National Cancer Institute (NCI)
Division of Cancer Control and Population Sciences (Epidemiology and Genomics Research Program)
Division of Cancer Epidemiology and Genetics

2020 NCI Cohort Consortium Annual Meeting
Meeting Summary Report
Monday, November 16 – Tuesday, November 17, 2020
Virtual Meeting

Monday, November 16, 2020

Opening Remarks and Introductions
Dr. Kathy Helzlsouer, *National Cancer Institute (NCI), National Institutes of Health (NIH)*
Dr. Lynne Wilkens, *University of Hawaii Cancer Center*

Dr. Kathy Helzlsouer welcomed participants and introduced the session moderator, Dr. Lynne Wilkens, who is the current chair of the Cohort Consortium Steering Committee. Dr. Helzlsouer also introduced incoming 2021 Steering Committee Chair Dr. Roger Milne and 2022 Chair-Elect Dr. Heather Eliassen. In addition, Dr. Helzlsouer recognized current Steering Committee members and the NCI organizers of this meeting. She announced that Steering Committee elections will take place in 2021 and encouraged participants to join.

Dr. Helzlsouer announced two transitions happening within the Cohort Consortium. Executive Director Dr. Nonye Harvey will be leaving the group to work for the NIH Office of the Director. Camille Pottinger, a current research associate for the Cohort Consortium, will step into the role of executive director. Dr. Harvey received an Award of Merit in recognition of her 15 years of outstanding leadership and contributions to the NCI Cohort Consortium and facilitating advances in cancer epidemiologic research.

Next, Dr. Helzlsouer made the following announcements:
- Newly released funding announcements are live on the Division of Cancer Control and Population Sciences (DCCPS) website. Visit the website for more information.
- The Cancer Epidemiology Descriptive Cohort Database (CEDCD) has undergone major renovation and cohort data is currently being updated. Look for a new and improved CEDCD website in February 2021. Contact Joanne Elena if you have any questions.

Following the announcements, Dr. Helzlsouer turned the meeting over to the chair of the Cohort Consortium Steering Committee, Dr. Wilkens. Dr. Wilkens shared the list of activities that the Steering Committee has been working on in 2020. These activities include the following:
- Established the NCI Cohort Consortium Webinar Series. To date, they have hosted two webinars. In 2021, webinars will be featured on a quarterly basis.
- Added an Associate Member Council representative on the Steering Committee.
- Planned the annual meeting.
- Reviewed new project proposals:
  - Appendiceal Cancer Consortium (APPECC) – Andreana Holowatyj
  - Population Attributable Fraction (PAF) project – Robert MacInnis
- Participated in the International HundredK+ Cohorts Consortium (IHCC) Third Virtual Summit (May).
- Continued implementation of strategic initiatives of the NCI Cohort Consortium. These initiatives include communication, research facilitation, career development, filling scientific gaps, and addressing common challenges.

To close, Dr. Wilkens shared several of the activities that the Project Group subcommittee has focused on this past year. These activities include the following:

- Continued review of Project Group activities.
- Provided project updates to the Steering Committee on the Health Effects of Cigar, Cigarillo, Pipe and Hookah Smoking (Jyoti Malhotra), Colorectal Cancer Pooling Project (Peter Campbell), and Inflammation and Breast Cancer Risk and Mortality (Laure Dossus).
- Revised annual project progress report form for consortium project groups.
- Drafted authorship policy guidelines for the NCI Cohort Consortium.
- Explored the development of guidelines on data transfer agreement and material transfer agreements (work to continue in 2021).
- Developed the NCI Cohort Consortium Project Hub/Database (coming in 2021).

Acknowledging 20 Years of Service and Leadership on the NCI Cohort Consortium
Dr. Robert Croyle, *Director, DCCPS, NCI, NIH*
Dr. Stephen Chanock, *Director, Division of Cancer Epidemiology and Genetics (DCEG), NCI, NIH*

Dr. Robert Croyle awarded Dr. Deborah M. Winn an Award of Merit in recognition of her 20 years of service and outstanding leadership to the NCI Cohort Consortium and for facilitating advances in cancer epidemiologic research.

Dr. Winn is the senior advisor to the director of the Division of Cancer Prevention at NCI. Dr. Winn has held leadership positions at NCI, including deputy director of the DCCPS and acting associate director of the extramural Epidemiology and Genomics Research Program. Prior to that she was the deputy director of the Division of Health Interview Statistics at the National Center for Health Statistics, part of the Centers for Disease Control and Prevention.

She is internationally recognized for her epidemiologic research on tobacco and head and neck cancer. Her other research interests include environmental risk factors for breast cancer, evaluating the impact of epidemiologic findings on clinical practice and public health, and the development of resources, infrastructures, and policies to facilitate cancer research.

Dr. Stephen Chanock awarded Dr. Robert N. Hoover an Award of Merit in recognition of his 20 years of service to and outstanding leadership of the NCI Cohort Consortium, and for facilitating advances in cancer epidemiologic research.

Dr. Hoover is the former director of the Epidemiology and Biostatistics Program and currently serves DCEG as a scientist emeritus. Dr. Hoover is widely known as an international leader in cancer epidemiology and an expert in hormonal carcinogenesis, as a visionary scientist who has established multiple ongoing programs of research to understand the causes of cancer in human populations, as well as a generous mentor to scores of trainees.

Over the course of his career, Dr. Hoover has served NIH and the epidemiological community by providing expert assessments of epidemiologic evidence, evaluating plans and progress of scientific endeavors, sitting on search committees, and advising on training programs. He has often been called upon by NCI and NIH to brief members of Congress and their staff on issues related to cancer epidemiology and science policy. He has been an advisor to many other government entities, including the Food and Drug Administration, the Office of Management and Budget, the US Military Cancer Institute, and the Public Health Service.

Session I: COVID-19 Research in Cohort Studies
Moderators:
Dr. Katie O'Brien, *National Institute of Environmental Health Sciences (NIEHS), NIH*
Dr. Shine Chang, *University of Texas MD Anderson Cancer Center*

The International HundredK+ Cohorts Consortium (IHCC): Integrating Large-Scale Cohorts to Address Global Scientific Challenges
Dr. Geoffrey Ginsburg, *Duke University School of Medicine*

Dr. Geoffrey Ginsburg presented on the IHCC. The vision of IHCC is to create a global network of large cohorts (with multidimensional data from diverse populations) that will be maximally utilized to enhance scientific understanding of the biological, environmental, and genetic basis of disease in order to improve population health. This vision is meant to focus on discovery, translation, and population health sciences.

