UNIVERSITY OF CALIFORNIA HEALTH

UC HEALTH DATA GOVERNANCE TASK FORCE REPORT 2024

THIS PAGE INTENTIONALLY LEFT BLANK

Contents

Implementation Task Force Members

The Task Force Steering Committee and Work Groups contain members from across the UC system. A special thank you and recognition should be given for the input and work of each member for their time, energy, and passion in helping to bring these recommendations to fruition.

Steering Committee
Cora Han, Chief Health Data Officer, UC Health, UC Office of the President [Chair]
David Rubin, Executive Vice President, UC Health, UC Office of the President
Hillary Kalay, Senior Principal Counsel, UC Legal, UC Office of the President
Lars Berglund, Associate Vice Chancellor Emeritus for Biomedical Research, UC Davis
Theresa Maldonado, Vice President, Research and Innovation, UC Office of the President
Susan Cochran, Professor of Epidemiology & Biostatistics, UCLA
Van Williams, VP of Information Technology Services & Chief Information Officer, University of California

Work Group A – Justice-based Model of Health Data Use
Christine Cassel, Adjunct Professor of Medicine, UCSF [Co-Chair]
Laurie Herraiz, Associate Program Director, CTSI, UCSF [Co-Chair]
Barbara Koenig, Professor Emeritus of Medical Anthropology & Bioethics, UCSF
Shanda Hunt, Assistant Vice Chancellor for Research Compliance & Integrity, UCSD
Sarah Dry, Vice Chair, Biobanking and Research Services, UCLA Health
Vivek Rao, Faculty, Human-Centered Design, UC Berkeley Haas
Jason Phillips, Assistant Director, UCSF Health Experience, UCSF Health
Sara Ackerman, Associate Professor, Social & Behavioral Sciences, UCSF

Work Group B – Health Data Set Access Repository
Tom Andriola, Vice Chancellor of Information, Technology and Data & Chief Data Officer, UC Irvine [Chair]
Noelle Vidal, Healthcare Compliance & Privacy Officer, ECAS, UC Office of the President
Brian Russ, Research Policy Manager/Interim Contracts and Grants Operations Officer, RPAC, UC Office of the President
Nadya Balabanova, Data Scientist, UC Health Data Warehouse, UC Irvine

Work Group C – Data Sharing Guidelines
Jason Yeates Adams, Director of Data and Analytics Strategy, Associate Professor of Pulmonary, Critical Care, and Sleep Medicine, UC Davis Health [Co-Chair]
Albert Duntugan, Chief Data Officer, UCLA Health [Co-Chair]
Vanessa Ridley, Chief Compliance Officer & Chief Privacy Officer, UCSF
Erick Jenkins, Manager, Clinical Trials Contracts, UC Davis Health
Michael Hogarth, Chief Clinical Research Information Officer, UCSD Health

The UC Health Data Governance Implementation Task Force would like to thank those members of the Task Force and Work Groups who contributed their time and effort to the development of the report and recommendations but who were unable to continue their participation.

Carrie Byington (UCOP)
Rachel Nosowsky (UCSF Health)
Jennifer Chayes (UCB)
Susan Pappas (UCSF)
Aaron Neinstein (UCSF)
Lucila Ohno-Machado (Yale)

Special thanks to Pagan Morris and Jennifer Benbow of the Center for Data-driven Insights and Innovation for work group facilitation, project management, and for their design, compilation, and editing of this report.

Executive Summary

The importance of data in transforming health and medicine has never been clearer. As a leading public research and healthcare-delivery institution, the University of California (UC) has a duty to utilize its data assets, and to do so in a safe and responsible way. Extracting insights from large data sets involves collaboration and data sharing amongst clinicians, technologists, bioethicists, and others both within and outside UC. Critical to the success of these collaborations are replicable and transparent processes for managing and sharing health data responsibly.

The University of California has been on the forefront of health data governance for many years. Building upon the work of an earlier Presidential Task Force, President Michael Drake launched the UC Health Data Governance Implementation Task Force in October 2021 to develop recommendations for safe and responsible use and sharing of health data. Task Force members include scientists, clinicians, ethicists, privacy and regulatory experts, and IT specialists from across University of California Health (UCH), as well as UC Office of the President. Together, these experts developed a set of recommendations that will enable UC to be a strong data steward of its patients' data, a role that is ongoing and will evolve in the face of technological and regulatory changes.

The Task Force developed three sets of recommendations:

A. Development of a Justice-based Model of Health Data Use

- Development of a strategic plan for ongoing approaches to transparency in how UC uses patient data.
- Development of educational materials on Health Data, including what it is, what purposes it can be used for, why it is valuable, and what steps UCH patients and Californians can take to learn more.
- Enhanced governance structures, created with input from experts in bioethics, data science, and artificial intelligence.
- Creation of a public council, with representatives from communities throughout all UC medical center catchment areas, to advise the UCH Health Data Office and local governance offices on data use, data sharing, and public communications about these topics.

B. Tracking of health data collaborations

- Development of a pilot Health Data Set Access Repository managed by the UCH Health Data Office that will capture key elements of data sharing agreements.
- Development of resources within the UCH Health Data Office regarding executing data sharing agreements, including guidance, templates, and education around utilization of the Health Data Set Access Repository.
- An annual or biennial systemwide data conference, hosted by the UCH Health Data Office.

C. Updated Data Sharing Guidelines

Following a comprehensive review of the Interim Operating Guidelines, which established a provisional framework for assessing data sharing collaborations with third parties, the Task Force generated a definitive set of Data Sharing Guidelines and recommends their adoption systemwide.

Introduction

As a leading healthcare-delivery and public research institution, the University of California (UC) has developed extensive Health Data[1] assets across its campuses. Insights gleaned from the analysis of these data assets will drive advancements in healthcare delivery, discovery, and education that will in turn improve health outcomes, patient and provider experience, value, and health equity, both for UC patients and the general public. UC has both a duty to utilize these data assets to advance science and clinical care and an equally strong duty to act as a steward of this precious resource.

Data-intensive projects are already transforming health and medicine. Tools based on artificial intelligence have enhanced disease detection and prevention. Data-intensive projects provide opportunities to make delivery of care more effective, efficient, and equitable. During the coronavirus pandemic, data have been vital for tracking the disease, predicting outbreaks, and conducting studies that have advanced clinical care and saved lives.

Enabling data-driven improvement of human health requires both scientific research and translation of knowledge into new innovations. Extracting key insights from large, complex Health Data sets has become an invaluable but challenging part of the discovery process. This process requires a comprehensive approach, involving scientists and technologists spanning multiple disciplines, both from within UC and externally, as well as collaboration with government, nonprofit, and for-profit industry partners. Data sharing has thus become an indispensable mechanism by which the University works to fulfill its mission to improve the public's health.

Society stands to benefit from large-scale data efforts, but it is critical that UC address the complexity and challenges of this intricate terrain, including ethical, legal, privacy, and security risks, and genuinely engage patient communities in making decisions concerning how their data are used. To reap the full benefits of advanced data science, including artificial intelligence and machine learning, UC must build productive relationships between the guardians of patient Health Data and those who know how to make use of it. The University's role in promoting trustworthy, data-intensive collaborations is growing, and UC must ensure that this model of data use is safe and responsible. This model must also be consistently grounded in the principle of justice. This requires building trust with under-served and disenfranchised communities about how their Health Data are collected, used, and shared; fairness in the utilization of data; and the addressing of inequities so that the benefits of data use flow to all.

In response to the challenge of navigating these issues, UC President Michael Drake and former UC President Janet Napolitano launched the UC Health Data Governance Implementation Task Force ("Task Force") in October 2021 to develop recommendations for the responsible use and sharing of Health Data. This systemwide task force included scientists, clinicians, bioethicists, privacy and regulatory experts, and IT specialists. It builds upon the work presented in an earlier Ad Hoc Task Force Report on Health Data Governance released in 2018 ("Ad Hoc Task Force Report"). The Ad Hoc Task Force Report proposed initial recommendations and Interim Operating Guidelines for assessing data-intensive collaborations with third parties while long-term solutions were being considered and developed.

Since the Ad Hoc Task Force Report was released, UC Health (UCH) has made significant progress, including the establishment of a system-level Health Data Office located within UCH's Center for Data-driven Insights and Innovation that convenes a system-wide "Tiger Team" to evaluate projects involving

[1] The Task Force defines Health Data as any information pertaining to the health, care, and treatment of UC Health patients and plan members that (1) results in a report used in treatment or monitoring of a patient; (2) generates a claim or bill for services provided; (3) is used for operations, financial management, population health activities, or quality metrics; or (4) any derivation of these data, irrespective of how trivial or complex the derivations may be.

the disclosure of Health Data to a third party, and the development of local data governance groups at each health campus.

In October 2021, the Task Force began its work with the goal of accomplishing the remainder of the recommendations in the Ad Hoc Task Force Report. Specifically, this Task Force was charged with developing recommendations to address the following:

- Development of a patient-informed, justice-based model of Health Data use at UC and demonstration of the necessity and benefits of more active data use.
- Development of a system for tracking data set access (Health Data Set Access Repository or HDAR) that will capture the existence of relevant data sets, as well as all transactions with external parties (current and proposed) allowing access to them.
- Review and revision of the Interim Operating Guidelines consistent with feedback from local campus data governance groups and other UC stakeholders, and establishment of a system-level Health Data Governance Committee to replace the interim Tiger Team.

The Task Force formed three Work Groups to address each of these tasks. Based on this work, Part I of the report sets forth the key elements of a justice-based model of Health Data use and overarching UC principles for responsible Health Data governance. Part II sets forth practical steps UCH and UCH locations should take to operationalize the principles, including how to track data sharing with third parties and establish local and systemwide processes to ensure that the data are shared responsibly.

UC Responsible Health Data Governance Principles

The Ad Hoc Task Force Report set forth several overarching principles that govern transactions enabling access to UC Health Data. Drawing upon input from Task Force members and local data governance groups, the Task Force proposes these updated UC Responsible Principles for Health Data Governance, which clarify existing language and account for similar concepts that are shared across the principles.

1. **Attention to the University's Unique Responsibility and Mission**. As a public research and educational institution, the University's fundamental mission is to create and share knowledge broadly, but also responsibly and strategically. UC must safeguard sensitive information while continuing to provide broad use and dissemination of knowledge to benefit the public.
2. **Sharing Data Outside UC for Public Benefit**. The University must collaborate with others to create new knowledge and benefit the public. This means partnering with all sectors, including non-profit, for-profit, and government entities. The University must articulate a clear public benefit in any arrangement in which Health Data are shared outside of the University, especially for commercial transactions. Financial gain and commercial development, by themselves, are not a clear public benefit.
3. **Justice**. The University recognizes that there are significant inequalities in our society. Developing a shared concept of the public good requires authentic, sustained, and dynamic engagement, including outreach with the public to ensure that disenfranchised communities participate in setting health research, clinical, and policy-making agendas.
4. **Transparency and Patient Engagement**. The University should empower data subjects and be transparent about University activities involving data. At minimum, each UC Health location is encouraged to develop and publish data governance policies and procedures in accessible language and format on a publicly available UC Health location.

5. **Responsible Stewardship**. Generation of new scientific knowledge and related innovations in healthcare are increasingly dependent upon the analysis of large Health Data sets, which may require sharing data with academic, non-academic, and industry collaborators to enable use of the most advanced scientific methods available. To do so responsibly requires careful data management including a rigorous and skeptical assessment of the promised benefits of sharing Health Data, attention to patient privacy and data security, promoting the widespread dissemination of scientific findings, and ensuring that the benefit of UC Health patients and society in general are the primary motivations for any sharing of Health Data.

Part I – Development of a Patient-informed, Justice-based model of Health Data Use

Recognizing that Health Data are a valuable public good that should be used to advance the health of residents of our state and reduce health inequities, the Task Force developed recommendations aimed at developing sustained outreach with the public to ensure that patients from the diverse communities served by UC have a voice in setting health research, clinical, and policy-making agendas. The Task Force determined that elements of a justice-based model include transparency, education, and community engagement.

