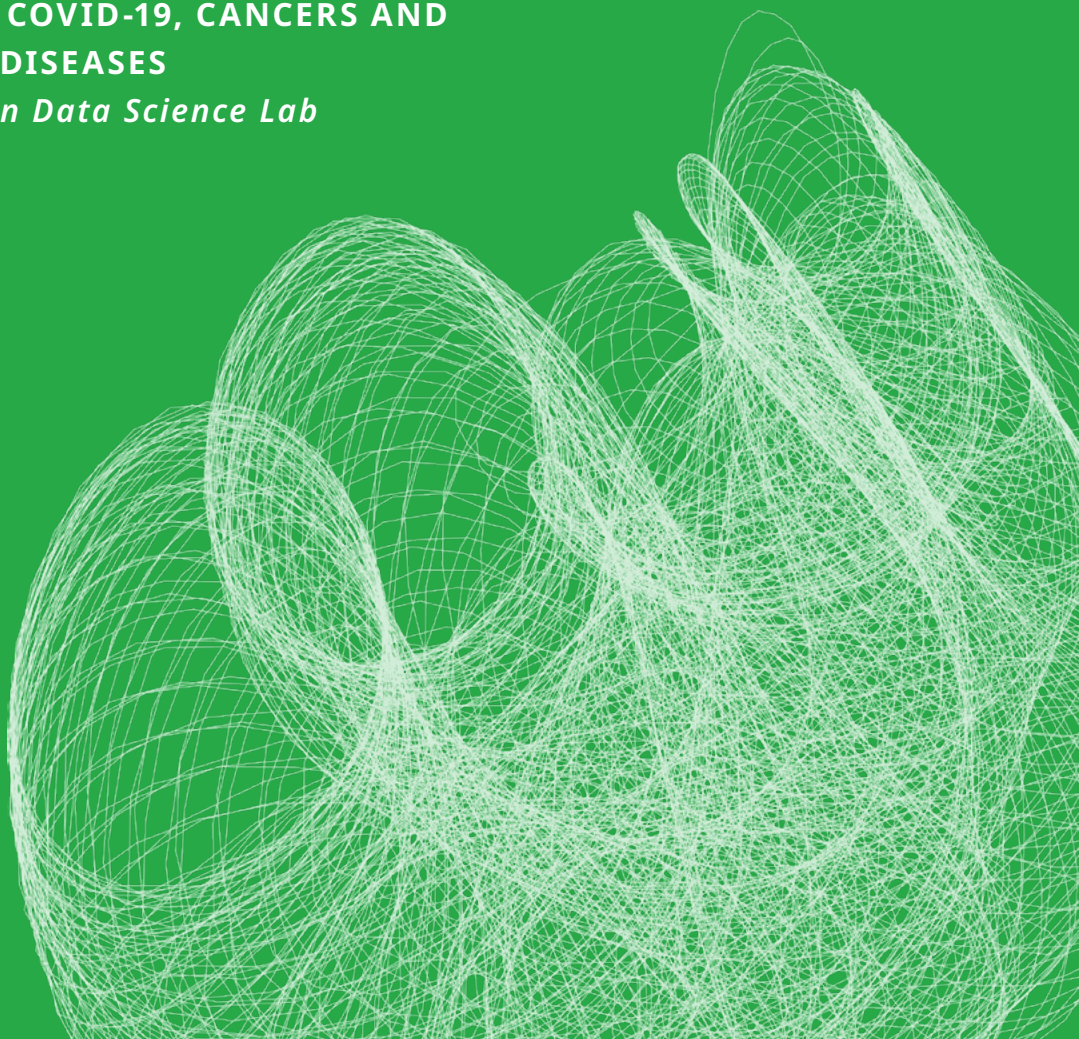
IQVIA™
INSTITUTE
FOR HUMAN DATA SCIENCE

Evolving the Understanding of the Natural History of Disease

PERSPECTIVES ACROSS COVID-19, CANCERS AND NEURODEGENERATIVE DISEASES

Insights from The Human Data Science Lab

July 28, 2020



OCTOBER
2020

Table of contents

The Human Data Science Lab Participants

EXTERNAL PARTICIPANTS

Caleb Alexander, MD
Professor, Epidemiology, Johns Hopkins Bloomberg School of Public Health

Roy D. Baynes, MD, PhD
Senior Vice President and Head, Global Clinical Development. Chief Medical Officer, Merck Research Laboratories

Pierfranco Conte, MD
Full Professor of Oncology, University of Padova, Italy
Director of the Division of Medical Oncology 2, Istituto Oncologico Veneto, Padova
Chairman of Rete Oncologica Veneta

Michael Gold, MD, MS
Vice President, Neuroscience, Abbvie

Inmaculada (Inma) Hernandez, PharmD, PhD, FACC, FAHA
Assistant Professor, Pharmacy and Therapeutics, School of Pharmacy, University of Pittsburgh

Sharon Hesterlee, PhD
Executive Vice President, Chief Research Officer, Muscular Dystrophy Association

Charles B. Nemeroff, MD, PhD
Matthew P. Nemeroff Professor and Chair, Department of Psychiatry and Behavioral Sciences, Mulva Clinic for the Neurosciences
Director, Institute of Early Life Adversity Research, Dell Medical School, The University of Texas at Austin

Oliver Ottmann, MD, FRCPath
Professor and Head of Haematology, Division of Cancer and Genetics, Cardiff University

David Paydarfar, MD
Professor & Chair, Department of Neurology. Director, Mulva Clinic for the Neurosciences. Dell Medical School, The University of Texas at Austin