The premise of IHCC arose from the notion that for decades, large cohort studies have been established worldwide. Due to the size, ancestral origins, and geographic boundaries of cohort studies, there are analyses limitations. But by combining these studies, it enables researchers to address pressing global health questions no one cohort study can answer alone. Ultimately, the combined data enhances the value of each study and leverages the investments in them.

The first International Cohorts Summit was held at Duke University in March 2018. The goals of the meeting were to explore prospects for collaborations and compatibility of cohorts, discuss a searchable global cohort registry, and to understand the barriers to data and specimen sharing. The criteria for cohort participation were (1) greater than or equal to 100,000 participants, (2) not selected for disease, (3) available biospecimens, and (4) potential longitudinal follow-up. Some exceptions were made for middle- and low-income countries.

From the beginning, meeting participants quickly realized several challenges:
- Combining cohorts presents complexity when there is limited documentation of available data.
- Lack of standardization and harmonization of questionnaires.
- Inability to move, send, receive, or utilize data/samples due to regulatory restrictions and national laws.
- Lack of standards for phenotyping and health outcomes.
- Cross-cultural differences in risk tolerance and privacy.

Possible solutions to these challenges include:
- Articulating data standards to encourage and facilitate sharing.
- Standardizing data collection prospectively using tools such as PROMIS and PhenX.
- Moving analysis to data sets rather than vice versa.
- Developing automated approaches to phenotyping based on electronic health records.
- Using digital health devices or apps for phenotypes.
- Using existing frameworks to help address privacy, security, and consent, such as *All of Us* and GA4GH.

While undertaking the work, participants considered the compelling scientific questions that could be addressed with these types of studies, such as:
- Risk associated with rare exposures and outcomes.
- Generalizability of risk factors and associations.
- Population-specific determinants of health.
- Social or cultural determinants of health.

- Country- or cohort-specific risk predictions.
- Human knockout identification and phenotyping.

Following the first meeting, the Cohort Consortium moved forward with the following three overarching goals and workstreams: (1) develop the information technology to promote "discoverability" of the cohorts and sharing of data; (2) develop a robust scientific plan that leverages cross-cohort collaboration and maximizes input from the global scientific community; and (3) develop the policy agenda to advance data sharing and engagement with industry.

In the second annual Cohort Consortium meeting, participants met in Reykjavik, Iceland. Attendees included 117 representatives of 67 cohorts from 29 countries. From that meeting, the group left with the following actions:

- Finalize information technology and policies to support the science.
- Pilot science and demonstrate that it can be done.
- Develop a formal collaborative structure and charter.
- Create rules of engagement for biopharma and tech firms.

From the second annual meeting came the proof of concept study on polygenic risk scores led by Patrick Sleiman of the Children's Hospital of Philadelphia. The pilot demonstrated cross-cohort data to develop polygenic risk scores for asthma, blood pressure, BMI, type 2 diabetes, and height. Eight cohorts' data from five counties were used for a federated analysis. The trans-ethnic scores outperformed all population-specific scores in non-European cohorts with similar predictive values. The federated approach of generating scores at each site had challenges, but the process worked well for efficient analysis.

An outcome of the IHCC work has been the IHCC Cohort Atlas Project. The project aims to survey and collate data dictionaries for all IHCC cohorts, semantically harmonize the cohort metadata, and develop an online cohort atlas to enable discovery across IHCC cohorts. The project has brought together several axes of cohort data, e.g., disease status, data use, sample collection parameters, genotype, and phenotype. It also has gathered a highly diverse set of more than 100 cohort data dictionaries.

The third annual meeting was held virtually in 2020 and attendees included 177 representatives of 87 cohorts from 27 countries. They coalesced around a visionary charter and the path forward, introduced the IHCC Cohorts Atlas, finalized the IHCC charter and governance, and identified new scientifically meritorious cross-cohort research projects. A new activity came out of the third annual meeting, which was to develop an IHCC scientific agenda for COVID-19.

Dr. Ginsburg closed the session by offering a brief overview of two of IHCC's COVID-19 studies.

- The first study, led by Dr. Sarah Baumeister from the University of Oxford, is reviewing COVID-19's widespread impact on mental health. The aims of this project are to catalogue the availability of IHCC cohorts that have included mental health assessments during or after the outbreak of the COVID-19 pandemic, conduct cross-cohort analyses to address key questions, and characterize the neuropsychological and cognitive manifestations or complications of COVID-19 infection.
- The second project, led by Dr. Michele Ramsay of the University of Witwatersrand, is studying why COVID-19 appears to have lower prevalence in African countries. The study aims to explore host genetics of COVID-19 infection and disease progression in South African populations to generate knowledge to inform a precision medicine approach to the COVID-19 pandemic.

Looking ahead, IHCC hopes to continue its work in developing scalable demonstration projects. The group aims to expand the Atlas project and increase access throughout the broader scientific community to encourage collaborations. Leaders hope to engage cohorts from low- and middle-income countries, increase publication, and continue to proceed toward sustainability.

Learn more about IHCC at ihccglobal.org.

Q&A Session
The virtual floor opened for participants to ask questions to Dr. Ginsburg. A participant asked about IHCC's funding sources. Dr. Ginsburg shared that NIH and the Wellcome Trust have equally funded several of the IHCC projects. Additionally, project funding is put up by other supporting organizations such as the Davos Alzheimer's Consortium. He also noted that if anyone has project or study ideas, to bring them to the attention of the group and they will consider them. Interested participants should visit the website for more information on sharing project ideas.

A participant asked how many of the cohorts are able to go back and get the data about study participant experience during the pandemic. Dr. Ginsburg shared that IHCC currently doesn't know, but will know when IHCC gets there.

A participant asked Dr. Ginsburg to elaborate on the lessening of cohort restrictions for low- and middle-income countries, specifically for African countries, South American countries, and indigenous populations. Dr. Ginsburg responded by stressing that an IHCC charter tenet is to build strong partnerships with low- and middle-income countries, both in science and publication authorship. More information about this effort is available on the IHCC website.

A participant asked, due to the regional response to the COVID-19 pandemic, how these studies have been adjusting for the variation in recommendations for the pandemic. Dr. Ginsburg shared that in the haste to get some of the projects set up, IHCC may not have considered this a high-priority factor. He noted that they will include information on local regulations, social distancing, and closures. He also stressed that like all things during the COVID-19 pandemic, they are learning things in real time.