As a first step in its process, the Task Force organized a two-day conference in April 2022 entitled, "Got Health Data? Moving Towards a Justice-based Model of Data Use" (University of California Health, 2022). This virtual conference brought together experts across different disciplines including researchers, bioethicists, privacy experts, clinicians, patients, and industry leaders to discuss issues around Health Data use. Topics included incorporating the patient voice meaningfully and sustainably into decisions around data use, data ethics and governance, defining public good in the context of data analytics, and collaborating with government and for-profit entities.[2] In addition to insights from this conference, the work group reviewed the relevant literature[3] and sought input from internal and external experts on community engagement, artificial intelligence (AI), and other topics.[4]

The Task Force concluded that a justice-based model of data governance requires the building of trust with the public, including under-served and disenfranchised communities, about how their Health Data are collected, used, and shared; fairness in the utilization of data; and the addressing of inequities so that the benefits of data use flow to all. As experts have explained, justice demands the use of public interest as a key objective for using data.[5] In practical terms, this requires transparency to build buy-in with the public around data sharing processes. It requires education to raise awareness and understanding of data sources and data use. It demands community input to shape scientific and health goals and to set priorities for research questions. It also requires UC to develop processes to consider and address community concerns in a meaningful and transparent manner. These concerns are particularly significant for underserved and underrepresented communities with the most pressing medical needs and burden of disease.

The existing regulatory regime is insufficient to address the need for a justice-based model. Privacy and security are the focus of the current regulatory and legal framework for data governance[6]; however, this is inadequate for the complex issues, including public good and public trust, that are related to Health Data acquisition and use. Moreover, experts have recognized that the need for such a model is driven by the emergence of risks previously unaccounted for in existing models of Health Data usage.[7] For example, existing models prioritize individual consent but fail to consider the societal benefits and risks of Health

[2] *See* Appendix 2 for "Got Health Data?" Conference Proceedings Paper; "Got Health Data? Moving Toward a Justice-Based Model of Data Use," Web Conference, University of California Health, April 19-20, 2022, https://www.ucop.edu/uc-health/departments/got-health-data-moving-toward-a-justice-based-model-of-data-use-conference-april-2022.html.
[3] *See* Appendix 1 for Overview of Academic Perspectives
[4] These experts from the Biomedical Ethic Research Program at Mayo Clinic include Karen Meagher, Ph.D., Assistant Professor of Biomedical Ethics; Austin Stroud, M.A, Biomedical Ethics Research Coordinator; Richard Sharp, Ph.D., Professor of Biomedical Ethics; and Susan Curtis, M.L.S., Biomedical Ethics Research Coordinator.
[5] Angela Ballantyne and G. Owen Schaefer, "Public Interest in Health Data Research: Laying out the Conceptual Groundwork," *Journal of Medical Ethics* 46, no. 9 (September 1, 2020): 610–16, https://doi.org/10.1136/medethics-2020-106152.
[6] Eike-Henner W. Kluge, "Secure E-Health: Managing Risks to Patient Health Data," *International Journal of Medical Informatics*, "Virtual Biomedical Universities and E-Learning" and "Secure eHealth: Managing Risk to Patient Data," 76, no. 5 (May 1, 2007): 402–6, https://doi.org/10.1016/j.ijmedinf.2006.09.003.
[7] Kieran C. O'Doherty et al., "If You Build It, They Will Come: Unintended Future Uses of Organised Health Data Collections," *BMC Medical Ethics* 17, no. 1 (September 6, 2016): 54, https://doi.org/10.1186/s12910-016-0137-x

Data use, including inappropriately managed secondary use and the increasingly blurred lines between scientific research and initiatives in for-profit enterprises.[8]

There are also gaps in the protections provided by federal and state laws addressing the privacy and security of Health Data.[9] For example, current law allows health care providers to share Health Data without patient consent when those data have been de-identified pursuant to certain legal standards; however, those data present a risk of reidentification and downstream misuse when combined with other large data sets. Data generated by patients through apps or devices not tied to their healthcare provider also are not covered by HIPAA, the primary federal law protecting patient privacy.[10] Finally, while the public overwhelmingly supports efforts to improve the quality of health care and health care delivery, the public may not understand that Health Data legally is used without patient consent for quality improvement and quality assurance activities, including those designed to evaluate and improve performance in a clinical area or department. Without intentional communications and community engagement, such practices still may undermine patient trust.

In addition, traditional human subjects research ethics are centered about the process of informed consent as a reflection of the principle of "Respect for Persons" as set forth in The Belmont Report, the founding document in which the principles of ethical conduct of research with human subjects is outlined.[11] Yet informed consent is often not legally required nor practicable in the context of large scale data analytics projects, given the voluminous number of patients and minimal identifiers involved. Moreover, consent documents may not capture the nuances of data sharing that would help patients better or fully understand how their data are being used/shared, and there is ample evidence in the literature that many patients do not read these documents carefully.[12,13,14] Where consent is not feasible, the principle of respect for patient autonomy remains an important goal, and a justice-based model presents an alternative path to achieving this goal.

Data analysis holds a key to advancing science, improving patient care, and promoting public health, and these goals cannot become the casualty of a crisis in trust. The motivation for data sharing is itself based on the ethical principles of stewardship and justice, which means that sequestering the data is not an option. A justice-based model of Health Data use that incorporates transparency, education, and community involvement into a strong governance model will build the trust necessary to achieve these goals.

---

[8] Shona Kalkman et al., "Patients' and Public Views and Attitudes towards the Sharing of Health Data for Research: A Narrative Review of the Empirical Evidence," *Journal of Medical Ethics* 48, no. 1 (January 1, 2022): 3–13, https://doi.org/10.1136/medethics-2019-105651.
[9] Kristen Rosati and Jay Shaw. "Ethics of Data Use and Governance." Panel Discussion presented at the Got Health Data? Moving Toward a Justice-Based Model of Data Use Conference, April 20, 2022. https://www.youtube.com/watch?v=z15O5eM658M&t=14560s.
[10] "Examining Oversight of the Privacy & Security of Health Data Collected by Entities Not Regulated by HIPAA" (U.S. Department of Health and Human Services, July 19, 2016), https://www.healthit.gov/buzz-blog/privacy-and-security/examining-oversight-privacy-security-health-data-collected-entities-not-regulated-hipaa.
[11] "The Belmont Report" (Office for Human Research Protections (OHRP), April 18, 1979), The Belmont Report also discusses conceptions of justice as relevant to research involving human subjects, asserting, for example, that the selection of research subjects should be examined so as not systematically to be limited to certain classes of persons. https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/index.html.
[12] Mazor, Kathleen M., et al. "Stakeholders' views on data sharing in multicenter studies." Journal of comparative effectiveness research 6.6 (2017): 537-547.
[13] Beverly H Lorell et al., "Informed Consent in Clinical Research: Consensus Recommendations for Reform Identified by an Expert Interview Panel," *Clinical Trials* 12, no. 6 (December 1, 2015): 692–95, https://doi.org/10.1177/1740774515594362.
[14] Laura M. Beskow and Elizabeth Dean, "Informed Consent for Biorepositories: Assessing Prospective Participants' Understanding and Opinions," *Cancer Epidemiology, Biomarkers & Prevention* 17, no. 6 (June 16, 2008): 1440–51, https://doi.org/10.1158/1055-9965.EPI-08-0086.

Elements of a Justice-Based Model

Transparency

Openness and transparency are key ingredients to demonstrating respect for UCH patients and to building meaningful engagement. Transparency in Health Data governance can build individual and community buy-in around data sharing processes and will lead to greater support for broad data sharing. Indeed, research indicates that most individuals, while having privacy concerns, are proponents for sharing de-identified data if it results in societal gain and benefits the greater good.[15]

Meaningful transparency demands that UCH do more to optimize the patient experience, starting with visibility into how Health Data are used. The current process of notifying patients about the use of their Health Data, while legally compliant and widely employed across the industry, could be enhanced to provide greater transparency. For example, UCH's current notification practice consists of annually presenting patients with the Terms & Conditions of Service and Notice of Privacy Practice documents for signature prior to a medical encounter. Findings from an observational and interview study conducted in multiple UCSF clinics indicate that some patients did not recall ever having seen the forms.[16] For those who did, many acknowledged they did not read them before signing. Creating more user-friendly, accessible, and interactive materials, and disseminating these materials broadly through the educational methods described below would be a positive addition to the legally required forms and would present an opportunity to build trust with patients.

Transparency around the potential benefits of data sharing, including specific, relatable examples beyond the general promise of big data research is also important. Many such examples already exist within UCH and should be promoted though new and existing channels.

Finally, while Health Data research has the potential to address health inequity by increasing representation of people who have been medically underserved and historically discriminated against, UC must keep historical facts in mind. Some groups have a well-founded mistrust of medical institutions, and such reservations need to be explored and addressed through balanced, tailored communications. UC must be fully transparent with historically marginalized populations about what data are shared, and why and how they are shared. UC also must ensure that inclusive communication and meaningful transparency efforts include input from and participation with historically marginalized populations through the community engagement techniques described below.

Community Engagement

Community engagement is another pillar of a justice-based model. Effective engagement can take many forms, including, as this group recommends, direct community representation in governance processes. Ideally, patient and community representatives in governance groups should include a wide range of patient voices from diverse communities, including those who may be unaware of the varying uses of their Health Data. For community voices to be most effective, however, it is also important for representatives in governance groups to have a working knowledge of how Health Data may be collected and used for quality improvement, research, and public health efforts. There are advantages and disadvantages to selecting individuals with prior knowledge and experience in these issues, versus identifying inexperienced individuals and providing them with the necessary education to make meaningful contributions. While seeking some members who have experience with data use is helpful, it does not mitigate the need to provide adequate education about the scientific issues in Health Data

[15] Nanibaa'A Garrison et al., "A Systematic Literature Review of Individuals' Perspectives on Broad Consent and Data Sharing in the United States," *Genetics in Medicine* 18, no. 7 (July 1, 2016): 663–71, https://doi.org/10.1038/gim.2015.138.; Mhairi Aitken et al., "Public Responses to the Sharing and Linkage of Health Data for Research Purposes: A Systematic Review and Thematic Synthesis of Qualitative Studies," *BMC Medical Ethics* 17, no. 1 (November 10, 2016): 73, https://doi.org/10.1186/s12910-016-0153-x.

[16] Sara L. Ackerman et al., "Patients Assess an eConsult Model's Acceptability at 5 US Academic Medical Centers," *The Annals of Family Medicine* 18, no. 1 (January 1, 2020): 35–41, https://doi.org/10.1370/afm.2487.

science as well as the legal and regulatory contexts to all. In addition, for the specific purposes of governance input, groups of diverse individuals who are representative of the state's populace, including both UC Health patients and non-patients, should be included.

There is significant literature and experience with evidence-based approaches to deriving input directly from patients and from other relevant stakeholders. One well-described model is called deliberative democracy,[17,18,19] an extensive, iterative process that has been used toward various goals especially focused on controversial public policies. Other engagement models include patient and family advisory groups and community advisory boards, which have been established in most healthcare organizations to provide patient and community input into clinical operations, research, and institutional policies. Particularly relevant to the issues around data use is the last decade of experience with clinical research using genetic data, which is highly sensitive and where standards of patient engagement have been developed and published.[20]

The work group recognizes that the 'how' of community engagement will be specific to each institution and its network, the patients and communities served, and other stakeholder constituencies. One potentially useful paradigm for developing appropriate processes is participatory design. Participatory design engages stakeholders—in this case, patients—as co-researchers and co-creators, leveraging creative methods to envision a desired future state[21,22,23,24] The approach has been used to support ethical algorithm design, envision more effective ways of displaying patient-generated Health Data in electronic health records, and to develop infographics to improve engagement with communities with diverse levels of health literacy. [25,26,27] Regardless of whether any of these specific methods are appropriate for a given institution, a generalizable takeaway from participatory design is that patients and stakeholders are empowered not only in a consenting or feedback role, but also in actively contributing to the creation of systems, policies, and processes.