Edmund Pezalla, MD, MPH
Founder and CEO, Enlightenment Bioconsult

Cristiana Sessa, MD
Professor, Oncology Institute of Southern Switzerland

Samik Upadhaya, PhD
Research Analyst, Cancer Research Institute Venture Fund and Clinical Accelerator, Cancer Research Institute

Annie Yu, PhD
Data Scientist, Cancer Research Institute Venture Fund and Clinical Accelerator, Cancer Research Institute

IQVIA PARTICIPANTS

Murray Aitken
Executive Director, IQVIA Institute
Senior Vice President, IQVIA

Stig Albinus
Senior Advisor, IQVIA Institute

Jennifer Christian, PharmD, PhD, FISPE
Vice President, Clinical Evidence and Epidemiology, Real World Solutions, IQVIA

Greg Dennis, MD
Global Head, Senior Vice President, Therapeutic Science and Strategy Unit, IQVIA

Jeffrey Hodge
Vice President, Development Solutions, Oncology Center of Excellence, IQVIA

Alan Metz, MD
Senior Vice President, Strategic Partnerships, IQVIA
Chief Medical Officer, the Americas, IQVIA

Cristina Oliva, MD
Vice President, Oncology and Head of Oncology Centre of Excellence, IQVIA

Jeffrey Spaeder, MD
Chief Medical and Scientific Officer, Senior Vice President, IQVIA

Introduction

The ongoing impact of the global coronavirus pandemic on communities, health systems, and economies around the world has elevated the urgency of understanding of SARS CoV-2 and COVID-19 related illnesses. However, despite a tsunami of scientific studies, now more than 67,753 published papers (as of July 27,2020)[1], COVID-19 still remains largely a mystery, as the origin of the virus, the pathogenesis, is still unknown, the clinical characteristics of the diseases are complex, and the understanding of the natural history of disease and related comorbidities continue to change. As we come to terms with the learnings from the pandemic, we find ourselves facing a renewed imperative for a paradigm shift in other disease areas.

Across many disease areas beyond COVID-19, there is a growing call for a better understanding of the natural history of disease, in particular, the role of prodromal disease, i.e., the early onset of disease before symptoms become apparent.

The gaps in understanding the natural history of disease are well-described in rare diseases, but the challenge is increasingly rising to the top of the research agendas in cancers and neurodegenerative disorders. There are similar challenges in many other diseases where the movement to precision medicine is leading to an enhanced understanding of diseases, previously thought as single diseases, as clusters of many separate conditions. Furthermore, factors beyond genetics, such as the environmental, social, behavioral and dimensional, are increasingly playing a role in the understanding of the progression of disease.

To discuss the challenges and explore new pathways for a shifting paradigm in disease understanding, the IQVIA Institute for Human Data Science brought together a multidisciplinary panel of 13 experts from various fields of academic research and medicine, including oncology, neuroscience, rare diseases, epidemiology and health economics.

The discussion was grounded on the evolving understanding of the natural history of disease in oncology and neuroscience using Human Data Science.

This paper brings highlights from the lively and inspiring discussion covering a broad array of perspectives and, in some cases, controversial topics such as, big data vs. phenotyping, the quality of data, the complexity of applying social determinants of health in research, and the value of knowing when there is no clinical utility.

The summary also includes some of the intriguing ideas about new pathways for disease understanding that were brought to bear during the virtual lab session.

HUMAN DATA SCIENCE AS THE VANTAGE POINT

The discussion at the virtual lab session used Human Data Science as the stepping stone for a holistic view of the natural history of disease drawing upon the capabilities available in both human science and data science and going beyond traditional clinical characteristics of disease to capture broader aspects of human health, wellness, social and environmental factors, and the impact of therapeutic intervention.

A Multidisciplinary Approach can Help Fill Knowledge Gaps in a Rapidly Evolving Environment

HUMAN SCIENCE

- Domain expertise
- Natural history of disease
- Genomics, proteomics
- Impact of diagnostics and therapeutic interventions
- Keeping people healthy, disease prevention, disease interception and reversal

DATA SCIENCE

- Data access, linkage, and management
- Machine learning, predictive analysis, advanced analytics
- Shared access to data and insights to drive alignment and improved decision-making

SETTING THE STAGE WITH COVID-19

The discussion further reviewed COVID-19 as the impetus for the topic for the first Human Data Science Lab. While a lot of the focus on the pandemic has been on the dramatic burden in terms of disease, death and economic costs, COVID-19 can also be viewed as a catalyst for positive change.

There are at least 10 key themes for Human Data Science that we can think about in terms of gaps in our current understanding and opportunities for change. One is the evolving understanding of the natural history of disease, which has been played out visibly with the respect to COVID-19. Other key themes have been arising around the need for new approaches to accelerate the discovery and development of new therapeutics and vaccines; recognition of the priority of point of care diagnostics to detect disease; and appreciation for the urgency of new models for collaboration between the private and public sector and among private entities to advance development of therapeutics and vaccines.

COVID-19 As a Catalyst for Positive Change – 10 Key Themes for Human Data Science

- Understanding the natural history of disease
- Overcoming barriers to accelerate discovery and clinical development
- Embracing new models for collaboration
- Applying digital technology to detect, track, and diagnose disease
- Sharing access to data
- Bridging gaps in understanding disease epidemiology
- Advancing patient point-of-care diagnostics
- Improving commercial viability of vaccines and curative medicines
- Getting back-to-basics in personal hygiene and vaccinations
- Accessing the global and local nature of a pandemic

Discussion of the Issues

The first section of the lab session was focused on the discussion of the issues pertaining to the natural history of disease.

