

**National Cancer Institute (NCI)
Integrated Canine Data Commons (ICDC) Steering Committee (SC)**

**Teleconference
Wednesday, November 13, 2019**

Participants (*Present)

External Committee Members

Matthew Breen*
Renee Chambers*
Dawn Duval*
Allison Heath*
Will Hendricks*
Warren Kibbe*
Debbie Knapp, ICDC-SC Chair*
Cheryl London
Jeff Trent*
Roel Verhaak*
Shaying Zhao

Internal Committee Members (NCI, National Institutes of Health, and Frederick National
Laboratory for Cancer Research [FNLRC])

Matthew Beyers*
Allen Darry
Toby Hecht*
Paula Jacobs*
Tony Kerlavage
Erika Kim*
Amy LeBlanc*
Christina Mazcko
Philip Musk
Elaine Ostrander*
John Otridge
Ralph Parchment, ICDC-SC Managing Secretary*
Connie Sommers*
Greg Tawa*

Others

Lori Lydard
Tara Whipp*
Mary Cerny (writer)*

Opening and Welcome

Drs. Knapp and Parchment opened the meeting at 11:33 a.m. ET and welcomed attendees.

Minutes of the October Meeting

The minutes of the October 16, 2019, ICDC-SC meeting were accepted as written.

Reports from the Working Groups

Data Governance Advisory Board (DGAB) Chair's Report

The DGAB's initial draft guidance and workflow document for prioritization of the type of data for the ICDC, as well as the evaluation criteria and the intake form for the first round of submission to the canine commons, have been simplified. The revisions focus on identifying the key pieces of information needed at this point, with options for investigators to provide narratives about the data they are collecting and to submit the rationale for why their data are important and should be shared. Next steps include converting the simplified intake form into a web page and including an email address that can be used to submit the intake form.

The DGAB prototype guidelines, evaluation criteria, and intake form are available in the ICDC-SC Box folder. ICDC-SC members were asked to review the revised/simplified intake form and files and send comments to Dr. Kibbe or Mr. Beyers within the next week so that the documents can be finalized. Additional comments will be incorporated to generate final drafts for approval.

Best Practices Subcommittee (BPS) Chair's Update

BPS and NCI Semantic Infrastructure (SI) Team

The BPS has taken the opportunity to align issues of harmonization and formatting of the ICDC through coordination with the NCI SI Team. A second meeting of the NCI SI Team and members of the BPS and DGAB was held on October 18. Harmonizing elements and simplifying components of the ICDC will facilitate access to and use of data for analysis. Some of the semantic services available from the NCI SI Team include the NCI Thesaurus (NCIt), the cancer Data Standards Registry and Repository (caDSR), Clinical Data Interchange Standards Consortium (CDISC) data standards, and the BRIDG model.

One of the goals of this collaboration is to try to identify the unique value of the canine commons for reanalysis of data in the data sets that are being curated. One strategy to achieve this goal is to create a unified approach to calls for a range of elements, including single-nucleotide variants, complex structural arrangements, germline polymorphisms, and fusions. The ICDC is also interested in taking a deep look at germlines across breeds, tumor type, and in affected versus healthy dogs.

Committee members recognized that researchers do things based on their specific projects and objectives but that such disparate approaches can lead to different interpretations and, in turn, present difficulties in doing meta-analyses across projects. To overcome these obstacles, it will be important to go beyond the raw data and standardize terminology and definitions and to

develop a uniform framework for a cloud-enabled pipeline for genomics data based on “gold standard” data sets.

No definitive conclusions or plans are in place yet, but the team will look to other platforms for guidance. For example, the larger NCI Genomic Data Commons (GDC) has had challenges in accessing data from three different data centers. To address discrepancies in how data are defined, the GDC has implemented a voting method to reach consensus on a common set of variance calls when there are differences in semantics, tools, and other key factors. The cBioPortal for Cancer Genomics was identified as another good example of a platform for visualization and download of large-scale human cancer genomics data sets for meta-analyses.

The semantic team initially will work separately from the individual working groups (WGs) and the ICDC-SC. The team will continue to discuss and update the ICDC-SC as the discussion progresses. The process will be opened up to the full ICDC-SC once key issues are addressed and decisions are made.