Experts consulted by the Task Force also emphasized the importance of community partnerships as a way of incorporating patient perspectives into decisions around data use. As one speaker at the "Got Health Data" conference stated, "It all comes back to that basic trust, dignity, and respect. If a community, an individual, or family does not feel that they have been treated by their health care system, or that institution, with trust and dignity, their likelihood of being interested in also sharing these very personal

[17] Andre Bächtiger et al., "Deliberative Democracy: An Introduction," in *The Oxford Handbook of Deliberative Democracy*, ed. Andre Bächtiger et al. (Oxford University Press, 2018), 0, https://doi.org/10.1093/oxfordhb/9780198747369.013.50.
[18] Kieran C. O'Doherty, Alice K. Hawkins, and Michael M. Burgess, "Involving Citizens in the Ethics of Biobank Research: Informing Institutional Policy through Structured Public Deliberation," *Social Science & Medicine* 75, no. 9 (November 1, 2012): 1604–11, https://doi.org/10.1016/j.socscimed.2012.06.026.
[19] Svenja Wiertz and Joachim Boldt, "Evaluating Models of Consent in Changing Health Research Environments," *Medicine, Health Care and Philosophy* 25, no. 2 (June 1, 2022): 269–80, https://doi.org/10.1007/s11019-022-10074-3.
[20] James Scheibner et al., "Data Protection and Ethics Requirements for Multisite Research with Health Data: A Comparative Examination of Legislative Governance Frameworks and the Role of Data Protection Technologies†," *Journal of Law and the Biosciences* 7, no. 1 (July 25, 2020): lsaa010, https://doi.org/10.1093/jlb/lsaa010.
[21] Amy Peterson et al., "The Health Equity Framework: A Science- and Justice-Based Model for Public Health Researchers and Practitioners," *Health Promotion Practice* 22, no. 6 (November 1, 2021): 741–46, https://doi.org/10.1177/1524839920950730.
[22] Malvika Pillai et al., "Toward Community-Based Natural Language Processing (CBNLP): Cocreating With Communities," *Journal of Medical Internet Research* 25, no. 1 (August 4, 2023): e48498, https://doi.org/10.2196/48498.
[23] Simona C Kwon et al., "Applying a Community-Based Participatory Research Framework to Patient and Family Engagement in the Development of Patient-Centered Outcomes Research and Practice," *Translational Behavioral Medicine* 8, no. 5 (September 8, 2018): 683–91, https://doi.org/10.1093/tbm/ibx026.
[24] Lisa M. Vaughn and Farrah Jacquez, "Participatory Research Methods – Choice Points in the Research Process," *Journal of Participatory Research Methods* 1, no. 1 (July 21, 2020), https://doi.org/10.35844/001c.13244.
[25] "Patient Innovation," accessed November 15, 2023, https://patient-innovation.com/.
[26] Mitchell L. Gordon et al., "Jury Learning: Integrating Dissenting Voices into Machine Learning Models," in *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI '22 (New York, NY, USA: Association for Computing Machinery, 2022), 1–19, https://doi.org/10.1145/3491102.3502004.
[27] Liedeke Koops and Richard I. Lindley, "Thrombolysis for Acute Ischaemic Stroke: Consumer Involvement in Design of New Randomised Controlled Trial," *BMJ* 325, no. 7361 (August 24, 2002): 415, https://doi.org/10.1136/bmj.325.7361.415.

and also politicized data points, this may be a non-starter unless we can assure them of the trust that we can maintain within the system. And certainly, that can be done through community partnerships."[28]

Patients themselves recognize the importance of bidirectional community engagement around data use, including input from groups of diverse individuals who are representative of the populace. For example, below are several quotes from patients interviewed as part of a 2022 UCSF study conducted with clinic staff and patients at four UCSF clinical locations:

> *I think it'd be a great idea to have patients on the panel and community leaders, people who are really aware of the possible harms… (participant #3)*
>
> *I think everybody who goes to UCSF already knows that there's some type of data share that's happening, but I think approaching people the way that you did in lobbies, really gives you guys a better opportunity to work with your patients and actually have your patients share in... You probably get more willingly shared information just by doing that. But I understand that's an awkward thing. (participant #6)*
>
> *"Let's just say, for instance, you have a focus group of like 10 or 15 people. All the [inaudible 00:25:46] place. So, a mixture, a conglomerate, of, we have a healthcare worker who's Asian, we have a community member that's African American, they have a social worker that is Hispanic, that [00:26:00] type of thing, and so mix it up so you can get the point of view from everybody." (participant #7)*

Given the above, the Task Force urges UCH to consider the following high-level principles for community engagement:

- Attention to diversity to ensure appropriate representation of the populace, including those affected by decisions made by UC.
- Making involvement accessible and affordable, e.g., through transportation assistance, reimbursement of expenses, language concordance, remuneration of expertise, and respectful use of time.
- Transparency and expectation management about how community input will be used, e.g., advisory vs. decision-making.
- Preparation of community members with adequate background information in clear and understandable language to enable full participation.

Education

Both transparency and community engagement require education, a third pillar of a justice-based model**.** Without a public educated on how Health Data may be collected, used, and shared, it is difficult to achieve full, diverse, and representative community engagement. General understanding of the role of data science in people's lives, including in health care, is very uneven among the public. UCH cannot have transparent and open communications about how (*i.e.*, current practices, regulations, and

[28] Julie Harris-Wai. "Incorporating the Patient Voice into Data Use." Panel Discussion presented at the Got Health Data? Moving Toward a Justice-Based Model of Data Use Conference, April 19, 2022. https://www.youtube.com/watch?v=z15O5eM658M&t=3836s.

protections) and why Health Data are used in research if the populace is unfamiliar with the basic issues. The University must inform patients, the non-patient populace, and UC staff and faculty about these issues so that bidirectional communication can be unitized effectively. Public education by UC about the role of data science in the lives of Californians would be a public good from which the entire state would benefit.

Improving health literacy around data use requires culturally responsive, accessible, and understandable content. Non-written formats, such as short videos, have been used very successfully for informed consent at some UC medical campuses; video formats would lend themselves well to public educational campaigns and should be considered, among other communication strategies. Finally, any health communications must be in multiple languages (and closed captioning considered) that reflect the diversity of California's population.

Educational efforts that lead to meaningful transparency should include definitions of fundamental terms that healthcare professionals may consider commonplace, such as "health data," "research," and "de-identification." Patients rightly have concerns about privacy risks, and these should be addressed by providing information in plain language on current UCH practices around how and why Health Data are collected, stored, shared, and used. This material should include realistic assurances about data security and privacy protections, and, if possible, should be developed with input from members of the intended audiences.

Transparency, education, and community involvement require that UCH develop processes to learn about patients' hopes and concerns; respond to patients' hopes/concerns by modifying UCH processes and/or addressing these in educational materials, and publicly update how patients' hopes/concerns have been addressed. Additionally, UCH should develop ways to share findings derived from Health Data, to help the populace better understand how these data help UCH improve public health. UCH will not be able to resolve all hopes/concerns in a manner acceptable to everyone. By acknowledging that these hopes/concerns have been heard and have been used to inform UCH processes, however, UC can demonstrate to its patients/populace that it is working for the public good and that the University has earned the public's trust.

Recommendations to Develop a Justice-based Model of Health Data Use

Based on the considerations described above, the Task Force recommends that UC engage in the following next steps to implement a justice-based model of Health Data use:

1) Communications

   Together with stakeholders across the UC system communications, leaders should develop a strategic plan for ongoing approaches to transparency in how UC uses patient data, including:

   a) how information is presented

   b) how patient input is incorporated into decisions around large-scale data analytics

   c) how UC engages in the national discourse on this issue

2) Education

   a) UC should work with health literacy experts to develop materials providing a brief education on Health Data, including what it is, what purposes it can be used for, why it is valuable, and what steps UCH patients and Californians can take to learn more.

      i. One suggestion raised by experts was that this education take the form of one or more brief videos that can be shared on UCOP websites, as well as UC campus websites.

b) UC should consider including education on use of Health Data within appropriate courses in the undergraduate/graduate curriculum and should consider sharing teaching materials with Cal State and California Community Colleges to encourage them to develop their own courses.

3) Governance

a) UC Health should incorporate into governance structures input from experts in bioethics, data science, and artificial intelligence[29] both within the UC system and external to UC.

b) UC Health should establish a council that will convene periodically and advise the system-level Health Data Office and local governance offices on data use and sharing and public communication. The council should include representatives from all communities with UC medical centers and have defined terms with opportunities for rotating membership. Ideally representatives should have a working knowledge of how Health Data may be used for quality improvement, research, and public health efforts, which could be provided through the educational efforts described above, if needed. Attention to diversity to ensure appropriate representation of the UCH populace is also paramount and should include historically disadvantaged populations, limited English speakers, and others.

c) UCH should review and assess governance structures regularly, but no fewer than every 3 years, and update structures accordingly.

d) Further research that includes patient focus groups should be conducted at individual campuses to assess the impact of these recommendations, especially with respect to diverse populations.

[29] As the Task Force work has progressed, interest in development and deployment of artificial intelligence (AI) tools in healthcare has increased substantially, and AI governance has become an important focus of UC. In October 2021, a UC Presidential Task Force released the first set of AI responsible principles in higher education in a report entitled Responsible Artificial Intelligence: Recommendations to Guide the University of California's Artificial Intelligence Strategy. UC has continued to work on operationalizing these principles in the health domain and more broadly. Since AI technologies rely upon data, the Task Force expects that AI governance will be integrated with ongoing UC data governance efforts.

Part 2 – Operationalizing Responsible Health Data Governance at UCH

In addition to developing the key elements of a justice-based model for Health Data use, the Task Force also developed responsible approaches UCH and UCH locations should consider when operationalizing the UC Health Data governance principles. These include tracking data sharing agreements with third parties, and local and systemwide processes to ensure responsible data sharing with third parties.

A) Tracking System for Data Sharing Agreements

Contracts and contractual provisions governing data sharing are fundamental to responsible data governance. Recognizing the importance of tracking the existence of data sharing agreements with third parties throughout a system as large as UC Health, the Ad Hoc Task Force Report recommended that the Task Force form a team of experts within UC to propose recommendations regarding a systemwide Health Data Set Access Repository (HDAR). UC staff could use the HDAR to become aware of data sets that include data from more than one UCH location, what entities have been allowed access to these data sets, and what contractual terms have been put in place to govern data sharing. The Ad Hoc Task Force Report also recommended that the team of experts assess whether similar local repositories should be formed. Additionally, it recognized that setting up a systemwide or local repository requires balancing the benefits arising out of the capture of necessary information and reducing administrative burden.

Based on this charge, the Task Force assessed the need, utility, and recommended features of an HDAR and set forth recommendations around how to proceed. As part of this evaluation, the Task Force reviewed the needs from the perspective of UC Health locations, the needs for systemwide reporting, the ability to collect and manage Health Data on an ongoing basis, and potential tools for implementation external and internal to UC. This review provided for a better understand of the benefits and challenges of creating a systemwide repository. The Task Force also solicited input from local campus governance groups, contracting groups, and technology transfer groups regarding the appropriate audience for this tool, what elements of data would be most useful to collect, how to balance burden and utility, and legal and compliance concerns.

Assessing the Need for a Data Set Access Repository

Among the tools the Task Force discussed was the website www clinicaltrials.gov, a publicly accessible registry for clinical trials maintained by the National Library of Medicine at the National Institutes of Health (NIH). Clinical trials conducted both nationally and internationally are required to be entered into this registry under federal law, and registration of trials may also be required by journals prior to publication. Task Force members appreciated the relatively easy interface; however, members also noted that queries to the site can be clunky and do not necessarily return all records. Moreover, there is no requirement for entry and no consequences for those who do not enter trials into the registry, so the registry is not comprehensive.

The Task Force also discussed an existing systemwide tool, the Faculty & Organization Profile System. This System provides two services.[30] First, the Faculty Profile System displays funding awards, technology transfer agreements and invention disclosures by UC principal investigators. It can aid a technology transfer office when assessing the researcher's portfolio for conflicting obligations. Second, the Organization Profile System shows UC research and technology transfer relationships with an institution or company. Since the system is UC-wide, it is a helpful tool if a campus receives an

[30] This System pulls data from the Contract and Grant (SPX) database and/or the Patent Tracking System (PTS). Both SPX and PTS are maintained by UC Office of the President units. SPX is maintained by Institutional Research and Academic Planning (IRAP), and PTS is maintained by Innovation Transfer and Entrepreneurship (ITE).]

agreement from a new sponsor and wants to reach out to another campus to compare terms (*e.g.*, Riverside gets a new contract from Janssen, has never worked Janssen before, and sees UCSF has many Janssen agreements, so Riverside connects with UCSF to review UCSF's agreements). Work Group members agreed on the utility of having this information in a centralized hub, but also recognized that it is not a central repository of agreements and does not track data provisions. Task Force members discussed the pros and cons of having agreements exist in a central repository. For example, campuses may not want their agreements open for review by others because of concerns about confidentiality and security, or because of competitive concerns. At the same time, given staff turnover and burden it can be challenging to locate relevant agreements across the system. Task Force members also recognized that any systemwide tool needs to be maintained and updated and socialized across the UC enterprise on an ongoing basis to account for staffing changes, among other things.