Session II: Lightning Talks
Moderators:
Ms. Camille Pottinger, *The Scientific Consulting Group, Inc.*
Dr. Heather Eliassen, *Harvard Medical School and Harvard T.H. Chan School of Public Health*

Appendiceal Cancer Consortium (APPECC)
Dr. Andreana Holowatyj, *Vanderbilt University Medical Center*

Dr. Andreana Holowatyj shared that appendiceal adenocarcinoma incidence rates have increased by 232% over the last two decades. Because of the rare nature of this digestive system tumor, it's challenging to conduct studies using an individual cohort approach, and as such, researchers felt that the NCI Cohort Consortium was uniquely suited to overcome this limitation and make advances in appendiceal cancer.

The overarching objective of this project is to understand risk factors, etiologies, and prognostic factors of appendiceal cancer and utilize this knowledge to reverse the increasing disease burden as well as inform clinical, molecular, and population-level features that contribute to appendiceal cancer disparities.
Key notes of APPECC:

- Nested case-control study
- 25+ international cohorts
- 1,240+ cases of appendiceal cancer
  - Will be matched 10:1 to non-cancer controls within cohorts
  - Will also be matched 2:1 to right/left-sided colon and rectal tumors within cohorts

How to Add Free Data Visualizations to Your Cohort's Website
Dr. James Lacey, *City of Hope*
Ms. Emma Spielfogel, *City of Hope*

Dr. James Lacey highlighted the importance of data visualization. He explained that data visualization can help answer key questions such as, *does your cohort have data on…?* Or, *how many biospecimens are available in your cohort?* He noted that along with data visualization being a goal in the June 2018 NIH Strategic Plan for Data Science, it also can be easy to implement with free and user-friendly tools.

Dr. Lacey shared an example of data visualization benefits by highlighting the California Teachers Study's use of data visualization. The research team began using data visualization as a way to present numbers and frequencies to new researchers, and later expanded its use to include interactive visualizations on the California Teachers Study website. The templates could be reused to answer common and reoccurring cohort questions.

Next, Dr. Lacey turned the presentation to Ms. Emma Spielfogel to provide a demonstration of the California Teachers Study's data visualization. Find all of the study's public visualizations at https://public.tableau.com/profile/California.teachers.study#!/.

Population Attributable Fraction Project (PAF)
Dr. Robert MacInnis, *Cancer Council Victoria*

Dr. Robert MacInnis opened the presentation by highlighting the strategy for reducing the cancer burden, which is to target key preventable risk factors, specifically the proportion of cancers prevented if exposure to its risk factors were removed or reduced. With this in mind, Dr. MacInnis explained that the overarching goal is to estimate PAFs for cancer incidence and mortality.

He explained the limitations of using published exposure-cancer associations, such as assumed independence between exposures, lack of subgroup data available, lack of interaction information, no accounting for competing risks, and differing lengths of follow-up.

He continued by explaining the limitations of estimating associations from cohort data, such as exposure prevalence from cohort is not representative of the general population, past burden is estimated, external representative prevalence data is often not contemporary, lack of subgroup data available, and no accounting for competing risks.

By using the Pooling Project of Prospective Studies of Diet and Cancer (DCPP) resource, the project plans to benefit from the large international data with their strength of association analyses, stronger evaluation of less common cancers, risk factors interactions, subgroup analyses, etc., and pair this information with country-specific exposure prevalence. Most of the data needed from DCPP has been collected and harmonized.
The project plans to apply for funding from World Cancer Research Fund in late 2020.

Opioid Cohort Consortium (OPICO)
Dr. Mahdi Sheikh, *International Agency for Research on Cancer*

Dr. Mahdi Sheikh shared that OPICO was created to address a growing global health concern—overuse of opioids. This project works to investigate the effects of regular opioid use on mortality and on cancer development. He highlighted the many limitations in the current evidence on opioid effects, such as confounding effects and biases from linkage studies, no opioid use data in most cohorts, limited number of opioid users in cohorts with opioid use data, and variation in patterns of use, legislation, etc. by country. He stressed that a comprehensive consortium-based approach is needed.

OPICO is supported by the IHCC and the Global Genomic Medicine Collaborative. OPICO's overarching aim is to build strong international resources for multidisciplinary scientific studies on the use of opioids and their long-term hazardous effects. The project looks to organize data on opioid use from large-scale prospective cohorts, compile new data on opioid use in large cohorts through linkage to national records, assess the type, distribution, and extent of opioid use across diverse populations, and determine the association of opioid use with cancer incidence and mortality.

Metabolomics of Obesity and Obesity-Related Cancers
Dr. Deirdre Tobias, *Harvard T.H. Chan School of Public Health*

Dr. Deirdre Tobias opened the presentation by sharing that excess body weight contributes to risk factors for more than 13 cancer sites, and 9% of the cancer burden in developed countries.

Dr. Tobias explained that the project aims to prospectively assess previously identified obesity-related metabolites with incident composite endpoints of obesity-related cancer risks and to identify novel obesity-related plasma metabolites and their association with incident obesity-related cancer risk. The precision prevention goal is to identify metabolites underlying the obesity/cancer relationship to enable targeted prevention strategies.

Q&A Session
The virtual floor opened for participants to ask questions of the presenters. The first question was for Dr. Holowatyj, asking if she has observed similarities or differences between appendiceal risk factors and colorectal cancer risk factors. Dr. Holowatyj explained that there were no observed risk factors with appendiceal cancer with the exception of surgical quality. She continued by sharing that because of this, there are significant advances to be made in appendiceal malignancies.

The next question asked Dr. Lacey what data visualization software was used. Dr. Lacey responded that Tableau was used in the California Teachers Study. A follow-up question asked about the amount of time required to clean up the data and prepare it for the visualization process. For the California Teachers Study, the project team used Tableau internally before using the free public version (links shared in the chat). In terms of the time required, Ms. Spielfogel explained that a lot of time was dedicated to ensuring that there wasn't any data that shouldn't be exposed.

The next question asked Dr. Tobias if her project has been linked to the Consortium of Metabolomics Studies (COMETS). Dr. Tobias shared that it was not linked to the COMETS Consortium, but that she would be interested in exploring the consortium and what it has to offer.

The next question was directed to Dr. Sheikh regarding illicit and recreational use of opioids and how they are capturing that while looking at these relationships. Dr. Sheikh shared that some cohorts (especially OPICO Asian cohorts) have already collected data on the illicit use of opium, which is a type of opioid that is usually used for recreational purposes. Their main concern is the lack of cohort analysis on opioid prescription use data. They are trying to look at these two sets of data in a new merged approach in an effort to harmonize the data, using examples in Australia and Scotland.