1. ONCOLOGY AND NEUROSCIENCE AS OPPOSITE ENDS OF THE SPECTRUM

Juxtaposing oncology and neuroscience served as the foundation for the main discussion of the event. Similar to the challenges pertaining to the understanding of COVID-19, there are significant issues in the understanding of the natural history of disease regarding many cancers and neurodegenerative diseases, such as Alzheimer's disease, Parkinson's disease, Lewy Body disease, ALS and Huntington's disease. Yet, while there are commonalities in the gaps of disease understanding between oncology and neuroscience - for example, the important role of pre-symptomatic, prodromal disease - there are also notable differences. These include the advances in molecular diagnostics and biomarkers that have helped the evolution of precision oncology, whereas biomarkers are still at relatively early stages of development for neurodegenerative diseases. In many respects, despite the similarities, oncology and neuroscience are at the opposite ends of the spectrum.

Discussing these issues in the broader context of Human Data Science can help advance the understanding of the natural history of disease across many disease states to ascertain the complex role of multiple factors that simultaneously impact disease progress from very early, pre-symptomatic stages to full-blown, late-stage disease.

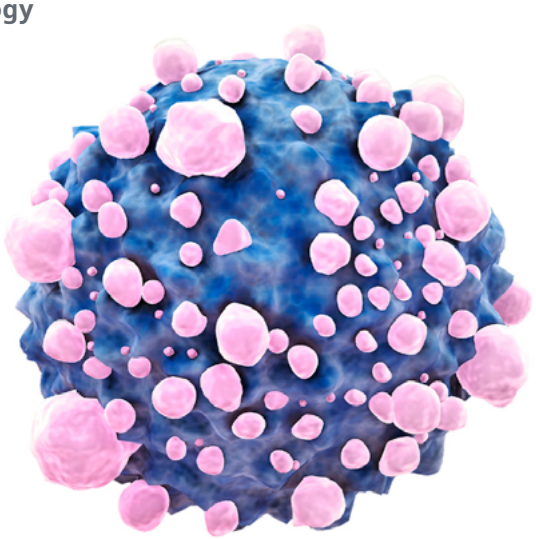
The acute need for understanding natural history of cancers

In the field of oncology, researchers and clinicians are confronted with the sobering fact that the natural history of many cancers is still elusive, despite the significant progress over the last couple of decades in genomics, cell-imaging, biomarker development and precision oncology. This is, in part, because we don't yet understand the complex interactions between genetic and biological systems with environmental, social and behavioral exposures.

There is an acute need to improve disease understanding in oncology—the natural history of every single type and sub-type of cancer. Participants acknowledged that there are better opportunities for progress today given the availability of actionable biomarkers following the major trend of transforming

The Growing Complexities Triggered by New Knowledge in Oncology

- Genomic, transcriptomic, proteomic, and metabolomic data power new insights
- Yet, these tools have opened a window to the incredible complexity of tumor heterogeneity and the challenge of accurately predicting the safety and efficacy of combination therapies based on clinical, pathology, and molecular profiles



> *"The fact is that we don't know the natural history of most of the molecular pathways of common tumors. We don't really know whether the rare molecular dysregulation of common tumors, such as lung cancer, breast cancer and colorectal cancer, has a natural history, which is similar, identical, or quite different from the natural history of what's known among tumors. This is why we need for international registries where rare dysregulations can be be annotated, not only to annotate dysregulation, but to annotate the clinical course of disease."*
>
> Pierfranco Conte, Professor of Oncology, University of Padova, Italy

oncology into the field of precision medicine. A lot of progress has been made, but there is still a long way to go in understanding what is actionable, what is relevant and what truly drives outcomes that matter among the complex parts of the proteomics and genomics components. One of the most relevant areas is the understanding of the oncogenic pathways.

The evolving natural history of neurodegenerative diseases

In neuroscience, there are similar challenges as in oncology in terms of the understanding of the natural history of many neurodegenerative diseases, but there are also many differences. When we look at many neurodegenerative disorders, such as Alzheimer's disease, Parkinson's disease, Huntington's disease and ALS, we must recognize that we are far away from truly understanding their natural history.

As in oncology, there has been progress in understanding the complex interactions in neuroscience across biology, genetics, neurotoxic pathways, and the origin of the degeneration of the brain and the early onset of brain cell death.

> *"It is just pitiful how little we know about the pathogenesis of these disorders, particularly Alzheimer's disease."*
>
> Charles B. Nemeroff, Dell Medical School, The University of Texas at Austin

Gut-brain neurotoxic pathways

There is also an exciting, emerging debate about expanding our understanding of the neurotoxic pathways of disease as illustrated in the brain-gut axis hypothesis. There is emerging evidence of this human microbiota gut-brain connection from a range of studies. For example, patients who decades ago had vagotomies to treat their severe peptic ulcer have shown significantly lower incidence of Parkinson's disease many years later compared to control patients who did not have this surgical intervention. Interestingly, the studies suggested that bacterial inflammation originating in the human gut was an important originator of neurotoxic impact on the brain cells.