After reviewing these resources and other expert input the Task Force recommends a pilot of an HDAR either managed by or in coordination with the UCH Health Data Office, If possible, this tool should leverage existing systemwide tools.[31]

Features of a Data Set Access Repository
The Task Force recommended that the HDAR include the following features:

- Collection of certain data elements including project name and short description; locations impacted; outside UC collaborator or sponsor; description of data elements accessed/shared; project status; point of contact; flags for whether approvals are needed and status of needed approvals (e.g., IRB, governance); data management provisions.
- Enable the UCH Health Data Office and others at UC to be a resource to help get information that locations are looking for and refer them to the right local contact.
- Provide an inventory of data agreements, contracts, and templates.
- Ability for traceability and auditability, to enable responses to records requests and compliance concerns.
- Controls to address privacy and security concerns, including ensuring that access is limited to approved UC staff and that strong cybersecurity safeguards and policies are in place.

Recommendations
Based on the considerations described above, the Task Force recommends that UC engage in the following next steps regarding establishment of a Health Data tracking system:

1) Develop a pilot for an internal facing HDAR, managed by the UCH Health Data Office. This pilot system should:
    a. Capture the elements described above.
    b. Utilize or build upon an existing tool, if possible.
    c. After one year, the UCH Health Data Office will assess success of the pilot, including adoption across locations, and feedback from users; make a determination of whether or not to continue; and, if appropriate, refine the HDAR to incorporate feedback.
2) Work with experts systemwide, including from legal, compliance, privacy, security, IT and risk services to assess the feasibility of developing a pilot of an external facing tool to increase transparency around data agreements. If a decision is made to move forward, the pilot would likely start development after six months and be released roughly one year after the internal HDAR pilot is launched.

[31] The Task Force also made other recommendations implicating contracting, including contracting guidelines set forth in Appendix 1, Exhibit B.

3) Grow systemwide data contracting expertise. The UCH Health Data Office should:
   Work with experts systemwide, including research administration, contracting, compliance, legal, and others to develop a community of practice and systemwide resources regarding executing data sharing agreements, including guidance, templates, and education around utilization of the HDAR.
4) The UCH Health Data Office should host an annual system-wide data conference with other collaborators across the UC system as conference topics suggest (*e.g.*, BRAID, the UC Real World Evidence Collaborative).

B) Data Sharing Guidelines

The Ad Hoc Task Force Report proposed Interim Operating Guidelines (IOG) for assessing data-intensive collaborations with third parties to address existing third-party data access requests. As part of these interim processes, the President appointed a small "Tiger Team" with legal, regulatory compliance, IT, and communications expertise to evaluate projects regarding access to Health Data by third parties, and also recommended the development of local data governance groups at each health campus. These interim guidelines were purposefully intended to be interim, a first step in an iterative process.

Since the release of the IOG, there have been major advancements in cloud computing, artificial intelligence, telehealth, and consumer-facing mobile health applications. Out of these developments has emerged an increasingly complex landscape of Health Data sharing partnerships, making the need for consistent, transparent processes for considering these collaborations essential.

The Task Force conducted an in-depth revision of the IOG based on feedback from local governance groups, the Tiger Team's experience over the past three years, and the evolving landscape of digital health. The updated DSG is attached as Appendix 1. In drafting the DSG, the Task Force made several material updates to improve local and systemwide processes. The Work Group recommends that the DSG be adopted systemwide, along with sufficient support to allow the Health Data Oversight Committees (HDOCs) to perform their functions in a thorough, consistent, and timely manner. The major updates contained in the DSG are:

- **Updated UC Responsible Principles for Health Data Governance ("Principles").** The Ad Hoc Task Force Report set forth six high-level principles. The Work Group revised the Principles to clarify existing language and account for similar concepts that are shared across the Principles.

  Among other things, the updated Principles recognize explicitly the importance of UC partnering with all sectors, including non-profit, for-profit, and government entities to create new knowledge and benefit the public. In so doing, the University must articulate a clear public benefit in any arrangement in which clinical data are shared outside of the University, especially for commercial transactions.

- **Updated Scope.** The Ad Hoc Task Force Report applied its governance framework to Health Data, which included two categories: Source Health Data, which is the primary and original data set, and Derived Health Data, which represents some transformation of the original Source Health Data. In addition, the Report set forth a carve-out to the definition of "Health Data" for "research data created exclusively for a sponsored clinical research study pursuant to Institutional Review Board (IRB) approval and collected pursuant to a patient authorization or consent." In response to feedback, the Task Force made several material revisions to this definition:

  - Revised the definition of Derived Health Data. The DSG defines Derived Health Data as any derivation of source Health Data irrespective of how trivial or complex the derivations may be. Derived Health Data may be produced by:
    - Mathematical calculations performed on Source Health Data OR
    - Deletion or substitution of Source Health Data components OR
    - Creation of synthetic versions of Source Health Data

- Updated the clinical research exception to the Health Data definition to focus on whether informed consent has occurred, rather than whether the data has been created exclusively for a sponsored clinical research study. The DSG defines Clinical Research Data (which falls outside the scope of the framework) as any data created exclusively for a clinical research study or clinical trial in which informed consent was obtained and specifies the non-UC third parties with whom the data can be shared. Clinical Research Data are:
    - Limited in scope as defined in an Institutional Review Board (IRB)-approved study; and
    - Created or collected during a study pursuant to a patient consent that is approved by the IRB and explicitly describes intent to share patient data with one or more named non-UC third parties.

- **Guidance for Local Health Data Oversight Committees (HDOC).** Over the past three years, five local health campuses (UCSF, UCD, UCI, UCLA, UCSD) have set up local processes, including Health Data oversight committees, to review third-party Health Data access requests. These HDOCs have now reached a certain level of process maturity. The DSG is intended to provide expanded guidance regarding the structure and functions of the HDOCs, informed by the collective experience of each campus's committees over the years since their formation. The guidance is expected to be interpreted and adapted as appropriate to local circumstances while encouraging consistency across sites wherever possible.
- **Revised Contracting Guidance**. In revising the contracting guidance contained in the IOG, the Task Force balanced the need for flexibility given the significant variety and volume of collaborations that come through local contracting offices with the need to ensure that contract negotiations with respect to Health Data always consider the Responsible Principles for Health Data Governance. The DSG contains revised guidance with respect to Health Data transactions that is intended not to duplicate, supplant, conflict with, or unnecessarily delay other requirements (*e.g.*, IRB review), but to add an additional oversight specifically for transactions involving data that may not otherwise receive a full review.
- **Updated Role of Systemwide HDOC.** The Task Force updated the role of the UC Health Systemwide Health Data Oversight Committee ("Systemwide HDOC," formerly the "Tiger Team") to have two primary functions.
    - First, the Systemwide HDOC, convened by the UCH Health Data Office within CDI2, reviews certain high-risk requests to share data with non-UC third parties that have been escalated from one or more UC Health campuses, as well as systemwide data sharing requests. The DSG sets forth the process for escalation to the Systemwide HDOC. The Systemwide HDOC should include representation from the local HDOC of each health campus, divisions including UC Legal, UCOP compliance, UCOP Research Policy Analysis and Coordination (RPAC), and experts in communications, bioethics, and cybersecurity. In addition, the Systemwide HDOC may bring in expertise on an ad hoc basis from other technical, business, or research units across UC. The Systemwide HDOC should also regularly incorporate input from the newly created Patient Council, described in Part A, above, so that the patient voice can be meaningfully integrated into decisions around data use.
    - Second, the Systemwide HDOC works with relevant stakeholders from local governance groups and UCOP offices to develop and disseminate guidance around data sharing issues that raise common concerns across locations.

Conclusion

Data are central to UC's mission to improve the health and well-being of all people living in California, the nation, and the world. The recommendations set forth in this report reflect UCH leadership's continued focus on balancing the importance of utilizing data at scale while also acting as a steward of this resource. As the health data sharing landscape continues to evolve, responsible data governance will require ongoing, cross-functional efforts at campuses, and coordination of these efforts systemwide. The collective adoption of measures to promote the safe and responsible sharing of Health Data will enable UC to facilitate scientific collaboration, discover new treatments, and deliver exceptional care.

Appendix 1 – UC Health Data Sharing Guidelines

UC Health Data Sharing Guidelines

Version 1.0

February 2024

Introduction

As a leading healthcare-delivery and research institution, the University of California (UC) has developed extensive Health Data assets across its campuses. Insights gleaned from the analysis of these data assets will drive advancements in healthcare delivery, discovery, and education that will in turn improve health outcomes, patient and provider experience, value, and health equity, both for UC patients and the general public.

Enabling data-driven improvement of human health requires both scientific research and translation of knowledge into new innovations. Extracting key insights from large, complex Health Data sets has become an invaluable but challenging part of the discovery process that requires a comprehensive approach, including scientists and technologists spanning multiple disciplines both within and outside of UC, and collaboration with both government, nonprofit, and for-profit industry partners. Data sharing has thus become an indispensable mechanism by which the UC works to fulfill its mission to improve public health.

Despite potentially transformative benefits, sharing Health Data comes with risks. Maintenance of patient privacy is necessary to protect patients and avoid legal, ethical, reputational, and financial ramifications to the University. Even when compliant with laws and UC policies, data sharing may still carry risks related to health equity, protection of vulnerable populations, effects on our reputation, and breach of the public trust. In other words, even when we *can* share Health Data, it's not always clear that we *should* share Health Data.

UC policies will not always keep pace with the rapidly evolving ways in which data are captured and used, including advances in healthcare-specific and consumer-facing software and devices. Therefore, the University needs a consistent, streamlined, transparent, and replicable framework that protects patient privacy, ensures data security, promotes responsible partnerships with third parties, and effectively balances the risks and benefits of data sharing before engaging in such efforts.

Background on Data Sharing Oversight

Governance of UC Health Data has been a focus of the University for several years. In 2018, a Presidential Ad Hoc Task Force on Health Data Governance released a report that began to address this complex terrain. The Ad Hoc Task Force Report recommended, among other things, the formation of an interim system-wide mechanism to evaluate requests to share Health Data with non-UC third parties, and creation of an analogous health data oversight process at each UC Health campus. The Ad Hoc Task Force Report recognized the potentially transformative benefits of sharing UC Health Data with non-UC third parties, but also identified multiple challenges, including but not limited to:

- Defining the scope of Health Data.
- Developing a mechanism for determining how UC Health should balance the competing interests in access to its Health Data.

- Determining how the University should identify and monitor such potential competing interests among its various systemwide initiatives and across campuses, particularly multiple transactions involving the same industry partner and, potentially, the same Health Data.
- Determining how the University should address the ethical, regulatory, compliance, and reputational implications of sharing Health Data for use by or in partnership with third parties.
- Identifying whether there are specific activities concerning the use and disclosure of Health Data in which UC clearly should not engage.

The President subsequently appointed a small "Tiger Team" with legal, regulatory compliance, IT, and communications expertise to establish an interim process and to review projects escalated from the UC Health campuses. In addition, each UC Health location has created a Health Data Oversight Committee (variously named, but herein referred to as HDOC) to govern the sharing of Health Data with non-UC third parties.

The resulting Interim Operating Guidelines (IOG) were developed with input from the Office of the General Counsel, Research Policy Analysis and Coordination, as well as leaders from the original Task Force. The IOG included a description of the interim process and guiding principles, contracting guidance, and sample language for agreements under which UC provides access to its Health Data. The IOG was intended to enable the University to advance its mission to improve human health by sharing its Health Data with outside entities while also being accountable for public benefit in every arrangement.

In the years since the release of the Interim Operating Guidelines, the world has witnessed an unprecedented global pandemic from the SARS-CoV-2 virus with profound and long-lasting impacts on human health and society. The pandemic has highlighted the potential of widespread data sharing to drive the rapid development of new diagnostic, therapeutic, and preventive innovations; but it has also illuminated the critical importance of social justice, health equity, transparency, and our fundamental responsibilities to maintain the public trust and protect the most vulnerable among us.