Session III: Cohort Consortium Project Group Updates
Moderators:
Dr. Hazel Nichols, *University of North Carolina Gillings School of Global Public Health*
Dr. Holly Harris, *Fred Hutchinson Cancer Research Center*

Associate Member Council (AMC) Update
Dr. Holly Harris, *Fred Hutchinson Cancer Research Center*

The AMC is a representative body of early-career investigators within the NCI Cohort Consortium where membership is targeted to post-doctoral fellows, researchers, and early-career investigators, generally within 10 years of terminal degree. By default, all investigators who meet these criteria are considered members of the AMC.

The purpose of the AMC is to engage and support early-career investigators through professional development, career networking opportunities, and research collaborations with the cohort consortium, and position members of the next generation to attain leadership roles within the consortium. The AMC was started following the 2018 annual meeting under the leadership of Celine Vachon and is a part of the career development portion of the NCI Cohort Consortium's strategic plan.

AMC has completed several activities over the past two years. Members have created the AMC charter, held an AMC kick-off session and roundtable event at the 2019 annual meeting, compiled an email list for the AMC, and hosted a webinar on supporting early-career investigators in cancer, which had over 120 attendees.

In the future, they plan to continue career development webinars and grow the AMC email list.

Updates on the Biliary Tract Cancers Pooling Project (BiTCaPP)
Dr. Jill Koshiol, *NCI, NIH*

BiTCaPP is providing new insights into the biliary tract cancer etiology. To date, they have 29 prospective studies with 2.9 million participants, and more than 5,000 incident biliary tract cancer cases. BiTCaPP has published several papers over the course of the past two years, covering a wide range of topics related to biliary tract cancer, such as family history, smoking and alcohol use, reproductive factors, etc.

Currently, BiTCaPP is undergoing studies on exogenous hormone use and diabetes. In the future, researchers plan to learn more about aspirin and statin use.

Dr. Jill Koshiol presented an example on the association between statin use and biliary tract cancers in the United Kingdom's clinical practice research datalink.

Additionally, she featured a proposal to examine circulating hormones and biliary tract cancer risk. They aim to examine the associations between sex-steroid hormone levels and gallbladder cancer (GBC) and extrahepatic bile duct cancer (EHBDC) risk in postmenopausal women and men older than 50 years of age. They hypothesize that the increased pre-diagnostic estrogen levels will be associated with an increased risk of GBC, while increased pre-diagnostic androgen levels will be associated with increased risk of EHBDC.

Liver Cancer Pooling Project (LCPP)
Dr. Peter Campbell, *American Cancer Society*

When LCPP started, the incidence of liver cancer had been increasing in the United States for over 30 years. The reasons for the increase of primary liver cancer were not clear, as no large epidemiological studies of liver cancer in the United States had been constructed. As a result, LCPP was formed in 2008 to pool data from existing US cohorts.

To date, LCPP has published nine key papers. The topics have spanned aspirin use, excess body weight, coffee intake, smoking and alcohol, and reproductive hormones.

Dr. Peter Campbell shared an example of an ongoing project within LCPP: the gut-liver axis and how it relates to the risk of liver cancer. Samples have been received from over half of the cohorts, and samples are in the process of being selected and sent from the remaining cohorts.

Additional exploratory efforts within LCPP include viral exposure signature, perfluoroalkyl substances, helicobacter species, and magnesium metabolite signatures.

Ovarian Cancer Cohort Consortium (OC3): New Initiatives and Future Directions
Dr. Renée Fortner, *German Cancer Research Center*
Dr. Britton Trabert, *NCI, NIH*

Dr. Renée Fortner opened the presentation by sharing a brief history of OC3. OC3 1.0 was established in 2012 and was originally funded by the United States Department of Defense (DOD). Dr. Shelley Tworoger established the data coordinating center at the Channing Division of Network Medicine at Brigham and Women's Hospital. Originally, the baseline data were harmonized to enable ovarian cancer risk factor studies with detailed consideration of histotype, aggressiveness, and anatomic site. Further studies were conducted pooling existing biomarker data from nested case-control studies in the consortium. The project also established the genetics database, and to date, OC3 has published 10 papers and one is currently in progress.

OC3 2.0 is moving beyond the OC3 1.0 baseline data with an extension and expansion through recently funded projects, including a second DOD grant in 2019 to support collecting updated follow-up data and tumor tissue collection. There has been a noticeable expansion in the racial and ethnic diversity of the OC3 via NCI supplemental funding. Additional adjacent DOD-funding has supported early detection consortium. Several biomarker studies are currently planned for the coming years, specifically using blood-based markers and extended tissue collection (TMAs).

Dr. Britton Trabert shared examples of data collected to date and the data harmonization process.

Q&A Session

The virtual floor opened for participants to ask questions of the presenters. The first question was for Dr. Trabert about the data sharing agreements with OC3. Dr. Trabert shared that with the transition of the data coordinating center from Brigham and Women's Hospital in Boston to the H. Lee Moffitt Cancer Center in Florida, it required OC3 to reestablish data sharing agreements with its cohorts. In this renewal, cohorts were given the option of reestablishing their agreements with Brigham and adding Moffitt versus establishing a new agreement with Moffitt.

The next question, directed to all presenters, asked how interested people could apply to receive access to their data. Dr. Koshiol shared that they would need to speak with the cohorts and review the data sharing agreements that they have in place before they could consider sharing the data. Dr. Campbell echoed Dr. Koshiol's explanation. Dr. Trabert explained that OC3 has a proposal system that interested parties may submit for consideration and approval. For each project, cohorts have the option to opt-in or opt-out of data sharing opportunities.

The next question for Dr. Trabert was, *how would the data transfer or use agreements (DTAs/DUAs) process work for new cohorts that join?* Dr. Trabert explained that new cohorts that are interested in contributing data to OC3 would set up an agreement with the Moffitt Data Coordinating Center.

The next question, directed to all presenters, asked how the projects engage early-career investigators. Dr. Campbell shared that when he started with LCPP, he was an early-career investigator. He explained that if researchers build a good resource, early-career investigators will come to you. LCPP has had a lot of success with this. Dr. Trabert echoed Dr. Campbell's statement. She added that opportunities such as the Cohort Consortium are great opportunities to get the word out about ongoing studies, research findings, etc. Dr. Fortner highlighted the enthusiasm from senior investigators in the OC3 since its inception as an important factor, together with their encouragement of early-career investigators to be engaged and contribute from the outset. Dr. Koshiol echoed the group's sentiments. She also acknowledged that early-career investigators are vital in moving the work forward.

The closing question, directed to all presenters, asked how to increase underrepresented minority early-career investigators. Dr. Campbell suggested adding fellowship funding as a possible approach. Dr. Koshiol shared that her team is having ongoing conversations about diversity and inclusion. Dr. Campbell suggested the Cohort Consortium explore diversity and inclusion in next year's meeting.

