.

¹² Adelle Simmons, Andre Chappel, Allison R. Kolbe, Laina Bush, and Benjamin D. Sommers. Health Disparities by Race and Ethnicity during the COVID-19 Pandemic and Policy Approaches. ASPE Issue Brief. March 26, 2021, <https://aspe.hhs.gov/system/files/pdf/265206/covid-equity-issue-brief.pdf>

¹³ Hood, C. M., Gennuso, K. P., Swain, G. R., & Catlin, B. B. (2016). County Health Rankings: Relationships Between Determinant Factors and Health Outcomes. *American journal of preventive medicine*, 50(2), 129–135. <https://doi.org/10.1016/j.amepre.2015.08.024>.

¹⁴ Centers for Disease Control and Prevention: National Center for Chronic Disease Prevention and Health Promotion. (2020, October). Social Determinants of Health. U.S. Department of Health and Human Services. <https://www.cdc.gov/chronicdisease/programs-impact/sdoh.htm>.

-
- ¹⁵ Office of Disease Prevention and Health Promotion. (2020, October). Disparities. U.S. Department of Health and Human Services. <https://www.healthypeople.gov/2020/about/foundation-health-measures/Disparities>.
- ¹⁶ Executive Office of the President of the United States. (2021, January 21). Executive Order on Ensuring an Equitable Pandemic Response and Recovery. <https://www.whitehouse.gov/briefing-room/presidential-actions/2021/01/21/executive-order-ensuring-an-equitable-pandemic-response-and-recovery/>.
- ¹⁷ Ibid.
- ¹⁸ May, T. (2020, May 21). How America's Health Data Infrastructure Is Being Used to Fight COVID-19. Medium. <https://travismay.medium.com/how-americas-health-data-infrastructure-is-being-used-to-fight-covid-19-e04893bea0f6>.
- ¹⁹ U.S. Food and Drug Administration. (2020). Real-World Evidence. U.S. Department of Health and Human Services. <https://www.fda.gov/science-research/science-and-research-special-topics/real-world-evidence>.
- ²⁰ Kent, J. (2020, August 21). How Healthcare Is Leveraging Real-World Data to Improve Outcomes. Health IT Analytics. <https://healthitanalytics.com/features/how-healthcare-is-leveraging-real-world-data-to-improve-outcomes>.
- ²¹ Assaf, D., Gutman, Y., Neuman, Y., Segal, G., Amit, S., Gefen-Halevi, S., Shilo, N., Epstein, A., Mor-Cohen, R., Biber, A., Rahav, G., Levy, I., & Tirosh, A. (2020). Utilization of machine-learning models to accurately predict the risk for critical COVID-19. *Internal and emergency medicine*, 15(8), 1435–1443. <https://doi.org/10.1007/s11739-020-02475-0>.
- ²² Wosik, J., Fudim, M., Cameron, B., Gellad, Z. F., Cho, A., Phinney, D., Curtis, S., Roman, M., Poon, E. G., Ferranti, J., Katz, J. N., & Tchong, J. (2020). Telehealth transformation: COVID-19 and the rise of virtual care. *Journal of the American Medical Informatics Association: JAMIA*, 27(6), 957–962. <https://doi.org/10.1093/jamia/ocaa067>.
- ²³ Naudé, W. (2020). Artificial intelligence vs COVID-19: limitations, constraints and pitfalls. *AI & Soc* 35, 761–765. <https://doi.org/10.1007/s00146-020-00978-0>.
- ²⁴ Christian, J., Dasgupta, N., Jordan, M., Juneja, M., Nilsen, W., & Reites, J. (2018). 21st Century Patient Registries: Registries for Evaluating Patient Outcomes: A User's Guide: 3rd Edition, Addendum [Internet]. Agency for Healthcare Research and Quality. <https://www.ncbi.nlm.nih.gov/books/NBK493822/>.
- ²⁵ Robeznieks, A. (2019). Wearable, the FDA and patient advice: What physicians should know. American Medical Association. <https://www.ama-assn.org/practice-management/digital/wearables-fda-and-patient-advice-what-physicians-should-know>.
- ²⁶ Social Determinants of Health Database (Beta Version). (December 2020). Agency for Healthcare Research and Quality, Rockville, MD. <https://www.ahrq.gov/sdoh/data-analytics/sdoh-data.htm>.
- ²⁷ RedHat. (n.d.). What is an API? <https://www.redhat.com/en/topics/api/what-are-application-programming-interfaces>.
- ²⁸ Oxford University Press. (2020). Artificial Intelligence. In Lexico.com. <https://www.lexico.com/about>. Retrieved March 23, 2021 from https://www.lexico.com/en/definition/artificial_intelligence
- ²⁹ Vaccaro, L., Sansonetti, G., & Micarelli, A. (2021). An Empirical Review of Automated Machine Learning. *Computers*, 10(1), 11. <https://doi.org/10.3390/computers10010011>.
- ³⁰ Bellary, S., Krishnankutty, B., & Latha, M. S. (2014). Basics of case report form designing in clinical research. *Perspectives in clinical research*, 5(4), 159–166. <https://doi.org/10.4103/2229-3485.140555>.