With this in mind, the present Data Sharing Guidelines (DSG) are intended as an update to the IOG with lessons learned from a changed world and from the constantly evolving landscape of Health Data, Data Science, and Health Technology. These Data Sharing Guidelines provide expanded guidance regarding the structure and functions of the HDOCs, and the responsibilities of the "Tiger Team," newly renamed as the Systemwide HDOC, informed by the collective experience of each campus's committees over the years since their formation. The IOG was originally intended as a truly interim document and so too are these updated guidelines. Please note that these guidelines are intended to undergo periodic updates to keep pace with substantive changes in their scientific, regulatory, and societal context. Feedback is encouraged and expected -- and will help shape subsequent versions for the better.

Scope

***The DSG apply to transactions involving the sharing of UC "Health Data" with third parties outside the UC system*, including industry, non-profit, governmental, and academic organizations**. Data sharing includes release of Health Data from the University and providing access to Health Data that may reside in UC-controlled storage environments. Health Data includes both identifiable and deidentified data and is defined by two categories including Source Health Data, which is the primary and original data set, and Derived Health Data, which represents some transformation of the original source data.

Source Health Data refers to any information pertaining to the health, care, and treatment of UC Health patients and plan members that:

1. Results from treatment or monitoring of a patient; or

2. Is contained in a financial claim or bill for services provided; or

3. Is used for clinical operations, risk management, performance improvement, population health activities, or care quality metrics.

Derived Health Data refers to any derivation of source Health Data irrespective of how trivial or complex the derivations may be, though does not include summary statistics derived from patient-level data. Derived Health Data may be produced by:

1. Mathematical calculations performed on Source Health Data; or

2. Deletion or substitution of Source Health Data components; or

3. Creation of synthetic versions of Source Health Data.

There are two additional limits on the scope of the DSG. First, the Task Force has carved out from the definition of "Health Data," and thus the DSG, a category of data termed Clinical Research Data. Clinical Research Data are any data created exclusively for a clinical research study or clinical trial in which informed consent was obtained and specifies the non-UC third parties with whom the data can be shared. Specifically, Clinical Research Data are:

1. Limited in scope as defined in an Institutional Review Board (IRB)-approved study; and

2. Created or collected during a study pursuant to a patient consent that is approved by the IRB and *explicitly describes intent to share patient data with one or more named non-UC third parties.*[32]

While Clinical Research Data fall outside of these guidelines, disclosure of these data is still governed by law, existing UC policy, and the clinical research agreement, and overseen by each campus's applicable research regulatory office(s) such as the Institutional Review Board (IRB) and Sponsored Projects Office. Moreover, the DSG does cover other research uses of Health Data. For example, sharing of Health Data with non-UC third parties, as part of research studies that involves a waiver of authorization or exemption from full IRB review, falls within the DSG's scope.

Second, while the formal governance committee structures required by the DSG and described below are not required of non-UC Health academic health center/campus locations, non-health locations should still establish processes for evaluating Health Data transactions consistent with the principles and risk considerations described below. The Task Force recognizes that procedures may vary across locations and depend on the project at issue. If a project is a joint one between a health location and non-health location, for example, the project might be reviewed by the health location's HDOC. Other high-risk projects might be escalated to the Systemwide HDOC for review.

Adherence to Principles Governing All Health Data Transactions

In considering transactions under which outside entities gain access to UC's Health Data, the University must ensure that certain key principles are adhered to. In its report to the President, the Ad Hoc Task Force on Health Data Governance originally established six high-level principles. Based on experience across the UC Health campuses since their inception, the principles have been revised to the five principles below. All Health Data transactions must have a clearly articulated benefit that is consistent with these Principles and the University's mission.

[32] As local and Systemwide governance processes continue to mature, the definition of Health Data and exception for Clinical Research Data may be updated as needed.

1. Attention to the University's Unique Responsibility and Mission

As a public research and educational institution, the University's fundamental mission is to create and share knowledge broadly, but also responsibly and strategically. UC must safeguard sensitive information while continuing to provide broad use and dissemination of knowledge to benefit the public.

2. Sharing Data Outside UC for Public Benefit

The University must collaborate with others to create new knowledge and benefit the public. This means partnering with all sectors, including non-profit, for-profit and government entities. The University must articulate a clear public benefit in any arrangement in which Health Data are shared outside of the University, especially for commercial transactions. Financial gain and commercial development, by themselves, are not a clear public benefit.

3. Justice

The University recognizes that there are significant inequalities in our society. Developing a shared concept of the public good requires authentic, sustained, and dynamic engagement, including outreach with the public to ensure that disenfranchised communities participate in setting health research, clinical, and policy-making agendas.

4. Transparency and Patient Engagement

The University should empower data subjects and be transparent about University activities involving Health Data. At minimum, each UC Health location is encouraged to develop and publish data governance policies and procedures on a publicly available UC Health location website.

5. Responsible Stewardship

Generation of new scientific knowledge and related innovations in healthcare are increasingly dependent upon the analysis of large Health Data sets, which may require sharing data with both academic, non-academic, and industry collaborators to enable use of the most advanced scientific methods available. To do so responsibly requires careful data management including a rigorous and skeptical assessment of the promised benefits of sharing Health Data, attention to patient privacy and data security, promoting the widespread dissemination of scientific findings, and ensuring that the benefit of UC Health patients and society in general are the primary motivations for any sharing of Health Data.

Key Functions for Local Health Data Oversight Committees

Local HDOCs are intended to mitigate the risks of sharing Health Data with non-UC third parties where existing laws or policies may be unclear or where data sharing is permissible under existing law or policy but potentially violates the five Principles governing data sharing delineated in this document. Key HDOC functions include (1) development and implementation of processes to review certain Health Data sharing agreements and (2) development and dissemination of policies and guidance around data sharing issues. HDOCs should work collaboratively with other local offices in these functions and avoid creating duplicative or redundant processes.

1. Review of Health Data Sharing Agreements.

In a rapidly changing scientific, technological, regulatory, and business landscape, specific requests to share Health Data with non-UC third parties may not fall clearly within existing laws or policies. When such cases are escalated or otherwise brought to attention, HDOCs may be expected to perform an interpretive "can we?" function to determine if proposed terms and conditions are permissible. HDOCs are also expected to perform a critical "should we?" function in cases where data sharing is deemed legally permissible but where unacceptable risk may still exist regarding public trust, conflicts of interest, or other ethical concerns.

To achieve these objectives, HDOCs should develop tools aimed at systematically identifying and assessing risks that prompt escalation of data sharing requests to the HDOCs for review. Examples of such risk mitigation tools include self-assessment check lists for those requesting to share data and "red flag" escalation check lists for IRBs and local contracting offices to use when reviewing agreements. Local contracting offices include Industry Sponsored Research/Contracts and Grants, Procurement, Clinical Trials, Technology Transfer, and Information Technology. Examples of specific data sharing red flags include but are not limited to:

- Apparent sale or barter of Health Data in exchange for goods, services, or other benefits
- Redisclosure or reuse of Health Data or deidentified data with additional third parties without restriction (e.g., using Derived Data for secondary purposes, including training of artificial intelligence models)
- Requests for Health Data beyond the minimum necessary required to provide specified services or answer specific scientific hypotheses.
- Health Data sharing requests involving especially large volumes of data, dealing primarily with highly sensitive data (e.g., reproductive health data, prisoner data, minors) or sharing with recipients with a known history of questionable data stewardship.
- Sharing of potentially identifiable biometric data such as genomic data, retinal images, or detailed head/face imaging.
- Sizable conflicts of interest related to Health Data exchange.
- Unusual or unique terms and conditions of data sharing unfamiliar to the principal contracting office.
- Terms and conditions that specify exclusive data sharing rights with a single third party. Third party has had a privacy or security data breach or incident associated with a regulatory enforcement fine or public media scrutiny.

If a collaboration raises one or more of these flags, local contracting offices should work with relevant stakeholders to clarify and mitigate identified risks as needed. If residual "can we" or "should we" uncertainties persist, the agreement should be escalated for review by the local HDOC to identify additional risk mitigation measures, including contractual protections and technical solutions, and a decision on whether the agreement should move forward. If the local HDOC is unable to arrive at a consensus decision, a final decision should be made by the site's Chancellor or Vice Chancellor for Health, potentially after additional evaluation by the systemwide HDOC (see below for additional details). Exhibit A, based on current governance risk assessment tools developed by local HDOCs depicts a nonexclusive list of criteria that is meant to assist in the development of self-assessment check lists and escalation check lists at each UCH member campus.

In addition to reviewing data sharing requests that have been escalated pursuant to a risk assessment process, local HDOCs are expected to review all Health Data sharing agreements with non-UC third parties (1) involving a non-research partnership; (2)involving the sharing of Health Data for purposes beyond those required for a HIPAA Business Associate (BA) to provide a specified service to UC and meet their legal obligations as a BA; and (3) outside of a covered entity to covered entity agreement or relationship (e.g., treating the same patient). The Task Force recognizes that there may be very few agreements that satisfy all three of these criteria, but those that do are likely to present unusual circumstances that warrant HDOC review.

Wherever possible, HDOCs should not duplicate the work of the IRB or local contracting offices that have authority to sign other agreements related to Health Data sharing but instead should assess such

agreements when escalated to the HDOC from these local offices.[33] In this capacity, HDOCs are expected to approve, deny, or further escalate requests of concern. Evaluation of cases should occur through well-defined processes to ensure transparency, consistency, fairness, and efficiency. Potential components of HDOC processes include use of intake forms, scoring rubrics to evaluate potential risks and benefits, standardized templates to present issues to the HDOC, and repeatable escalation workflows to guide the resolution of disagreements. HDOC determinations should be summarized and recorded in committee proceedings and communicated clearly to key stakeholders. Local HDOCs should also analyze approved collaborations to understand and assess the impact on underserved and sensitive patient populations.

2. Development and Dissemination of Policies and Guidance

In addition to developing internal processes around escalation and review of data sharing agreements, HDOCs should work collaboratively with local offices including the IRB, local contracting offices, legal, and compliance to develop relevant policies, standard operating procedures, and tools to ensure that agreements involving data sharing with non-UC third parties stay current with industry best practices and emerging risks. HDOCs should ensure that new or existing data sharing policies are consistent with applicable federal and state law, University of California policies, and consistent with the Principles outlined in this document. HDOCs should provide consultative services to local offices where needed to ensure that practices stay up to date with new technologies, scientific knowledge, and business practices.

HDOCs should also work closely with their respective Information Technology departments to optimize data privacy and cybersecurity processes, and data management practices throughout the data lifecycle, these processes should include encouraging the development of HIPAA-compliant secure computing enclaves that enable collaborative data analysis without data egress from UC-managed environments.

Local HDOC Membership and Reporting Structure

HDOC membership should represent stakeholders charged with protecting the privacy and security of Health Data, those with an understanding of the ethical and strategic implications of Health Data sharing, and those with experience negotiating agreements involving Health Data sharing. As academic health systems, HDOC membership should represent key organizational missions including clinical care, healthcare operations, research and education, strategy and growth, and service to the community.