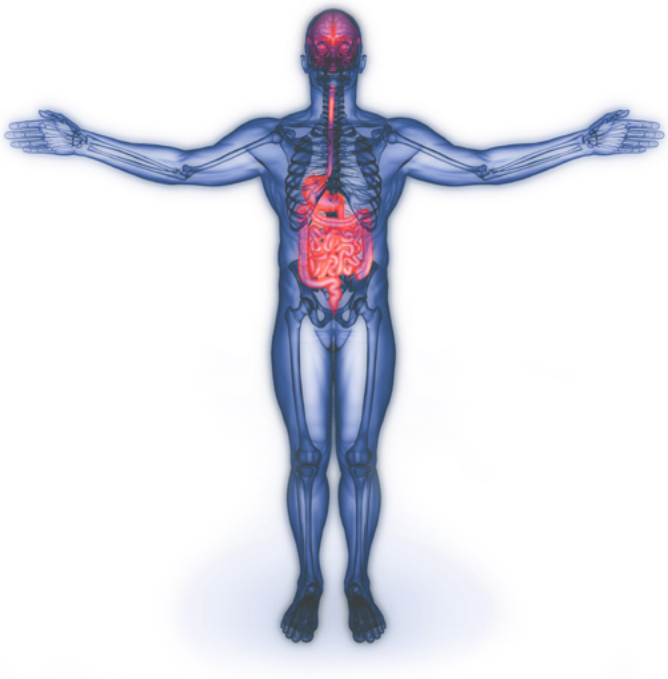
The human brain-gut axis hypothesis, if true, will have major implications for disease prevention and therapeutic intervention whether through lifestyle changes, such as diet, or antimicrobial pharmacotherapy addressing the gut microbiome and to halt neurotoxic, bacterial attacks on the brain. Thus, the vagotomy story sheds light on how an improved understanding of the natural history of disease may hold the key to future therapeutic breakthroughs.

Participants found the vagotomy study interesting as it is a falsifiable test of a hypothesis that had been floating around for a long time. Parkinson's disease may actually have an enteric origin – and very relevant to this discussion – that if you look at patients very early in the evolution of their Parkinson's, you find predominant dysautonomic features, for example, orthostatic hypotension, bowel and bladder dysfunction, and sexual dysfunction. There were all sorts of hints that there was a problem with the autonomic nervous system, and then the vagotomy study provided supportive data. This is a good example of the benefit of paying attention to the natural history of disease early on, and developing a hypothesis-generating framework that allows you to generate testable hypotheses that can create the foundation for insights about the spread of disease and potential intervention.

For a disease such as Alzheimer's, even if it were one single disease, there is growing evidence that many factors play a role in disease onset, progression and phenotypic including genetics, early childhood education and nutrition, lifestyle factors (e.g., smoking, drinking) and comorbid conditions such as hypertension and diabetes.

The Expanding Uncertainty of Neurotoxic Pathways

- Uncertainty about the natural history of disease within neurodegenerative disorders such as Alzheimer's disease, Parkinson disease, Huntington's disease, and ALS
- Complex interactions across biology, genes, neurotoxic pathways, microbiota, diet, aging, lifestyle, and behavior as exemplified in the emerging science around the brain-gut axis



"Oncology and neurosciences are at opposite ends of the spectrum. In oncology, we can access tumors and hematologic malignancies pretty directly. We biopsy them. We sequence them. We are able to interrogate them with sophisticated molecular tools. And then on the other extreme, in the neuroscience world, it is very hard to get after the tissue and the rigorous biology. We are often relying on indirect measures such as PET imaging."

Roy Baynes, Chief Medical Officer, Merck Research Laboratories

2. LARGE DATA SETS VS. PHENOTYPING

The role of large data sets is very controversial in neuroscience.

Concerns were expressed about the notion of pan-collection of data. A lot of genomic data sets have yielded very little tractable insights in terms of actionable information. Almost all neuropsychiatric disorders in the GWAS (Genome-Wide Association) studies[2] have shown us that a large number of genes have very small effects, relative to vulnerability. 300,000 enrolled patients may provide statistical significance, but it is clinically meaningless.

Trying to understand how 50 genes and their SNPS individually interact with environmental factors for increasing vulnerability to Alzheimer's, Parkinson's, schizophrenia, depression, etc., is complex. This is why we have moved to polygenic risk scores instead of looking at single genes which contribute very small effects on disease risk.

The general feeling of uncertainty was expressed. The hot areas of data science – artificial learning and machine intelligence algorithms – are in their infancy right now. At best they can identify potential pathways, but they have not even established the basics of how you can make an inference on causality versus correlation with large data sets or time series over time. Furthermore, highly promoted attempts to apply the power of supercomputing technology to enable clinical decision-making in oncology have been either unimpressive at best, or failures at worst, perhaps unfairly tainting the promise of advanced data science, such as AI and ML, due to the lack of appreciation of human science expertise in the development of the algorithms.

"I am a big fan of big data, but unless you do really significant phenotyping, the big data you collect is an absolute disaster. And you would probably agree with me that in all the studies where you just use EMR data to generate a diagnosis, the GWAS data look like crap. The fact is that in Alzheimer's, Parkinson's, and depression, you can't throw out phenotyping. You end up with the old garbage in, garbage out phenomenon."

Charles B. Nemeroff, Dell Medical School, The University of Texas at Austin

"I am relatively unimpressed by the Big Data Story. Data sciences provide interesting hypothesis generating tools. But there are issues with the data and there are very few high-quality data sets. And as a result of that, the process is noisy, and it is relatively low-level evidence. From a drug developer's point of view, I must say that we are a long way from real world evidence supplanting traditional drug development and randomized trials."