Tuesday, November 17, 2020

Session IV: Participant Engagement
Moderators:
Dr. Minouk Schoemaker, *The Institute of Cancer Research, London*
Dr. Erika Rees-Punia, *American Cancer Society*

The All of Us Research Program: Engagement Strategies and Lessons Learned
Mr. Justin Hentges, *All of Us Research Program, NIH*

The *All of Us* Research Program is a longitudinal effort to gather data from 1 million or more people living in the United States to accelerate research and improve health. By taking into account individual differences in lifestyle, socioeconomics, environment, and biology, researchers will uncover paths toward delivering precision medicine—or individualized prevention, treatment, and care—for us all. The program's mission is to accelerate health research and medical breakthroughs, enabling individualized prevention, treatment, and care for all of us.

To date, the program has enrolled 356,000 participants, of which 271,000 participants have completed the initial steps. Approximately 80% of current participants self-identify as belonging to one or more populations that has been historically underrepresented in biomedical research.

The program's engagement strategy is to meet participants where they are and to build trust. By using a trusted messenger, tailored trainings, engaging tools, and personalized messaging, the program has seen engagement impact. Examples of activities include webinars, informational booths, e-newsletters, etc.

Mr. Justin Hentges featured two engagement model case studies: (1) Facilitating Community Conversations: LULAC, ACCESS, and UnidosUS; and (2) Partnership Between Community Organizations and Health Care Provider Organization – YMCA and University of Wisconsin Madison.

To date, more than 2.4 million people have been engaged, 4,500+ hours of conversations have been had, 2,555 activities have taken place, and the program has received 15.3+ million impressions. Due to COVID-19, the program's in-person recruitment halted in March 2020, but in-person recruitment has restarted over the past couple months.

Positive Partnerships: Engaging Black Women in Health Research
Dr. Kimberly Bertrand, *Boston University School of Medicine and Slone Epidemiology Center*

The Black Women's Health Study (BWHS) was established in 1995 among 59,000 African American women across the United States. Every two years, the program shares a follow-up questionnaire. To date, the response rate is over 80% at each follow-up cycle. To accomplish this high response rate, Dr. Kimberly Bertrand explained that it takes a substantial amount of work by her team. Dr. Bertrand stressed the importance of building trust as a key component of the high response rate.

The program follow-up procedures include up to seven mailings, three months apart, with at least one mailing sent via Priority Mail. Participants can also complete the survey via the web platform or on the phone. While on the phone, program staff are able to build strong relationships with their participants.

Dr. Bertrand highlighted direct quotes from program staff and program participants.

The program shares its findings in a biannual newsletter, through its website (frequently visited pages include the study publications, program video, research team information, and program history), social media (Facebook), and in-person community events (events are often facilitated by request of the participants).

In 2019, the program commissioned a short film to highlight the history of BWHS and to share its plans of the future. It was shared at the American Public Health Association conference.

In 2020, the program celebrated its 25-year anniversary. Given the COVID-19 pandemic, anniversary plans moved virtual with a 2020 webinar series.

Looking ahead, the program plans to engage participants in program research in a more formal way through the Participant Advisory Group. This group provides input and advice on questions included in biannual surveys, serves as a sounding board for future research directions, and participates in pilot studies.

Breast Cancer and the Environment Research Program (BCERP): Community Engagement and Participant Retention
Dr. Susan Pinney, *University of Cincinnati College of Medicine*

The Cincinnati BCERP puberty cohort study was designed to determine the effect of environmental exposures around the timing of events of puberty. Conducted at three sites, the program enrolled girls between the ages of 6 and 7 years old and followed them for 12 years.

The retention methods used included:
- Cards and newsletters to parents and advocates.
- Cash (amount dependent on procedures completed).
- Gift cards (for completion of questionnaires and diet recall interviews, participants received a $50 gift card).
- Thank-you gifts (T-shirt and age-appropriate gifts, e.g., Polly Pockets, books, crafts, etc.).
- Community engagement activities (gala, parent information sessions, etc.).

While attending clinical visits, the program provided participants craft activities and movies while waiting, growing up female coloring books, handheld electronic games to use between exam stations, etc. In addition, participants participated in creating videos about the clinical visit.

Overall, the program of 379 participants had a 60% retention rate at the end of 12 years of follow-up. Through a retention survey of 230 of the girls, researchers learned about the best retention strategies. The retention rate varied by age group, but overall the best retention method was the cash method. Using thematic analysis, researchers asked the parents why they kept their daughters in the study. Trending words included cancer study, help, research, and important, among others. When asked a similar question, the participants stated that they stayed involved to help advance medicine and because of external motivation.

The program is currently working to extend its data and biospecimens to the scientific community for collaborations. To gain access to the data, interested collaborators need to first file an application for collaboration. Contact Dr. Susan Pinney at susan.pinney@uc.edu if you have any questions or are interested in a collaboration.

Q&A Session

The virtual floor opened for participants to ask questions of the presenters. The first question was directed to Mr. Hentges. The question asked, *how do you assess the effectiveness of programs like this?* Mr. Hentges explained that they assess the project at the local level by leveraging the relationships they have with their partners. They are currently developing ways to assess the project at the national level.

The next question was directed to Dr. Bertrand and asked, *how did the program build the community?* Dr. Bertrand explained that the community grew over time through outreach and building on existing relationships within the community. Facebook has also been a helpful tool to connect participants across the program, but because of the importance of maintaining participants' confidentiality, investigators and staff do not participate in connecting participants to each other. She stressed that the mailings, newsletters, and phone conversations contribute to the high response rate of the follow-up survey.

There was a follow-up question directed to Dr. Bertrand about *whether participation in events (e.g., community events, webinars) over time is associated with higher response to surveys and data collection projects within the BWHS?* Dr. Bertrand responded that although it hasn't been studied directly, local community events probably attract too few attendees to really make a difference in response rates and that the webinar series is too new to tell. She also noted that the goal of these events is not to increase response rates but to disseminate information and "give back" to the community, consistent with the mission of the BWHS to improve the health of Black women.

There was also a question in the chat about how the BWHS prioritizes in-person events, given the national distribution of the study population. Dr. Bertrand responded, "Priority for in-person events is largely dependent on resources (time & $): we almost always accept an invitation to speak if local. If travel is required, then it might depend on whether travel costs are reimbursed."

The next question asked the panelists if *any of them had published their findings*? Dr. Pinney shared that BCERP has a manuscript that they have submitted to several journals without success. She asked for suggestions from the group.

