Integrated Canine Data Commons
Data Governance Advisory Board
Draft Guidelines for Evaluating Data Submissions During Prototype Phase

November 12, 2019

Abbreviations

Term	Meaning
CRDC	Cancer Research Data Commons – a network of nodes brought together by the NCI to share cancer related data
DCF	Data Commons Framework – reusable framework that provides Authentication and Authorization and indexing services
IACUC	Institutional Animal Care and Use Committee
ICDC	Integrated Canine Data Commons - part of the CRDC that contains canine study data for broad sharing
IRB	Institutional Review Board
DGAB	Data Governance Advisory Board – this body which evaluates submissions for inclusion in the ICDC
FFRDC	Federally Funded Research and Development Center
FNL	Frederick National Laboratory for Cancer Research– an FFRDC that facilitates cancer research on behalf of NCI
NCI	National Cancer Institute of the NIH
NIH	National Institutes of Health

Introduction:

The Integrated Canine Data Commons (ICDC) will receive data from many projects and provide the community with relevant access to that data. During the prototype phase, not all requests to upload data to the ICDC can be accommodated due to the ICDC's focus and the effort and costs associated with bringing the data into the ICDC. Therefore, the ICDC must develop, document, and adhere to a Data Governance Process. This Data Governance Advisory Board will define and administer that process.

DGAB Mission:

Advise NCI Senior Advisory Committee, based on measurable criteria, on the priority for input into the ICDC of various data sets as received from potential submitters. Devise and publish a submission process for potential submitters. Publish results of submission evaluations and priority to ensure transparency. Report to the Steering Committee on a regular basis.

Process:

1. Initiation of submission request – submitter completes submission packet for data set. Submissions are accepted on a scheduled basis to be reviewed by the DGAB. Any submissions made after the deadline for that submission period will be reviewed on the next submission period.

2. FNL submission packet review – FNL staff will review the submission packet for completeness and provide feedback to the submitter as needed. Only completed packets will be forwarded to the DGAB for evaluation.

3. DGAB evaluation – On a regular basis, the DGAB will meet virtually and review, evaluate, and prioritize the submission packet. Prioritization will be done across the entire portfolio of submissions and will be updated as new submissions are received.

4. DGAB recommendation – Following the DGAB meeting, a recommendation on priorities will be made to the NCI Senior Advisory Committee.

5. NCI Senior Advisory Committee recommendation – The NCI Senior Advisory Committee will determine final status and priority of submissions to the ICDC. Projects selected for submission will be published on the ICDC website. Based on these recommendations, the ICDC Data Management Team will begin working with the submitters to prepare their data for entry into the ICDC.

Request Initiation.

Submitters will provide a narrative describing their study and its scientific benefit for inclusion in the ICDC. At a minimum, the following points should be addressed:

1. Name/Identifier of Study
2. Grant ID and funding source (if applicable)
3. IACUC/IRB approval numbers (if applicable)
4. Scientific Point of Contact (Name, Phone, Email)
5. Data Manager Point of Contact (Name, Phone, Email)
6. Data access policy (choose one): Open-access – no-embargo, Controlled-access – no embargo, Open-access – embargo, Controlled-access - embargo
7. Cancer type(s) included in study
8. Number of subjects included in study
9. Data types included in study (check all that apply): Imaging, genomics, proteomics, immunology, clinical, other (specify)
10. Amount of data (in TB)
11. The overall scientific benefit of including this study in the ICDC prototype.
12. Any publications associated with this study, if any.
13. Time constraints on processing/loading/releasing the data
14. Data standards used, if any (e.g., SEND)
15. Anticipated budget needed to prepare data set for submission.

The following documentation should be attached to the submission:

1. Data Dictionary
2. Data Model/Schema diagram indicating how collected data relates to subjects, visits, samples, etc.

Submissions should be sent to: icdchelpdesk@mail.nih.gov.

Evaluation Criteria:
Data may be submitted to the ICDC that span the breadth of clinical, pathologic and -omics studies, aimed at advancing our collective knowledge of cancer in humans and dogs. While all data sets submitted must be associated with minimum requirements for common data elements (CDEs), a subset of data sets will be selected for ICDC assistance to transform the data into the optimum format needed to maximize their impact.

Studies will be evaluated on the clarity of the biological question being asked, the quality and quantity of the data obtained and the perceived value of the data set with regards to further explaining/treating cancer in humans and canines.

The ultimate goal of the DGAB is to choose studies that will aid the ICDC in creating a 'Rosetta stone' for mapping disease, data types, interventions, and response to human disease, so that the analysis of that data will identify cancers that are likely to have similar/different biological drivers and thus respond or not respond to a given therapy similarly between canines and humans. These studies are likely to elucidate novel and hopefully informative mechanisms. The drivers for the ICDC are to:

1. Create a high value resource that will aid investigators in primary and secondary analysis.
2. Allow investigators to compare their data with other canine and comparative datasets.
3. Be a trusted, high quality, highly accessed, highly shared resource for canine oncology datasets.
4. Encourage collaborative research and provide tools for the canine oncology community to collaborate on problems that are important in comparative oncology.

Prioritization:
Preference is given to data sets which can be fully public and do not require any application process or data use agreements. Controlled access data is not currently in scope for ICDC.

The DGAB will evaluate each study individually on its merits and will then assign a priority to the study relative to studies previously evaluated. In theory, this means that a particularly valuable study, as determined by the DGAB evaluation, could interrupt the resource allocation to existing studies "in-process" and be given a higher priority. In practice, it is likely that studies will be allocated resources based on submission timing. Using the advice provided by the DGAB, the NCI Senior Advisory committee makes the final decision on project inclusion and prioritization.