Roy Baynes, Chief Medical Officer, Merck Research Laboratories

3. UPSTREAM VS. DOWNSTREAM IN DISEASE PROGRESSION

The major question was introduced about the extent and way in which research can move more upstream to generate a better understanding of the early origins of disease rather than chasing the butterflies downstream in the disease progression when intervention is almost too late.

It was suggested to focus on at-risk populations before clinical disease manifests itself. The problem with most of the studies is that they focus on patients when they are already quite ill, in the late trajectory of disease, which makes it difficult to intercept disease before they reach that point. For that reason, we need to focus on high-risk populations.

This elevates the urgency of looking simultaneously at multiple studies and data sources, there is not one single study that will help with this. It calls for a triangulation of findings coming from imaging, long-term registries, social determinants and environmental factors, and for synthesizing information from scientists around the world.

It is imperative to learn more about the biology of disease to intervene earlier: Even when you perform a so-called "early diagnosis", you diagnose the biological phenomenon that started years before. We need to have a lot of early information, clinical imaging, biological data from the high-risk populations prior to the development of disease.

"How do we screen populations? How do we even know where to look? How do we make sure that people present themselves early as opposed to when they are symptomatic? How do we know that they are getting worse? How do we know that an intervention is improving for these patients? How do we set public policy as well as insurance policy to pay for the screening? That's why understanding the natural history of disease is so important."

Edmund Pezalla, CEO of Bioconsult, former Aetna

4. A NEW FRAMINGHAM STUDY

The question was being asked: What are the commonalities and the differences between neuroscience and oncology when we talk about prodromal disease, in other words, disease before it is a disease?

Participants admitted that we are struggling with the notion that it is possible to identify people with a pathology before it manifests itself.

The question is how to enable large-scale population screening with regular meaningful follow-up. People have been trying to do this for a long time. It would be great if we could intervene with a safe, cheap and effective therapy when somebody has a pre-symptomatic or pre-clinical diagnosis. Vaccines are the model for that kind of intervention.

"I am not sure we can identify people with a particular pathology who are completely asymptomatic, "normal", unless we are actually going to mass population screening. If you are trying to pick people who are at risk because of some known factor, for example, the APOE4 gene, you are already biasing the selection of the sample. This is really fascinating, but this is where I am struggling with these large datasets and AI. To me, its garbage in, garbage out."

Michael Gold, Vice President, Neuroscience, AbbVie

It was suggested that the single most influential study of that type is the Framingham Heart Study[3]: Having a large, engaged population over several generations is key to success. To this day, there are still discoveries coming out from the Framingham study.

5. THE VALUE OF KNOWING WHEN THERE IS NO CLINICAL UTILITY

The evolution of advanced diagnostic tools that can detect the early onset of disease many years before symptoms occur is controversial when there are no therapies available to intercept, halt the progression of, or modify, the disease. This raises the fundamental question whether there is value in knowing when there is no clinical utility.

The question is also whether all people will want to know. Today, many people with Huntington's disease prefer not to know if they are a carrier of this devasting condition for which there is no effective therapy.

In Alzheimer's disease, there are only a few treatments available and they deliver very limited symptom relief, often with significant adverse events. Will people want a PET-scan to find out whether they have an amyloid plaque build-up in their brain when there is no true disease-modifying therapy available? On the flipside, if we don't want to know, how can we advance scientific discovery? Knowing early, even when there is no treatment available, may be the way to ultimately develop the therapeutic solution.

According to neurologists and psychiatrists treating patients with neurodegenerative disorders, patients and their families want to know because it removes ambiguity and uncertainty around the disease. It also helps narrow the discussion around prognosis, and it helps inform the discussion about the type of risk the patient's children may face.

It was also stated that in Western cultures at least, it is universally accepted that the patient will want to know the details about a cancer. This might be quite different for newer degenerative disease where people don't

want to know. Furthermore, there are cultures where disclosure of serious diagnoses to only the patient is not the norm.

6. SOCIAL DETERMINANTS OF HEALTH ARE AT THE FOREFRONT

The need to incorporate social, economic and ethnic factors has come to the forefront during recent events, in particular the role of health disparities in the exposure to COVID-19.

The natural history of disease can help facilitate the understanding of the role of social determinants and how they are interconnected from the very early onset of disease to late-stage disease, thus facilitating a broad spectrum of potential interventions – both pharmaceutical and non-pharmaceutical - along the entire disease continuum from primary prevention, over disease interception to therapy for late-stage disease.

However, it is challenging to apply big data and social determinants of health to help understand a disease when you don't understand the pathology. This is an interesting contrast, which raises the need to think about the different ways big data applies to different therapeutic areas.

"I remember learning in medical school that most of health doesn't have a lot to do with healthcare. This is an area where human data science has a real opportunity by underscoring the dimensions of health outcomes that aren't a function of what doctor you see or what treatment you get, but rather where you live, who you know, what you eat. It's all these social determinants. When you start peeling back the onion and looking at some of these specific diseases, thinking about the social determinants become quite compelling. Not just the risk factor for disease, but the risks of the outcomes conditional upon having a given diagnosis."

Caleb Alexander, Professor, Epidemiology, Johns Hopkins Bloomberg School of Public Health

Discussion of the Implications

The second part of the lab session was dedicated to discussing the implications of the challenges pertaining to the uncertainties around the natural history of disease.

1. APPLYING REAL WORLD EVIDENCE

There are different areas where real world evidence can support the evolving understanding of the natural history of disease.