-
- ³¹ Centers for Disease Control and Prevention. (2020, August 28). Coronavirus Disease 2019 (COVID-19) Associated Hospitalization Surveillance Network (COVID-NET). Retrieved from: <https://www.cdc.gov/coronavirus/2019-ncov/covid-data/covid-net/purpose-methods.html>.
- ³² Centers for Disease Control and Prevention, Agency for Toxic Substances and Disease Registry. (2019, December 6). CDC SVI Fact Sheet. Retrieved from https://www.atsdr.cdc.gov/placeandhealth/svi/fact_sheet/fact_sheet.html.
- ³³ Centers for Medicare and Medicaid Services. (2020, December 1). CMS Office of Minority Health Mapping Medicare Disparities. Retrieved from <https://www.cms.gov/About-CMS/Agency-Information/OMH/OMH-Mapping-Medicare-Disparities>.
- ³⁴ E Care. (n.d.) Segen's Medical Dictionary. (2011). Retrieved April 8 2021 from <https://medical-dictionary.thefreedictionary.com/E+Care>.
- ³⁵ HL7 International. (n.d.). FHIR® Overview. <https://www.hl7.org/FHIR®/overview.html>.
- ³⁶ Open Data Science. (2020, April 3). What Is Federated Learning? <https://medium.com/@ODSC/what-is-federated-learning-99c7fc9bc4f5>.
- ³⁷ SIREN. (2020, September 11). The Gravity Project. Retrieved from <https://sirenetwork.ucsf.edu/TheGravityProject>.
- ³⁸ HIMSS. (n.d.). Interoperability in Healthcare. <https://www.himss.org/resources/interoperability-healthcare>
- ³⁹ HL7 International. (n.d.). Introduction to HL7 Standards. <https://www.hl7.org/implement/standards/>.
- ⁴⁰ Nabavi, N. (2020). Long covid: How to define it and how to manage it. *BMJ*;370:m3489.
- ⁴¹ Oxford University Press. (2020). Machine Learning. In Lexico.com. <https://www.lexico.com/about>. Retrieved March 23, 2021 from https://www.lexico.com/en/definition/machine_learning.
- ⁴² SAS Institute Inc. (n.d.). Natural Language Processing (NLP) What it is and why it matters. https://www.sas.com/en_us/insights/analytics/what-is-natural-language-processing-nlp.html.
- ⁴³ National Institutes of Health, National Center for Translational Sciences. (2021). Social Determinants of Health (SDoH). Retrieved from: <https://covid.cd2h.org/social-determinants>.
- ⁴⁴ Foundation for the National Institutes of Health. (n.d.). Observational Medical Outcomes Partnerships (OMOP). <https://fnih.org/what-we-do/major-completed-programs/omop>.
- ⁴⁵ Grant, S., Collins, G., & Nashef, S. (2018, August). Statistical Primer: developing and validating a risk prediction model. *European Journal of Cardi-Thoracic Surgery*, 54(2), 203-208. <https://doi.org/10.1093/ejcts/ezy180>.
- ⁴⁶ MIT Distributed Learning and Collaborative Learning. (n.d.). Split Learning: Distributed deep learning without sharing raw data. Retrieved from <https://www.media.mit.edu/projects/distributed-learning-and-collaborative-learning-1/overview/>.
- ⁴⁷ Computational Health Informatics Program. (n.d.). What is SMART? <https://smarthealthit.org/an-app-platform-for-healthcare/about/>.
- ⁴⁸ Gartner. (n.d.). Synthetic Data. <https://www.gartner.com/en/information-technology/glossary/synthetic-data>.
- ⁴⁹ Health IT. (n.d.). United States Core Data for Interoperability (USCDI). Office of the National Coordinator for Health Information Exchange. <https://www.healthit.gov/isa/united-states-core-data-interoperability-uscdi>.