The next question asked for more information on engaging the Latinx community, specifically, *what activities have the panelists done to engage with this diverse community?* Mr. Hentges responded by acknowledging how diverse the Latinx community is and stressed that there really are communities within this larger Latinx community that should be considered. The *All of Us* Research Program, with the help of its local and key national partners, has held fireside chats, roundtable discussions, etc. to engage with and really understand this community. He emphasized the importance of having strong partnerships within the community they are trying to reach.

The next question asked the panelists *how they reengage registered participants who have stopped engaging with their program*? Dr. Pinney shared that they had to go back to the institutional review board (IRB) to

ensure the methods they wanted to use with the participants were OK. They also kept note of the times when girls left the program. Dr. Pinney noted that as girls got older, got jobs, and became more involved in after-school activities, etc., their participation in the study declined. In response to that, the program became flexible and tried to accommodate the participants needs and schedules by scheduling Saturday and evening clinical visits. Dr. Bertrand shared that consistency is key. They sent seven mailings and would follow-up via the phone. Dr. Bertrand also noted that the personalization of the outreach was key with handwritten notes, images of the team, etc. Mr. Hentges shared that retention is a priority for the *All of Us* Research Program. The program is currently trying to improve retention with follow-ups via email and phone. In recent months, the program has offered to share genetic health information as a way for participants to reengage.

The next question asked the panelists about their survey dissemination process and about the length of their surveys. Mr. Hentges explained that they use digital surveys that vary in length. They are considering the use of paper surveys in some outreach activities. Dr. Bertrand shared that reducing the survey from four pages to two pages significantly improved their response rates.

Session V: Data Sharing in Cohort Studies
Moderators:
Dr. Lynne Wilkens, *University of Hawaii Cancer Center*
Dr. Samuel Antwi, *Mayo Clinic*

Biologic Specimen and Data Repository Information Coordinating Center (BioLINCC)
Mr. Sean Coady, *National Heart, Lung, and Blood Institute (NHLBI), NIH*

In order to take full advantage of NHLBI supported clinical trials and epidemiologic studies and maximize their research value, data should be made available, under appropriate terms and conditions, to the largest possible number of qualified investigators in a timely manner. With this thinking, NHLBI created its data repository. Over time, and as the website grew, BioLINCC was established to coordinate the activities of the NHLBI Data Repository and the NHLBI Biorepository.

- The NHLBI Data Repository stores and distributes data sets from NHLBI clinical studies. It launched in 2000.
- The NHLBI Biorepository stores and distributes biospecimens from NHLBI clinical studies. It launched in 1975.

The NHLBI Data Repository has grown slowly over time, with heightened growth in the last 3–5 years. The current controlled access portfolio of the NHLBI Data Repository has 54 observational studies (781,633 participants) and 170 trials (465,651 participants).

Mr. Sean Coady shared several utilization metrics of the NHLBI Data Repository. Key points of interest include:

- New hypothesis generating research questions.
- Proposed data analysis.
- Review of statistical methods.
- Comparison group information.
- Clinical prediction/Risk prediction.
- Data shared into the data repository typically receives a data service request within 6–9 months of posting, with the majority receiving at least one request and publication within the first three years.

- Of completed data access requests, 23% indicates primary users have 0–5 years of research experience.
- Of completed data access requests, 59% did not indicate funding support.
- Of completed data access requests, 59% did not indicate training support among publications with funding acknowledgement.

Mr. Coady shared NHLBI's data sharing model. With this model in mind, NHLBI completed a comparison of journal publication category among articles published by study investigators and articles published by repository investigators. The comparison included general population cohort studies through 2014.

Order Online for Data Sharing: Menus, Self-Service, and Automated Delivery in the California Teachers Study

Dr. James Lacey, *City of Hope*
Ms. Emma Spielfogel, *City of Hope*

Dr. Lacey noted the universal challenge of getting data from a cohort to a cohort consortium project. At City of Hope, they have been asking themselves how to improve data exploration and data analysis.

While working with the California Teachers Study, their goal has been to improve data analyses by speeding up the process. The overall program infrastructure of the California Teachers Study includes the public website, controlled access (remote desktop environment), and study management. Each piece is designed for different users, holds different data and documentation, and has different levels of accessibility (submitted access forms are required). For example, to access aggregate data, users visit the public website. To access individual data, users visit the controlled-access site.

City of Hope plans to launch the California Teachers Study Secure Web Application in the coming months. Researchers can use the app to configure their data and receive access to web app's URL, among other things. The app automatically creates five custom data and documentation deliverables for every project: (1) a project-specific dataset and formats file, (2) a summary PDF with study population story, (3) an SAS data call with starter code, (4) an R data call with starter code, and (5) a data dictionary matching the project data. Ms. Spielfogel provided a brief demonstration of the web application.

Q&A session

The virtual floor opened for participants to ask questions of the presenters. The first question asked Dr. Lacey and Ms. Spielfogel *how much time it took them to prepare the data for their application?* Ms. Spielfogel explained that their team did a lot of upfront work to identify variables and present them the way they needed to be presented. A follow-up question was if City of Hope developed this themselves. They shared that they worked with a developer to create the application. Design and development began in March 2020.

Dr. Lacey shared that City of Hope will present the application in a webinar in January 2021, with more information to come.

The next question for Mr. Coady asked about the lag time between the cohort collecting the data to depositing the data into the repository. Mr. Coady explained that it takes place over a two-year time period and the data file type is SAS.

Another question for the panelists asked how a user would request genetic and nongenetic data together. Mr. Coady explained that both are in dbGaP. City of Hope explained that the modular approach would correspond to data domains and would allow people to request their data as needed.

European General Data Protection Regulation (GDPR) Updates
Moderators:
Dr. Marc Gunter, *International Agency for Research on Cancer (IARC)*
Dr. Joanne Elena, *NCI, NIH*

Update on Navigating the GDPR
Mr. Robert Eiss, *NIH*

Mr. Robert Eiss explained that in May 2018, the GDPR superseded the EU Data Protection Directive, which was adopted in 1995. GDPR applies to the 28 member states of the EU, Iceland, Liechtenstein, and Norway. Together they make up the European Economic Area (EEA). He noted that the United Kingdom is implementing GDPR despite Brexit.

Personal data under GDPR is a set of data to which GDPR applies more broadly than that under the Health Insurance Portability and Accountability Act (HIPAA). GDPR does not apply to anonymized data such as the following conditions:

- Identifiability is judged on a facts and circumstances test, taking into account all the means reasonably likely to be used.
- Pseudonymized data remains personal data.
- Non-anonymization safe harbor under GDPR, unlike HIPAA.