Drs. Kibbe and Trent will continue to take the lead on the semantic team, but all ICDC-SC members are welcome to participate. There was a suggestion to include Heather Gardner, D.V.M., Tufts University, in these discussions, in particular in the development of the cloud-based pipeline for genomics data.

Imaging WG Update

Drs. LeBlanc and Jacobs are working with the Virginia Tech IT and neurosurgery staff on writing a script for how to curate canine images that includes a series of steps for transfer and inclusion into the ICDC. Dr. LeBlanc continues to take the lead on developing a test process using two of the approximately 200 cases in the canine clinical glioblastoma multiforme (GBM) cohort with relatively complete data sets across all of the designated parameters (clinical/pathology, imaging, genomics, immunology) to determine how to integrate these components into the ICDC. The location of MRIs for dogs in the GBM cohort has been identified. The process of selecting the best “true” test cases for the initial prototypes is underway, taking into account the fact that not all of the cases in the GBM cohort have genomic data or the full complement of elements to be included in the ICDC.

The current focus is on creating gold standard data sets, using the first two cases from Virginia Tech. Other ongoing tasks include cleaning up and formatting the data to determine how best to enter the data sets into the ICDC. The team is also looking into identifying the next site to contact and work with, most likely either UC Davis or Texas A&M.

Genomics WG

The WG is developing a framework to identify cancer-related pipelines. Further discussion is needed as to how to define data input and how to harmonize data. There may be some value in doing reanalysis for harmonization. The WG agreed that the ICDC should be agnostic about whether the data are tumor vs. normal and that the commons should include germline as well as cancer-specific data but that the focus should be on cancer data.

Clinical and Pathology WG

The logistics for collection of data through REDCap are working well. The group is currently focused on standardization and will be taking the Clinpath libraries from REDCap to establish normal ranges. These activities will interface with the Immunology WG. The WG has discussed scanning PDF versions of pathology reports, but it is not clear yet where this information would be uploaded and stored. Another point to consider is whether additional comments about the reports should be allowed.

Immunology WG

The WG will put together a list of reliable biomarkers and collaborate with the Clinical and Pathology WG as noted above.

Uploading/Storing ICDC Data

Committee members identified several places where various data for the commons could be transferred and uploaded, such as the bioinformatics database Sequence Read Archive (SRA) for genomics data and The Cancer Imaging Archive (TCIA) for slides and images. There was concern, however, that uploading data to various sites or repositories, instead of a central location, could be confusing for both data submission and access and would not be the most efficient use of available resources. Further discussion and a decision on a central location to store and upload data are needed.

Other Issues

Upcoming Moonshot Initiative Meeting

Several members of the ICDC-SC are planning to attend the Cancer Moonshot Collaborative Meeting, November 18–20, 2019, in Bethesda, Maryland.

Administrative Items

Upcoming ICDC-SC Meetings

The ICDC-SC is planning to skip the December meeting to give the subcommittees and working groups time to focus on their tasks and activities. The regular ICDC-SC meeting schedule will resume in mid-January 2020. The date and time of the January teleconference are TBD. Dr. Parchment will forward meeting information and materials ahead of the January meeting. If needed, an ad hoc meeting will be convened in December.

DGAB and BPS Meetings

Dates and times of upcoming meetings are TBD. Details will be distributed to committee members when available.

Honoraria

A copy of the honorarium reimbursement form for ICDC-SC meeting participation was distributed ahead of the teleconference. External members were reminded to forward paperwork for their honoraria to Ms. Lydard. Anyone having problems with the form or reimbursement should contact Dr. Parchment.

Action Items

- The revised draft DGAB guidance/workflow/intake documents have been uploaded to the ICDC-SC folder on the Box site. Committee members were asked to review these materials and submit comments to Dr. Kibbe or Mr. Beyers within a week of the current meeting. Dr. Kibbe and Mr. Beyers will address the ICDC-SC's feedback and upload the updated drafts for final review and approval.
- Drs. Kibbe and Trent will contact Dr. Gardner about working with the BPS, DGAB, and the NCI SI Team.
- Topics for future meetings should be forwarded to Drs. Knapp and Parchment.
- Dr. Parchment will forward logistics information and materials before the January teleconference and notify members if a decision is made to convene in December.

Adjournment

The meeting was adjourned at 12:06 p.m. ET.