One of the broader points is that we will not learn everything from controlled clinical trials and that real world evidence is necessary to better understand in a naturalistic setting what medications patients are taking and what the social and behavioral determinants are. These can be better measured through real world evidence than via electronic medical record or insurance claims data captured in routine care and practice.

In terms of clinical development, natural history of disease studies are useful prior to clinical trials to help inform the prime study objective and determine the key critical end points. Natural history studies are used, in part, to understand the disease itself, and as a strategy to create a more collaborative environment and relationship with clinicians and trial sites. Finally, natural history of disease studies can serve as a potential historical control or external comparator to the single-arm studies that are evaluating treatment.

However, there are also challenges with this approach: If a Phase I or II study is underway, many patients will not want to continue in a natural history study if there is an opportunity to receive a new, potentially effective treatment. So, what happens is that we may not have the length of time to fully understand the natural history. Studies such as the Framingham Study have been valuable to provide a holistic, long-term assessment on the complex interactions between genetics, clinical, social and environmental factors.

2. THE PRECISION MEDICINE REVOLUTION IN ONCOLOGY

The oncology world has been revolutionized by precision medicine. The very specific understanding of mutations oftentimes has enabled drug discovery and probable treatment targets. Drugs tend to break down into two major categories: those which have protean or broad-based applicability, and those which are highly specific. And the more specific, the greater the need for a precision medicine tool. With the protean, the better the opportunity for treating the general population.

It is an exciting time in the oncology arena. We are starting to see major movement on survival. The American Cancer Society published data this year showing the single biggest drop in mortality from cancer ever reported - particularly in diseases like melanoma and lung cancer where major advances have been made.

However, we have progressively moved away from traditional histologic diagnosis of cancer as the mechanistic basis for therapy: As an example, it has been well-recognized that patients who have defects of DNA repair accumulate a large number of mutations over time. These are often the most immunogenic cancers that are quite resistant to chemotherapy, and these patients are uniquely sensitive to checkpoint inhibition. This has also turned out to be the same for patients who have had high mutational burdens in certain cancers. Therefore, we are starting to see drug approvals, which are based upon the precision medicine diagnosis rather than histological diagnosis.

While we have made progress in some cancers where natural history studies don't play a role anymore, we will continue to need more natural history of sub-types of cancer so we can continue the precision medicine revolution and develop treatments for areas where none exist today.

"I would argue, from my point of view, that natural history no longer plays a role in oncology, at least in the strictest of terms. I can remember one disease, chronic myeloid leukemia, where the natural history of disease ended in the '90s with the advent of a targeted agent, a kinase inhibitor called imatinib. But since the goal post is continuously moving, not only for hematological malignancies, but for all tumors, there might be a few exceptions where we really have nothing in terms of treatment."

Oliver Ottmann, Professor and Head of Haematology, Division of Cancer and Genetics, Cardiff University

3. DETERMINING VALUE WITH NO CLEAR ENDPOINT

The discussion of implications turned to the challenges of reimbursement or funding of therapy when there is no natural history of disease.

Participants acknowledged the difficulty of determining value when there is no clear endpoint for treatment. Natural history is extremely important to payers. First of all, what kind of outcomes can a patient expect with a standard-of-care therapy? And do some therapies work better than others? A lot of times payer don't actually have information about the comparative effectiveness of what's available. The understanding of natural history of disease also helps target patients. In a particular area of cancer there are going to be patients who are doing better than others.

There may be additional implications from a social point of view that are not just related to biology. If payers understand that, they may be more accepting, for example, of multiple forms of delivery of a drug. Therefore, advancing our understanding of value requires learning from what happens in the real world in terms of targeting a particular group of patients.

4. IMPROVING DATA QUALITY AND DATA INTEGRATION

There was general consensus about the challenges regarding poor quality of data and data integration. There was also agreement that a measure of data quality would be helpful, including standards for transparency and protocols for data sharing.

Payers and health technology assessment groups are used to looking at clinical trials, but they are not used to natural history of disease and real world data, so they need to understand how the data was derived. It is important to explain the data to people who are not full-time statisticians.

"Data integration is very poor, even when the data quality is good, which often isn't the case. As cancers and rare diseases increasingly are being broken down in smaller and smaller quantities, we are simply not able to test new treatments with big randomized trials with additional endpoints of overall survival. Therefore, we need innovative trial methodologies"

Oliver Ottmann, Professor and Head of Haematology, Division of Cancer and Genetics, Cardiff University

"We are incredibly inefficient in collecting data. We do collect a lot of natural history and biomarker data, but because of the silos in data methodologies, it's not being combined and maximized in a useful way. This is certainly a significant issue in rare disease where those data are hard to come by."

Sharon Hesterlee, Executive Vice President, Chief Research officer, Muscular Dystrophy Association

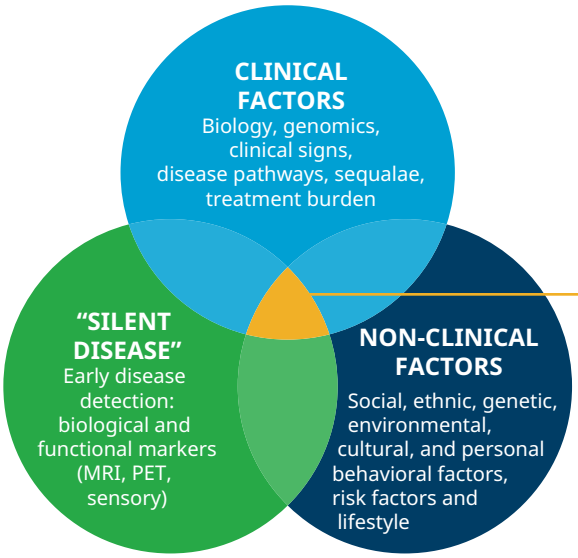
5. APPLYING HUMAN DATA SCIENCE IN THE UNDERSTANDING OF NATURAL HISTORY

There is an opportunity to test the value and promise of Human Data Science in better exploring, identifying and motivating work on the social determinants of health. So much of life does not happen in the exam room and a lot of health doesn't have to do with healthcare.