Implications if GDPR applies include a prohibition on processing special categories of personal data absent an applicable exception. Special categories of personal data include racial or ethnic origin, data concerning health, data concerning a natural person's sex life or sexual orientation, genetic data, biometric data used for the purpose of uniquely identifying an individual, and political opinions, religious or philosophical beliefs or trade union memberships. Additionally, fines from infringements under the GDPR can be extensive, e.g., Google.

GDPR requires a legal basis be in place to permit the transfer of personal data from the EEA to jurisdictions lacking adequate data protection legislation, e.g., the United States.

Mr. Eiss presented GDPR commentary examples, including one featuring Dr. Francis Collins of NIH.

Since 2018 there have been challenges to GDPR. For example, Schrems II challenged whether US law regarding government access to personal data for national security purposes adequately protects EU citizens.

Mr. Eiss highlighted some the creative challenges and tensions of GDPR.

- Data privacy vs. open science.
- Harmonization vs. member state autonomy.
- Data anonymization vs. scientific utility.

Next, Mr. Eiss featured a sampling of projects affected by GDPR data transfer requirements, such as the Finnish National Institutes of Health and Welfare Genomic Studies of Type 2 Diabetes, Clinical Trial Readiness

for Spinocerebellar Ataxia Type 1 and Type 3, and the International Blood and Marrow Transplant Research Program, among others.

The presentation concluded with a brief summary of the major challenges and ambiguities for the US scientific community and identified possible paths forward.

Challenges:
- Appropriate legal basis for processing personal data including special categories.
- Varying standards of anonymization across EEA.
- Requirements for consent for future research uses.
- Compliance with right to withdraw, while meeting ethical and legal obligations to retain data.
- Legal basis for data transfers outside EEA when standard contractual clauses are not feasible.

A path forward:
- Pursue code of conduct that creatively reconciles differences in EU-US data privacy standards, for EU approval.
- Advocate for revisions in the standard contractual clauses.
- Seek a more uniform and workable definition of anonymization.
- Develop template data use agreements and consent for clauses to be adapted by US-EU investigators and consortia.
- Develop platforms to share preferred or best practices in implementing GDPR, to reduce current risk-averse behaviors across Europe.

Possible Data Transfer Solutions and Where We Go from Here
Dr. Giske Ursin, *Cancer Registry of Norway*

Dr. Giske Ursin opened by thanking Mr. Eiss for his presentation and acknowledging his work in this field. She continued by echoing several of Mr. Eiss's statements around the challenges and tensions of the GDPR.

On November 10, 2020, the European Data Protection Board (EDPB) released a six-step roadmap for data transfer. The steps are:
1. Know your transfers.
2. Identify the data transfer tool mechanism.
3. Assess whether the tool is effective (or if any laws of the data importer's country impinge on the tool).
4. Identify and adopt supplementary measures.
5. Complete procedural steps if you have identified effective supplementary measures.
6. Reevaluate at appropriate intervals.

During Dr. Ursin's presentation, she discussed the first four steps. Key points include:
1. Know your transfers.
    a. Issues related to data sharing affect direct sharing and remote access, as well as future sharing and already shared data (ongoing collaborations). This impacts current data sharing and future data sharing.
2. Identify the data transfer tool mechanism.
    a. Rely on appropriate safeguards (Article 46). See the guidelines published by EDPB. Several of these safeguards conflict with US federal law.

    b. Key challenges for NIH with the GDPR and US federal laws include archiving laws in the US and the redress mechanism (in the case of serious security breach of a federal US institution involving research data rights of US citizens and European citizens).
        i. A possible solution would be to provide remote access to European data on a European server, combined with Article 46 mechanisms.
3. Assess whether the tool is effective (or if any laws of the data importer's country impinge on the tool).
    a. Consider if any laws or practices impinge on the effectiveness of the Article 46 transfer mechanism. Possible questions:
        1. Which federal/state laws apply to you that would impinge on standard contractual clause requirements?
        2. Mass surveillance – Do you have relevant previous experience with requests from public authorities to disclose data?
        3. What is the data storage method? (Cloud storage could be an impediment)
4. Identify and adopt supplementary measures.
    a. Necessary if the data exporter's assessment is that legislation in the data importer's country impinges on the effectiveness of the Article 46 transfer tool.
    b. Examples of technical supplementary measures include encryption when sending and pseudonymization.
    c. Strong pseudonymization procedures are important!

Dr. Ursin concluded with possible solutions, stressing that researchers let the syntaxes travel, not the data. Researchers should consider harmonizing data by using data dictionaries, analyzing data by sending syntaxes to the European institutions or to one European node, and sending researchers to Europe to conduct analyses on European data.

Q&A session

The virtual floor opened for participants to ask questions of the presenters. Dr. Joanne Elena asked the presenters if individual consent helps in this process? Mr. Eiss answered by sharing that explicit consent is one of the derogations under GDPR. In consent documents, if researchers indicate that the participants' data will be used or transferred out of the EEA, then consent works for GDPR. However, it becomes impractical for consent given before 2018. Dr. Ursin echoed Mr. Eiss's statements.

The next question asked why cloud storage would be an impediment to GDPR. Mr. Eiss and Dr. Ursin both noted that cloud storage is a very complicated legal question. Mr. Eiss continued by explaining that cloud storage policies or procedures can vary by country and by organization within each country.

The next question asked about bringing together data in cases where some may live on a US server while others may live on a European server. Dr. Ursin says it's a very tricky process and bringing together, or harmonizing, the data isn't as easy. Ideally, sending US-based researchers to Europe could improve the data harmonization process.

The next question asked if there would be interest in bringing together a group of Cohort Consortium members to discuss GDPR. Mr. Eiss and Dr. Ursin both agreed this would be a good idea and useful to all.

The next question asked the presenters where they go or who they go to when new information is released about GDPR policies and procedures. Dr. Ursin explained that she typically turns to her team of lawyers to help navigate the new information.

Dr. Elena opened the virtual floor for Mr. Eiss and Dr. Ursin to offer closing remarks. Mr. Eiss noted that NIH is making active efforts to meet with the EDPB to elevate these issues and pose possible solutions. There is a feeling among researchers that the those making GDPR decisions do not have a complete understanding of what researchers are working through. He hypothesized that things could relax with the recent changes in leadership in the US.