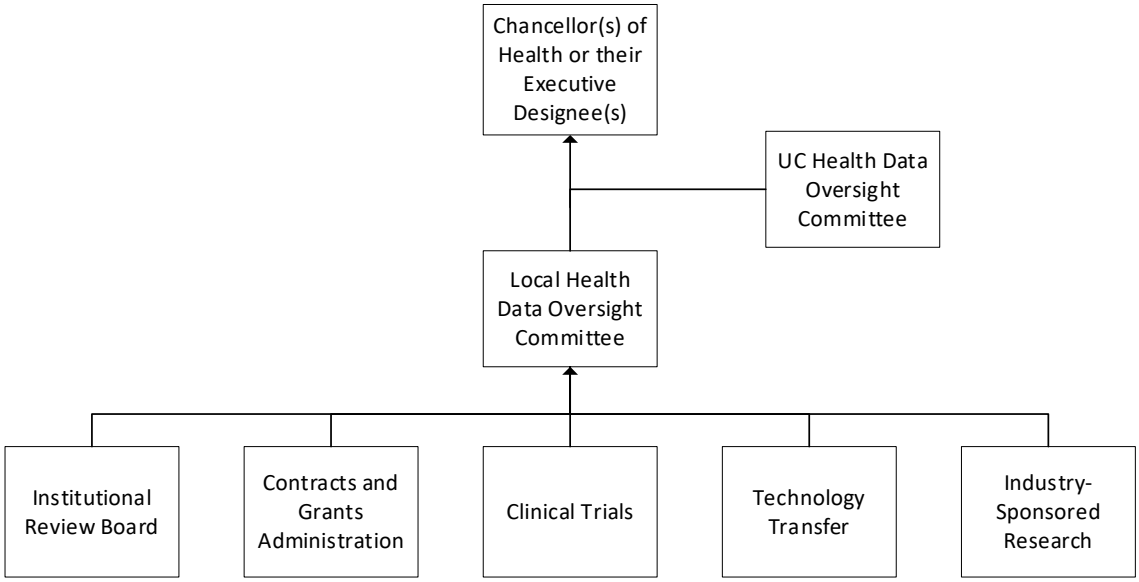
HDOCs should seek membership from local site officers with the authority to sign agreements such as procurement, contracting, sponsored programs, legal, and technology transfer. Additional membership from privacy, security, compliance, ethics, licensing, and IRB offices should be included, as should those with strategic interests in Health Data sharing including professional schools (e.g., School of Medicine, Public Health), Centers and Institutes and business development units, as appropriate. Given the “should we” oversight function, HDOCs should seek to incorporate representation from patients and community members, bioethicists, and experts in diversity, equity, and inclusion where feasible and appropriate through either ad hoc or consistent membership on the committees.

Given the risks associated with Health Data sharing, each local HDOC should report directly to its Chancellor or Vice Chancellor for Health, who should have ultimate decision-making authority at the local level, and to the Systemwide HDOC in some circumstances (Figure 1). Escalations of data sharing requests to local Chancellors/Vice Chancellors for Health should occur when local HDOCs are unable to achieve consensus, a requestor disagrees with an HDOC decision, or a data sharing request is

[33] For example, a local HDOC cannot and should not mandate IRB review; however, there may be situations in which a local HDOC learns of material information which may not have been available to the IRB when it conducted its review. In this situation, the DSG contemplates a recommendation from the local HDOC that the IRB consider further assessment based on this additional information.

determined to pose unusually high risk to the institution and/or to UC Health (such as described above). HDOCs should also report unusually high-risk requests to the Systemwide HDOC and when data sharing involves data from more than one UC Health campus. Finally, local HDOCs should consider consultation with the Systemwide HDOC in cases where data sharing involves circumstances likely to affect other UC Health campuses comparably, when particularly novel circumstances arise, or when denials of local data sharing requests may inform decision making at other campuses.

Figure 1. *Expected Escalation Structure for Local Health Data Oversight Committees*



Data sharing with non-UC third parties poses substantial risks to each campus and to the University of California. To ensure proper risk mitigation, HDOCs should be supported sufficiently to perform their functions in a thorough, consistent, and timely manner. Each campus should assess the need for HDOC support functions including analyst, project management, and administrative roles, with funding allocated by the Chancellor/Vice Chancellor for Health proportionate to need. It is also strongly recommended that each campus designate a responsible executive to implement and oversee its local review process and to serve as a point of contact for communications with the Systemwide HDOC.

UC Systemwide Health Data Oversight Committee

The UC Health Systemwide Data Oversight Committee ("Systemwide HDOC," formerly the "Tiger Team") is responsible for two primary functions. First, the Systemwide HDOC reviews high-risk requests to share data with non-UC third parties that have been escalated from one or more UC Health campuses, as well as data sharing requests originating from systemwide initiatives. Second, the Systemwide HDOC works with relevant stakeholders from campus governance groups and UCOP offices to develop and disseminate guidance related to Health Data sharing issues that may affect multiple UC locations.

Systemwide HDOC Review

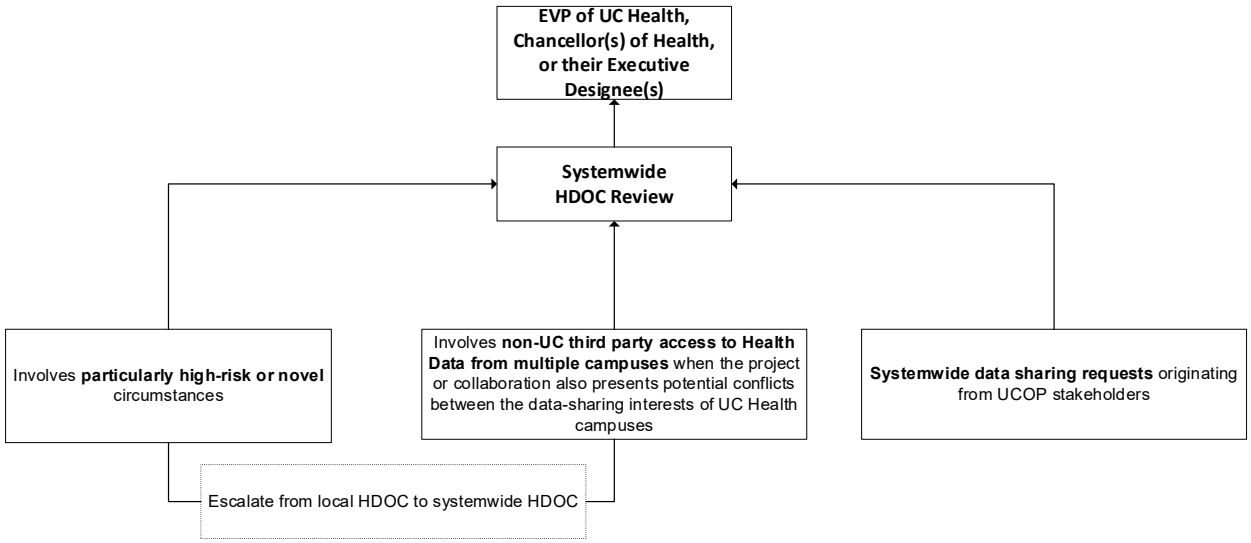
Local HDOCs should escalate the following high-risk requests they have received to the Systemwide HDOC:

- Projects involving particularly high-risk or novel circumstances; and
- Projects involving non-UC third party access to Health Data from multiple campuses when the project or collaboration also presents potential conflicts between the data-sharing interests of UC Health campuses.

The Systemwide HDOC must also assess systemwide data sharing requests originating from UCOP stakeholders, including University of California Health and UC BRAID initiatives. As described above, the Systemwide HDOC will also play a consultative role to local HDOC offices in cases where data sharing involves circumstances likely to affect other UC Health campuses comparably, when novel circumstances arise, or when denials of local data sharing requests may inform decision making at other campuses.

For projects submitted under the escalation criteria above, the Systemwide HDOC will perform "can we" and "should we" evaluative functions comparable to the local HDOCs as described above. The Systemwide HDOC will work with the campus(es) affected to also assess for internal conflicts or competing priorities, and to weigh the risks versus the overall benefits to the University and the patients and public we serve. The Systemwide HDOC will submit recommendations to the relevant Chancellors or Vice Chancellors for Health, as well as the Executive Vice President of UC Health, for their final decision (Figure 2).

Figure 2. *Sample escalation from local HDOC to systemwide HDOC*



Systemwide Guidance and Coordination

The Systemwide HDOC also serves as a high-level expert advisory group that works to develop consistent practices and data ethos across UC with respect to Health Data sharing for large scale data analytics projects. In this role, the Systemwide HDOC helps to develop consensus around underlying UC principles and identify best practices across the campuses for cross-pollination, especially when novel and community impacting questions arise with the evolution of technology, science, and the regulatory and public policy landscape. The Systemwide HDOC will also work to lead, develop and disseminate specific guidance related to innovation issues and strategic questions common across locations, including contracting guidance and sample language or data agreement templates vetted by relevant stakeholders. As with local efforts, these unifying activities should seek to be consistent, transparent, efficient, and non-redundant of other campus efforts.

The Systemwide HDOC will periodically review its data sharing recommendations to determine if they were made consistently and in alignment with the systemwide data sharing principles. Such reviews will also evaluate local HDOC processes and practices to ensure Systemwide HDOC recommendations are advisable from both top-down and bottom-up perspectives. Through these analyses the Systemwide

HDOC will identify current best practices, refine policies, and illuminate the positive and negative outcomes of projects the University has undertaken to promote accountability.

Convening and Composition of Systemwide HDOC

The UCH Data Governance Office, housed in the Center for Data-driven Insights and Innovation (CDI2) convenes the Systemwide HDOC. The Systemwide HDOC should include representation from the local HDOC of each health campus, UC Legal, UCOP Compliance, UCOP Research Policy Analysis and Coordination (RPAC), and experts in communications, bioethics, and cybersecurity. In addition, the Systemwide HDOC may include ad hoc expertise from other technical, business, or research units across UC. The Systemwide HDOC should also regularly incorporate input from patients and community members so that the patient voice can be meaningfully integrated into decisions around data use.

As with the local HDOCs, the Systemwide HDOC and UCH Data Governance Office should be supported sufficiently to perform their functions in a thorough, consistent, and timely manner with funding proportionate to need.

For questions or comments about these Data Sharing Guidelines or the 2024 Health Data Governance Report, please email **HealthData@UCOP.edu.**

EXHIBIT A - UC HEALTH LOCAL CAMPUS SAMPLE RISK CRITERIA

LOCAL CAMPUSES MAY LEVERAGE THIS LIST AS A STARTING POINT IN DEVELOPING DATA SHARING RISK ASSESSMENT TOOLS

Inappropriate Purposes

- COMMERCIAL INTENT to use UC data to exclusively develop or improve third party's products or services which neither provide a benefit to UC patients nor contribute to research for the public benefit.
- SALE, DISCOUNT or BARTER OF HEALTH DATA in exchange for direct payment, goods, services, UC benefit in-kind, volume discount, or item of value to UC.

Contractual Risks

- POLICY VIOLATION where third-party terms and conditions are in violation of UC local policies, including UC policies around data privacy, security, ethics, patient confidentiality, intellectual property, indemnification, or conflicts of interest.
- DILUTION OF DATA TERMS where a third party is unwilling to accept UC data, indemnification, or insurance risk coverage terms as-is or with reasonable modification, such as those included in the UC BAA (Business Associate Agreement) and Appendix DS (Data Security) templates.
- UNUSUAL DATA TERMS giving exclusive data rights, unique terms, or unusual conditions not typically granted by UC.

Reputational Harms

- SIZEABLE CONFLICTS OF INTEREST where a health data sharing arrangement could be perceived by the public as inappropriate since UC individuals involved have significant interest in the third-party (e.g., stock equity, partial ownership, or consulting income with third-party entity)
- REPUTATIONAL CONCERNS arising from recent third party's public media scrutiny, regulatory enforcement, privacy/security breach, or any other questionable data stewardship practices.

Heightened Data Handling Responsibilities

- HIGHLY SENSITIVE DATA such as genetics, biometrics, reproductive family planning, chemical dependency, mental health, sex life, uniquely identifiable body images, or data involving prisoners or minors.
- QUESTIONABLE DEIDENTIFICATION PRACTICES allowing the third party (rather than UC) to handle deidentification without an independent HIPAA expert statistician or failing to treat de-identified data as UC proprietary confidential information.
- UNNECESSARILY BROAD ACCESS involving a third-party request for UC Health data elements beyond the minimum necessary to provide specified services or answer specific scientific hypotheses for a large volume of UC patient records.