There is also a need to incorporate big data not only in terms of biology and clinical perspectives, but in terms of the impact of social, economic factors and how the health system is organized.

The COVID-19 story from Italy can serve as an illustration:[4,5] The two regions, Lombardy and Veneto, are very similar in terms of patient population, demographics, and health services. However, mortality rates from COVID-19 were five times higher in Lombardy than in Veneto despite the two regions having similar rates of infections. The two regions implemented different strategies for their lockdowns with Veneto opting for stricter containment and mass testing, but their approach to the delivery of care for COVID-19 patients was also very different. In Lombardy, most patients were hospitalized in intensive care units and they probably died because they were in intensive care units. In the Veneto region, most patients were treated at home and very few were hospitalized in intensive care units. So, in order to better understand the outcome of disease, we need to know not only the biology and the clinical aspects of a disease, but also the social environment, the economic environment and how the health systems are organized.

Enablers of Advancing the Understanding of Natural History of Disease Through Human Data Science



- Quintessential tools to get to the fundamental understanding of disease progression
- Moving beyond linking disparate datasets to curating new integrated data

New integrated multi-dimensional methodologies and tools

- Integrating clinical science (RCT, RWE), epidemiology and behavioral science with data science (AI, ML, NLP)

The Path Forward

Following the discussion during the Human Data Science Lab, there appears to be a number of important opportunities for further advancing the evolving understanding of the natural history of disease through a variety of academic research endeavors.

IQVIA and the IQVIA Institute for Human Data Science will explore research projects and continued discussions with researchers and other stakeholders in healthcare centered on the following aspects:

- **Advancing natural history of disease studies** will play an important role both in the early stage of discovery to prepare the design of clinical trials and in the later stages of the clinical development process, where continued studies of the natural history can provide important perspectives and potentially corrective insights about disease progression that are not uncovered through randomized clinical trials or real world evidence generation.
- **Building international disease registries** which are critically important sources of insights as they generate insights on the evolution of disease and ability to draw comparisons and distinctions across population segments that are highly diverse from geographic, cultural, genetic and environmental vantagepoints.
- **Exploring the opportunities for "A New Framingham Study"**, a population-based, longitudinal study that investigates the natural history of disease in an important cluster of inter-connected diseases. There is a compelling opportunity for designing a new Framingham Study with a population mix that is socially, ethnically and culturally more diverse than the original Caucasian, Northeastern Framingham cohort while simultaneously drawing from the exceptional epidemiological power of the original longitudinal, population-based study format.
- **Enhancing the understanding of the complex interactions** across genetic, social, economic, environmental and ethnic determinants of health and how they intersect with clinical determinants of disease realizing that health outcomes to a larger extent is determined by non-clinical rather than clinical factors.
- **Improving the quality of data and integration of data** is a quintessential endeavor in future-ready medical discovery, research and development. As the volume of data and new, diverse data-sources grow exponentially, there is a growing urgency to advance consensus and methods for broad standards and protocols for data quality. This is particularly important with the convergence of clinical, human science and data science that traditionally draw from different thought processes and methodologies.

For a perspective piece, *read "Modernizing the Natural History of Disease Research: IQVIA Perspectives from the Human Data Science Lab"*

References

1 https://www.natureindex.com/news-blog/the-top-coronavirus-research-articles-by-metrics

2 National Human Genome Research Institute: Genome-Wide Association Fact Sheet. Available from: https://www.genome.gov/about-genomics/fact-sheets/Genome-Wide-Association-Studies-Fact-Sheet

3 Framingham Heart Study: Three Generations of Research on Heart Disease. Available from: https://framinghamheartstudy.org

4 Fewer deaths in Veneto offer clues for fight against virus. Financial Times. April 2020. https://www.ft.com/content/9c75d47f-49ee-4613-add1-a692b97d95d3

5 COVID-19 deaths in Lombardy, Italy: data in context. Odone, A et al. Lancet Public Health. 2020 Jun; 5(6): e310. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7182509/

About the Institute

The IQVIA Institute for Human Data Science contributes to the advancement of human health globally through timely research, insightful analysis and scientific expertise applied to granular non-identified patient-level data.

Fulfilling an essential need within healthcare, the Institute delivers objective, relevant insights and research that accelerate understanding and innovation critical to sound decision making and improved human outcomes. With access to IQVIA's institutional knowledge, advanced analytics, technology and unparalleled data the Institute works in tandem with a broad set of healthcare stakeholders to drive a research agenda focused on Human Data Science including government agencies, academic institutions, the life sciences industry and payers.

Research Agenda

The research agenda for the Institute centers on 5 areas considered vital to contributing to the advancement of human health globally:

- Improving decision-making across health systems through the effective use of advanced analytics and methodologies applied to timely, relevant data.
- Addressing opportunities to improve clinical development productivity focused on innovative treatments that advance healthcare globally.
- Optimizing the performance of health systems by focusing on patient centricity, precision medicine and better understanding disease causes, treatment consequences and measures to improve quality and cost of healthcare delivered to patients.
- Understanding the future role for biopharmaceuticals in human health, market dynamics, and implications for manufacturers, public and private payers, providers, patients, pharmacists and distributors.
- Researching the role of technology in health system products, processes and delivery systems and the business and policy systems that drive innovation.