Session VI: Cohort Consortium Project Group Updates
Moderators:
Dr. Jeannette Schenk, *Fred Hutchinson Cancer Research Center*
Dr. Nonye Harvey, *NCI, NIH*

Lung Cancer Cohort Consortium
Dr. Hilary Robbins, *IARC*

Dr. Hilary Robbins opened the presentation by explaining the Lung Cancer Cohort Consortium's work to develop a protein biomarker panel for early lung cancer detection. The motivation of this work is rooted in the context of lung cancer screening. Two current lung cancer screening studies confirm that low-dose screenings can reduce lung cancer mortality among people at high risk.

Next, Dr. Robbins reviewed the paradigm for risk-stratified screening.

The INTEGRAL U19 funded program from NCI hopes to develop a custom protein biomarker panel for lung screening eligibility and management. The program is broken out into three smaller projects—the first project focuses on genetics, the second project is the cohort consortium study, and the third project involves screening studies. Project 2 (cohort consortium work) is designed to follow three phases: full discovery (November 2019; review of 1,200 proteins), targeted discovery (March 2020; narrowing down to 477 proteins), and validation and risk prediction (November 2020; custom panel developed of up to 21 proteins).

Key notes:
- The project uses Olink – high-throughput multiplexed immunoassays.
- The study population consists of current and former smokers, with cases occurring within three years of blood draw.
- Full discovery cohorts include EPIC and NSHDS.
- Targeted discovery cohorts include MCCS, SCHS, HUNT, CPS-II.
- Main results: 20 proteins were successfully replicated in the targeted discovery phase.
- The 21 proteins for the custom panel were selected by LASSO regression and have strong single-protein associations. They must be consistent in their association across cohorts, and the proteins must be specific to histology, lead time, etc.

Looking ahead, Dr. Robbins shared that the team looks forward to finalizing the selection of proteins for the custom panel, organizing the sample shipment to Olink for the all validation-phase cohorts, finalizing data harmonization, and beginning the validation phase of data analysis.

Dr. Robbins noted that the data will soon be accessible for analysis, securely and remotely, via the IARC Analytical Hub. Some analyses planned in the Lung Cancer Cohort Consortium database include the validation of lung cancer risk prediction models for screening in the UK, validation of risk models in Iranian

and Chinese cohorts, validation of risk models in US minorities, developing a risk model for Asians, and developing risk models for African Americans and Hispanic Americans.

HPV Cancer Cohort Consortium
Dr. Paul Brennan, *IARC*

The HPV Cancer Cohort Consortium was initiated in 2009 by IARC and NCI. It involved nine cohorts and the German Cancer Research Center. Some of the main initial findings include:

- HPV16 E6 antibody predicts risk of oropharyngeal cancer (OPC) 10+ years prior to OPC diagnosis.
- Confirmed in the NCI Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial.
- Extremely high specificity.
- HPV16 E6 antibody diagnosis also predicts anal cancers prior to diagnosis.

Currently, the HPV Cancer Cohort Consortium is investigating how long before diagnosis individuals with oropharynx cancer become seropositive, what the sensitivity of this marker is for both oropharynx and anal cancer when compared to "gold standard" tumor tissue markers, and what the absolute risk of oropharynx cancer is among individuals with an HPV16 E6 positive result.

HPV16 E6 antibody and subsequent OPC (published in Annals of Oncology 2019):
There are nine participating cohorts. From these cohorts, the consortium was able to receive 743 OPC cases and 5,814 controls, serial blood samples for 111 cases, and tumor tissues for 277 cases. Blood was collected an average of 11.4 years before diagnosis. Cases occurred between 1979 and 2015 with the median year 2004. 22 of 5,814 (0.4%) controls were HPV16 E6 positive and 195 of 743 (26%) of cases were HPV16 E6 positive.

- 30 cases were robustly HPV16 E6 seropositive through the follow-up up to 25 years before clinical diagnosis.
- 17 cases sero-converted during the follow-up period.
- HPVC3: Tumor tissue was collected on 277 cases from JANUS and EPIC cohorts.

HPV16 E6 antibody and subsequent anal cancer (in preparation):
There are nine participating cohorts. From these cohorts, the consortium was able to receive 387 anal cancer cases, serial blood samples for 95 cases, and tumor tissue was collected on 132 cases. Blood was collected an average of 11.6 years before diagnosis. Cases occurred between 1975 and 2015 with the median year being 2005. 85 of 387 (22%) cases were HPV16 E6 positive overall.

- 12 cases were robustly HPV16 E6 seropositive throughout the follow-up up to 20 years before clinical diagnosis.
- 24 cases sero-converted during the follow-up period.
- 89 of 127 (70%) tumors were HPV16 positive, indicating a strong increase in sensitivity toward diagnosis.

The key results of the study thus far:

- Specificity of HPV16 E6 antibodies in absence of HPV-related cancer confirmed to be extremely high (>99%).
- Sensitivity appears to be around 90% for HPV and OPC.
- Sensitivity appears to be lower for anal cancer, and is largely predictive of anal cancer within subsequent five years.
- Seroconversion is around 12 years prior to OPC diagnosis, but can be up to 30 years.

Next, Dr. Paul Brennan shared the preliminary analyses of the absolute risk for OPC after a serology test for HPV16 E6. In short, he noted that these risks are generally consistent with rescreening at a short interval for other cancers, or perhaps with diagnostic workup for older men. If we rescreen at a short interval, we need a non-invasive method.

The presentation concluded with a summary of the consortium's plans for the future. The team hopes to identify circulating proteins that predict imminent development of HPV-related oropharyngeal or anal cancer among people who are seropositive for HPV16 E6. Additionally, they are looking ahead to try and identify whether circulating HPV DNA predicts imminent development of cancer among people who are seropositive for HPV16 E6 and evaluate whether either circulating proteins or circulating HPV DNA predict prognosis for HPV-related oropharyngeal or anal cancers.

Q&A Session

The virtual floor opened for participants to ask questions of the presenters. Dr. Jeannette Schenk started by asking Dr. Robbins about the distribution of race and ethnicity within the study population for which the biomarker panels are being developed. Dr. Robbins responded by sharing that her team has done their best to include the cohorts who have a strong representation of minority groups. She noted that they may consider combining the data of all the minority groups from across all 24 cohorts for analyses focused on minorities. Minority representation continues to be a focus of the consortium.

The next question asked Dr. Brennan, *do you think the male/female difference is biological or some missing confounders?* Dr. Harvey added to the question, asking how the consortium plans to expand race and ethnicity in the study. Dr. Brennan shared that the risk for HPV for OPC is higher among men than it is among women. He noted that among HPV-positive women, there are a great number of false positives. He continued by sharing that the consortium continues to develop hypotheses of how to refine the risk estimates based on the biomarkers. The consortium continues to explore ways to expand the cohort and data in race and ethnicity.