EXHIBIT B – HDOC Contracting Guidelines and Review Procedures

Contracting plays a central role in enabling responsible Health Data transactions, and this section provides guidance for various contracting offices that may handle these transactions. It is important to recognize that there exist numerous laws, policies, delegations of authority, and oversight units (both central and local) involved in these transactions. The intent of these guidelines is not to duplicate, supplant, or unnecessarily delay other requirements, but to add oversight specifically for transactions involving data that may not otherwise receive a comprehensive review. The exemption for Clinical Research Data, for example, is intended to avoid duplication of the IRB's full review of patient data access in certain projects, but leaves open projects or even components of projects, that are exempt from complete IRB review. Additionally, these guidelines are intended to ensure that contract negotiations with respect to Health Data always consider the Principles identified in this policy as they apply to the contracting process:

1. University Mission: University agreements should always further the University's fundamental mission in serving the public while safeguarding the data of its patients and research subjects.
2. Public Benefit: University agreements should encourage collaboration and dissemination of data for purposes consistent with its missions, but all such use must have a clear public benefit and dissemination of patient and research subject data should never be solely for financial gain, commercial development, or private benefit.
3. Justice: University agreements must consider the significant inequalities in society and health care systems and ensure that the University is not agreeing to terms that would take advantage of or increase the effect of those inequalities.
4. Transparency: University agreements should never include terms that would undermine the public trust in the University or its healthcare operations with respect to the use of Health Data.
5. Responsible Stewardship: University agreements should always consider the University's role as steward of Health Data and actively ensure that such data are protected and managed appropriately, including after the data have left the University.

Any agreement that may violate one or more of these principles should be reviewed by the local HDOC.

HDOC Contract Review Procedures

HDOC review of a particular data transaction subject to these data sharing guidelines may be requested and occur at any point in the **contract life cycle** – prior to any discussion of a contract, during the contract negotiation (which may involve numerous iterations), after a contract is signed, after a contract is approved and a new data transaction is contemplated under that contract, or even when no contract is contemplated or deemed necessary. As such, HDOC review should occur as soon as it is requested based on the information then available. Additionally, a data transaction may be a small component of a much broader set of terms, conditions and obligations that are beyond the scope and expertise of the HDOC, and these broader terms should not be part of HDOC review.

When the HDOC receives a request to review, the HDOC should review the data transaction, request any additional information necessary, and make a determination to approve the data transaction, or refuse to approve the data transaction as presented subject to any conditions the HDOC determines necessary. These conditions may include, among any number of other requirements, execution of an agreement, limiting or modifying an agreement with respect to the data transaction, requiring inclusion of supplementary documents (such as a Business Associate Agreement or BAA) or provisions to an agreement, or, for high-risk transactions, even a secondary review by the HDOC. The HDOC would then be able to respond to inquiries with respect to the project and waive any such limitations if new information becomes available. An HDOC approval may have very detailed and restrictive conditions for

high-risk or not fully developed transactions or a less restrictive approval for final projects that present lower risk with respect to Health Data.

Examples of HDOC approvals might include:

"HDOC approves the fully de-identified data transaction as presented to HDOC, conditioned on the following (unless such conditions are later waived by HDOC): An agreement will be put in place with the second party which will include, at a minimum: 1) a Business Associate Agreement addendum with the University as the Covered Entity and the Second Party as the Business Associate unless this requirement is waived by the Privacy Office; 2) the agreement will specify that the data will be used solely for the purpose of the project and returned or destroyed at the end; and 3) the agreement will require that the second party will make no attempt to identify any of the individuals in the dataset. Any substantive change to the materials presented for this project must be re-reviewed by the HDOC."

"HDOC approves the fully de-identified data transaction as presented to HDOC, conditioned on the following (unless such conditions are later waived by HDOC): Full committee review and approval by the IRB and approval by IT is secured with respect to the security of the software exchange mechanism. Any substantive change to the materials presented for this project must be re-reviewed by the HDOC."

"HDOC approves the access to identifiable patient data as presented to HDOC, conditioned on the following (unless such conditions are later waived by HDOC): The current purchase agreement is amended and executed by the parties to attach a Business Associate Agreement Addendum and a Data Security Appendix (template language available here ___) and the technical requirements of the electronic system exchange are reviewed and approved by IT. Any substantive change to the materials presented for this project must be re-reviewed by the HDOC."

Conditions placed on initial HDOC approvals should always be subject to waiver by HDOC if circumstances surrounding the initial data sharing request change to sufficiently mitigate risk. In cases where proposed risk mitigation conditions are unacceptable to the requester, the HDOC should have a process for approving acceptable alternatives if appropriate. An HDOC may receive a request for approval of a data exchange that may present issues not within the scope of the HDOC, such as an unresolved conflict of interest, and may condition its approval on review and approval by a different designated official or authorized department responsible for such approvals. Additionally, all approvals should indicate that such approval is valid only for the transaction as presented to the HDOC, and any future change in circumstances may require additional review.

Data Sharing Repository

Each local HDOC should maintain an internal repository of data sharing requests containing, at a minimum, a description of the purpose of the request, the nature of the requested data, the proposed recipient, contact information for the requestor, and the HDOC's decision. The repository should also contain copies of relevant UCOP-approved data sharing documents (including the Data Sharing Guidelines) and samples of potentially reusable language from relevant agreements. These data should be used to maintain a transparent record of HDOC-approved and denied data sharing requests and for process improvement efforts to optimize the appropriateness, consistency, and efficiency of decision making by local HDOCs. The local HDOC repository may also serve to enable documentation of a subset of local data sharing requests in a central, UC Health-managed Health Data Set Access Repository (HDAR). It is expected that data recorded in both local and central repositories will evolve together over time as both local and Systemwide HDOCs continue to adapt to the changing data sharing landscape.

Data Sharing Agreement Provisions for Sharing of Patient-Level De-identified Data

Increasingly, collaborations with non-UC third parties within the healthcare industry or startup world are built upon the sharing of de-identified data and usage of publicly available datasets. While sharing de-identified data raises fewer risks than sharing PHI and thus warrants fewer restrictions, the potential for reidentification of de-identified patient-level data, particularly when combined with other large public or consumer datasets, presents privacy risk and a risk of damaging patient trust in the institution. In addition, de-identified data may also be considered the intellectual property of the institution requiring a UC licensing agreement.

Accordingly, agreements that share patient-level Health Data that have been de-identified pursuant to HIPAA should include language similar to the following with respect to the use and downstream disclosure of these data. **Substantial deviations from the requirements below should be discussed with the local HDOC**.

Notwithstanding any other provision in this Agreement, in addition to any requirements or limitations herein with respect to information collected or derived from patients, students or employees associated with <UC> which <UC> provides to <Contractor> or allows <Contractor> to access in the course of this Agreement ("Institutional Information"), <Contractor> agrees to comply with the following additional restrictions:

<Contractor> shall not, either during or after the term of the Agreement:

i) use such Institutional Information, alone or together with any other resources, to attempt to re-identify, contact, or market to any individual or group of individuals that are the subject of the Institutional Information without the authorization of the individual.

ii) sell, attempt to sell, or disclose in exchange for any financial or non-financial value any Institutional Information.

This Section shall survive termination or expiration of this Agreement for as long as <Contractor> has access to any Institutional Information.

Appendix 2 – White Paper on "Got Health Data" Conference

Got Health Data? Moving Towards a Justice-based Model of Data Use

> The University of California is a public land grant institution with a mission and a responsibility to serve, at a minimum, all Californians. We have also an opportunity, through our educational programs and research, to serve a wider constituency in the United States and across the world. Our data are key to our ability to serve.
>
> —**Carrie Byington, M.D.**, Opening Remarks

University of California Health (UCH) has generated data on the care of more than 8 million patients across the state of California, as well as elsewhere in the US and beyond. The breadth and depth of this health data offer insights for improving quality of care, generating new scientific knowledge, and improving public health. While data-intensive collaborations in health care present great new opportunities, these collaborations also pose challenges, including complex legal, ethical, and privacy issues around patient engagement, data use, inequities in healthcare, and public/private partnership. UCH leadership recognizes the unique role and duty of academic medical systems to be innovative and strategic in utilizing data at a scale to advance science and clinical care while at the same time safeguarding the data.

To navigate these challenges, UCH sponsored a two-day conference in April 2022 entitled, "Got Health Data? Moving Towards a Justice-based Model of Data Use" (University of California Health, 2022). This virtual conference brought together experts across different disciplines including researchers, bioethicists, privacy experts, clinicians, patients, and industry leaders to discuss issues around health data use. Topics included incorporating the patient voice meaningfully and sustainably into decisions around data use, data ethics and governance, defining public good in the context of data analytics, and collaborating with government and for-profit entities. This conference builds on a 2018 report from the *President's Ad Hoc Task Force on Health Data Governance* (Gulbranson et al., 2018).

Conference speakers and panelists emphasized that health data are a valuable public good that can be used to advance the health of residents of our state. The advances already made and the tremendous promise of knowledge that can be gained mean that the users of these large data sets must be responsible stewards of this precious resource. It is that "stewardship" responsibility that leads to the need to have strong and thoughtful data governance.

The Power of Data-Analytics to Improve Health Care

Kicking off the conference, **Atul Butte** highlighted the creation of massive health databases from electronic health records and other sources as a radical transformation in the last decade (Butte, 2022, 00:06:11). Advances in artificial intelligence and machine learning have dramatically enhanced the potential power of health data for improving care and for advancing medical science. Butte emphasized the responsibility the University of California has to the institutional community and to the public to provide trustworthy stewardship and enable optimal progress from use of the data.

In her keynote, **Patricia Brennan** urged conference members to consider the NLM as a partner in utilizing health data analytics to benefit the widest range of the population (Brennan, 2022, 00:24:50). Maximizing this benefit, Brennan argued, requires extensive and ongoing engagement of members of the community in data governance, and a focus on the whole person, not just the person as a patient.

Limitations of Current Regulatory Landscape

Issues of public trust, and of potential bias and misuse of data, are important to address through strong data governance and stewardship because the existing regulatory and legal framework does not adequately resolve them. As **Kristen Rosati** of Coppersmith Brockelman PLC explained during her presentation "The Legal Landscape for Data Sharing," "[E]specially in the US, there are real gaps in legal protection, which we, as good stewards of data, need to figure out how to fix" (Rosati, 2022, 04:11:37) One reason for these gaps, according to **Rosati and James Shaw**, is that the myriad of federal and state laws addressing health data focus on the actor rather than the type of data (Rosati, 2022, 04:08:27; Shaw, 2022, 03:42:43). As a result, the existing landscape is potentially both over and under protective. It may allow for broad sharing of certain categories of data, including de-identified data, without appropriate protections against reidentification and misuse of these data in a way that undermines public trust. The complex and evolving legal paradigm may also limit large-scale data efforts that could benefit patients and the public.

Justice as a Guiding Principle

A pillar of robust governance is the concept of "justice." Across the two days, multiple experts contributed diverse viewpoints on the topic of justice, revealing that the term is complex and may be interpreted differently depending on who and what is being discussed.

For example, **James Shaw**, ethicist and researcher from the University of Toronto, defined justice as "an orientation to the study and use of health-related data in ways that aim to redress: a) the exclusions of structurally marginalized communities from systems of healthcare and public health, and (b) the oppressions faced by communities when participating in such systems" (Shaw, 2022, 03:56:00). In addition, justice must include accountability and action from "(c) the institutions responsible for governing participation." Professor Shaw supported the idea of community and public engagement in the governance process, pointing out that "the aim to use health data for the public good or common good relies on understanding of what that might entail. I would suggest that solidarity is central to that…." Shaw highlighted Nancy Fraser's approach to justice, adding process considerations to this framework. Her approach views justice as parity of participation – "Justice requires social arrangements that permit all members of society to interact with one another as peers in social life. Overcoming injustice means dismantling institutionalized obstacles that prevent some people from participating on a par with others, as full partners of social interaction" (Fraser, 2005). This definition highlights the importance of (1) public participation as an organizing principle for health data justice and (2) the importance of dismantling obstacles to participation.

Patricia Brennan noted that enacting a justice-based model requires institutions to engage with society; it demands that community input help shape the goals of research and policy. She argued for a balance of interests, "A justice-based model doesn't call for the individual rights to dominate over the institutional goals...[nor] the privileging of one professional knowledge over another" (Brennan, 2022, 00:36:55). Instead, Brennan noted that various stakeholders should come together to provide services in a way that addresses the needs of everyone receiving the service.

When considering justice as a guiding principle, speakers **Ysabel Duron**, **Donna Cryer**, **Julie Harris-Wai**, and **Deirdre Mylod** urged participants to embrace the complexity of this principle. As Duron noted, "We are talking about complexity, because we are talking about lived experience…" (Duron, 2022, 01:13:28). Harris noted, "We have a lot of data that shows that different communities and certainly communities who've been excluded from the health care system, from the public health system . . . may have a very different perspective on sharing their data" (Harris-Wai, 2022, 01:14:23). Justice requires us to recognize the unique needs of different communities and provide the appropriate resources to ensure equitable collection, access, and use of health data.