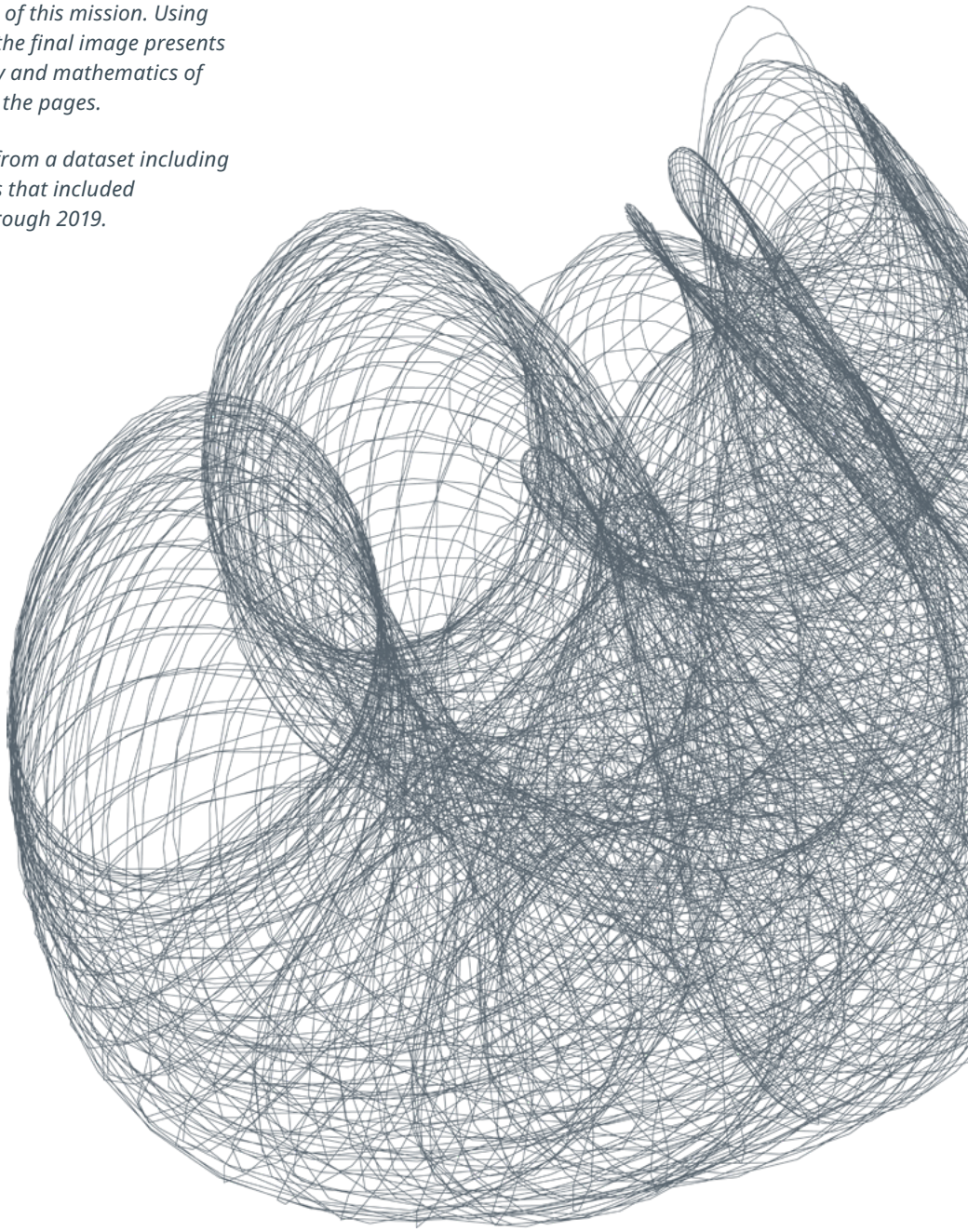
Guiding Principles

The Institute operates from a set of Guiding Principles:

- Healthcare solutions of the future require fact based scientific evidence, expert analysis of information, technology, ingenuity and a focus on individuals.
- Rigorous analysis must be applied to vast amounts of timely, high quality and relevant data to provide value and move healthcare forward.
- Collaboration across all stakeholders in the public and private sectors is critical to advancing healthcare solutions.
- Insights gained from information and analysis should be made widely available to healthcare stakeholders.
- Protecting individual privacy is essential, so research will be based on the use of non-identified patient information and provider information will be aggregated.
- Information will be used responsibly to advance research, inform discourse, achieve better healthcare and improve the health of all people.

The IQVIA Institute for Human Data Science is committed to using human data science to provide timely, fact-based perspectives on the dynamics of health systems and human health around the world. The cover artwork is a visual representation of this mission. Using algorithms and data from the report itself, the final image presents a new perspective on the complexity, beauty and mathematics of human data science and the insights within the pages.

The artwork on this report cover is created from a dataset including numbers and percentages of oncology trials that included harmacogenomic biomarkers from 2010 through 2019.



IQVIA™
INSTITUTE
FOR HUMAN DATA SCIENCE

CONTACT US
100 IMS Drive
Parsippany, NJ 07054
United States
info@iqviainstitute.org
iqviainstitute.org