Health Data as a Public Good

How do we determine whether data-driven health research benefits the public good – whose public and whose good are being served? (President's Ad Hoc Task Force on Health Data Governance, 2018).

Speakers recognized the multitude of ways in which health data benefit the public good. For example, health data can serve to identify and track public health threats. These can be infectious, as in the recent Covid-19 pandemic; behavioral, as in the epidemic of opioid use and mortality; or environmental, as in lead contaminated water or pesticide toxicity. Without large real time data availability, information about these and other threats would take longer to raise awareness and would suffer from inaccuracy related to incompleteness and delays. Health data are also a public good as a research resource to improve health care quality and patient safety, and to identify patients most likely to benefit from specific therapies, such as targeted cancer treatments.

As **Donna Cryer** and others recognized, authentic, sustained, and dynamic engagement with the public is necessary to develop a legitimate conception of the public good (Duron et al., 2022, 01:04:01). International authors (including **Shaw** and **Cassel**) refer to public trust as a "social license," essential to sustaining large business enterprises where collective and often tacit public support is necessary.[34] To maximize the benefit of health data, it is critical to develop a governance model that fosters collective support among stakeholders. The collection, generation, access, sharing and use of health data needs to recognize the needs of all, and requires equitable, just, and inclusive approaches. To establish and continually reinforce that health data are a public good, data users should seek to build public value, address knowledge gaps, and make progress towards the achievement of public health goals visibly and transparently.

Rethinking Consent: The Importance of Trust, Transparency, and Engagement

Traditional human subjects research ethics is centered about the process of informed consent as a reflection of the principle of "Respect for Persons" as outlined in The Belmont Report, the founding document in which the principles of ethical conduct of research with human subjects is outlined.[35] Speakers discussed the limitations of informed consent, however, in the context of big data analytics. Informed consent is often not feasible in this context given the voluminous number of patients. Moreover, consent documents may not capture the nuances of data sharing that would help patients better or fully understand how their data are being used/shared, and there is ample evidence in the literature that many patients do not read these documents carefully. While recognizing these limitations, as well as the fact that current regulations do not require informed consent when sharing deidentified data (see above), speakers recognized that the principle of respect for patient autonomy remains an important goal and discussed alternative ways to achieve this goal.

Several speakers recognized transparency as a tool for achieving respect for the individual and their role in their care. **Amy McGuire** stated that "transparency is probably the most important thing for making systems and entities trustworthy" (Brennan et al, 2022, 06:23:29). Achieving transparency requires expertise in communications for broad public communication as well as visible engagement of diverse community members and patients in ongoing governance processes. As suggested by **Deven McGraw**, this includes the idea of reciprocity; "You're not just taking, but you're also giving, and you're not assuming exactly what the community wants, but you've actually engaged directly with them" (Brennan et al, 2022, 06:17:32). **Ysabel Duron**, **Donna Cryer**, **Julie Harris-Wai**, and **Deirdre Mylod** similarly

[34] Shaw, J.A., Sethi, N. & Cassel, C.K. Social license for the use of big data in the COVID-19 era. npj Digit. Med. 3, 128 (2020). https://doi.org/10.1038/s41746-020-00342-y

[35] National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. (1978). The Belmont report: Ethical principles and guidelines for the protection of human subjects of research. [Bethesda, Md.]: The Commission.

highlighted that people want to make sure their information is secure, and that the exchange of data is transparent (Duron et al, 2022, 01:04:01).

In their panel on research uses of data, **Pagan Morris**, **Paul Boutros**, **Liz Salmi**, and **Kayte Spector-Bagdady** recognized that transparency is also important for improving understanding among the general public and patients about the sources of large data sets and how they are used for research purposes (Morris et al, 2022, 02:12:31). As Spector-Bagdady explained, "patients very much understand the difference between use for themselves versus research. However, they don't fully understand the nuances of different kinds of research…" (Morris et al, 2022, 02:35:35). This is why multiple speakers recommended improving health literacy.

Moving Forward in Collaboration and Partnership

Partnership amongst all stakeholders is critical for overcoming challenges in health data governance. The conference reflected the benefits of collaboration by bringing together experts across different disciplines including researchers, bioethicists, privacy experts, clinicians, patients, and industry leaders.

Conference members recognized that the patient is an essential partner in decisions around data use. As **Julie Harris-Wai** stated, "It all comes back to that basic trust, dignity, and respect. If a community, an individual or family, does not feel that they have been treated by their health care system, or that institution, with trust and dignity, their likelihood of being interested in also sharing these very personal and politicized data points, this may be a non-starter unless we can assure them of the trust that we can maintain within the system. And certainly, that can be done through community partnerships" (Duron et al, 2022, 01:15:06). **Donna Cryer** similarly emphasized that the engagement of patients/data contributors is necessary to the performance of high-quality research and that the process of exchanging health data should flow with the person (Cryer, 2022, 02:01:41). Speakers recognized that the practical challenges of achieving the goals of patient engagement are significant and require both resources and expertise; however, speakers suggested both peer mentors and representation as potential strategies for building trust and sharing knowledge about health data within varying communities.

Government and for-profit entities also present opportunities for fruitful collaboration, and speakers provided examples of successful collaborations as well as some of the concerns these collaborations raise. **John Wilbanks** explained how public private partnerships (PPPs) enable the balance of public interest and private innovation, where the combined goals of "share and protect" can be met (Wilbanks, 2022, 05:20:23). Two existing NIH sponsored examples are the "All of Us" initiative (collecting genome data on a million volunteers), and the National COVID Cohort Collaborative (N3C), which shares COVID-19 related patient data from more than 70 US Academic Health Systems, including the UC Health Systems. At their best, Wilbanks noted that PPPs represent an institutional form for figuring out different versions of share and protect depending on context, patient engagement, and public health requirements (Wilbanks, 2022, 05:32:48). Yet these PPPs are not without challenges and risks for data providers, health systems and patients.

Panelists **Niall Brennan, Deven McGraw, and Amy McGuire** agreed based on their research and experience that patients generally trust governmental data sharing activities more than commercial entities (Brennan et al, 2022, 05:38:51). Where there is concern about sharing with government entities, it is related to the potential impact on individual rights and freedoms. Concerns about private companies center on the assumption that data are only being used for purposes of making profit. Considering the advantages of data sharing collaborations with government entities and commercial companies, **John Wilbanks** advised designing governance processes to address those concerns (Wilbanks, 2022, 05:22:43). **Deven McGraw** similarly noted that protections are needed against bad actors, though these protections should not over-regulate the good actors or hamper the power of data sharing for positive progress in public health and health care (Brennan et al, 2022, 06:40:12).

Survey Results

Cultivating responsible data stewardship necessitates engagement with data users systemwide. To initiate this engagement with the broader UC community, Sarah Dry and Susan Pappas invited attendees to participate in an interactive survey each day of the conference (Dry & Pappas, 2022). Most of the participants (95%) believed that the public does not currently understand how their health data are used by care systems in California, and 88% believed that UC should play a role in increasing the public's understanding. According to 90% of participants, the most effective method for sharing this type of information would be short videos.

When asked to submit issues that they would like to further explore after the conference, participants most frequently cited health equity, followed by community engagement. Regarding approaches to community engagement, a large majority (71%) supported engaging with local community organizations at each medical center, and a smaller majority (57%) supported the creation of a UC Community Advisory Board.

Conclusion

Governance of health data use will require oversight of the rapidly evolving health data landscape and conference members discussed several key elements of this oversight, including what the concept of justice requires in this context, transparency, trust building, and authentic patient engagement. These are all part of building a governance model that will be both robust enough to remain relevant in the context of scientific progress and flexible enough to be responsive to evolving technologies and business models. It is UC's responsibility as a data steward to create and nurture these characteristics and enable trustworthy data collaborations.

References

All of US research program | National Institutes of Health (NIH). (n.d.). Retrieved from https://allofus.nih.gov/.

Brennan, N., McGraw, D., & McGuire, A. (2022, April 20) *Roundtable: Collaborating with Government and For-Profit Entities* [Conference presentation]. Got Health Data? Moving Toward a Justice-Based Model of Data Use. https://www.youtube.com/watch?v=z15O5eM658M&t=20334s

Brennan, P. (2022, April 19) *Keynote: Incorporating the patient Person's Voice into Decisions around Data Use* [Conference presentation]. Got Health Data? Moving Toward a Justice-Based Model of Data Use. https://www.youtube.com/watch?v=z15O5eM658M&t=1490s

Butte, A. (2022, April 19). *The Power of Data Analytics* [Conference presentation]. Got Health Data? Moving Toward a Justice-Based Model of Data Use. https://www.youtube.com/watch?v=z15O5eM658M&t=371s

Cassel, C., Platt, J., Rosati, K., Shaw, J. (2022, April 20) *Panel: Ethics of Data Use and Governance* [Conference presentation]. Got Health Data? Moving Toward a Justice-Based Model of Data Use. https://www.youtube.com/watch?v=z15O5eM658M&t=14560s

Cryer, D. (2022, April 19) *Fit for Purpose: Aligning Health Data Governance and Use with the Public Good* [Conference presentation]. Got Health Data? Moving Toward a Justice-Based Model of Data Use. https://www.youtube.com/watch?v=z15O5eM658M&t=7273s

Dry, S. & Pappas, S. (2022, April 19) *Interactive Session 1* [Conference presentation]. Got Health Data? Moving Toward a Justice-Based Model of Data Use. https://www.youtube.com/watch?v=z15O5eM658M&t=11723s

Dry, S. & Pappas, S. (2022, April 20) *Interactive Session 2* [Conference presentation]. Got Health Data? Moving Toward a Justice-Based Model of Data Use. https://www.youtube.com/watch?v=z15O5eM658M&t=24671s

Duron, Y., Cryer, D., Harris-Wai, J., & Mylod, D. (2022, April 19) *Panel: Incorporating the Patient Voice into Data Use* [Conference presentation]. Got Health Data? Moving Toward a Justice-Based Model of Data Use. https://www.youtube.com/watch?v=z15O5eM658M&t=3836s

Fraser, N. (2005, November). Reframing justice in a globalizing world. *New Left Review, 36*, 92.

Gulbranson, C., Stobo, J., Sherman, R., Berglund, L. F., Khosla, P., Nosowsky, R., Mazziotta, J. C., Engel, E., Mazziotta, J. C., Blum, M., Streitz, W., Federoff, H., Chalfant, J., Currie, R., Koenig, B., Longhurst, C., Martorano, R., Ohno- Machado, L., Troncoso, M., & Wajda, J. (2018). (rep.). *President's Ad HOC Task Force on Health Data Governance* . Retrieved from https://www.ucop.edu/uc-health/reports-resources/health-data-governance-task-force-report.pdf.

Morris, P., Boutros, P., Salmi, L., & Spector-Bagdady, K. (2022, April 19) *Panel: Research Use of Health Data* [Conference presentation]. Got Health Data? Moving Toward a Justice-Based Model of Data Use. https://www.youtube.com/watch?v=z15O5eM658M&t=7979s

Rosati, K. (2022, April 20) *The Legal Landscape for Data Sharing* [Conference presentation]. Got Health Data? Moving Toward a Justice-Based Model of Data Use. https://www.youtube.com/watch?v=z15O5eM658M&t=14690s

Shaw,J. (2022, April 20) *Health Data Justice: Definitions and Directions* [Conference presentation]. Got Health Data? Moving Toward a Justice-Based Model of Data Use. https://www.youtube.com/watch?v=z15O5eM658M&t=13328s

U.S. Department of Health and Human Services. (2022, May 18). *National Covid Cohort Collaborative (N3C)*. National Center for Advancing Translational Sciences. Retrieved from https://ncats.nih.gov/n3c.

University of California Health (2022, April 19-20). *Got Health Data? Moving Toward a Justice-Based Model of Data Use.* Virtual Conference Session. https://youtu.be/z15O5eM658M?si=YNAKfRXHIny3yBBr

Wilbanks, J. (2022, April 20) *Time for a Cambrian Explosion in Public-Private Partnership Structures* [Conference presentation]. Got Health Data? Moving Toward a Justice-Based Model of Data Use. https://www.youtube.com/watch?v=z15O5eM658M&t=19222s