SOP4CWD Project: Data Use Agreement

Cornell Wildlife Health Lab
April 2020

This Data Use Agreement ("**Agreement**") is entered into as of _____ (MM/DD/YYYY; "**Effective Date**"), between Cornell University, a non-profit, educational institution with corporate powers under the laws of the State of New York, having an office at 373 Pine Tree Road Ithaca, NY 14850 (hereinafter called "**Cornell**"), and _____ (State Agency Name, hereinafter called "**Provider**"), having an office at _____ _____ (State Agency Address).

WHEREAS, **Cornell** wishes to receive and collect, and **Provider** wishes to provide, data for research titled "Surveillance Optimization Project for Chronic Wasting Disease (SOP4CWD)" (hereinafter called "**SOP4CWD**");

NOW, THEREFORE, **Cornell** and the **Provider** hereto agree to the following terms and conditions:

Definition of Terms

1. "**Primary Data**" shall mean the original data provided by the **Provider** to **Cornell**, on whatever media they exist.

2. "**Limited Data**" shall mean the **Primary Data** excluding any Sensitive or Personally Identifiable Information ("**SPII**") as identified by the **Provider** or **Cornell**. Examples include, but are not limited to the following:
    a. Names
    b. Postal address information, other than town or city, State
    c. Telephone numbers
    d. Fax numbers
    e. Electronic mail addresses
    f. Social Security numbers
    g. Account numbers
    h. License numbers
    i. Any other unique identifying number, characteristic, or code for any human

3. "**Standardized Data**" shall mean the **Limited Data** standardized and formatted by **Cornell Research Personnel** and used per the **Research Plan**.

4. "**Open Data**" shall mean the set of **Standardized Data** deposited under a persistent identifier (i.e. doi) in the **Public Repository**.

5. "**Public Repository**" shall mean the final digital repository in the Cornell eCommons Library.

6. "**Data**" shall mean the entire set of **Primary Data**, **Limited Data**, **Standardized Data**, and **Open Data**.

7. "**Research Plan**" shall mean the research project and deliverables described in **Exhibit A**.

8. "**Cornell Research Personnel**" shall mean the individuals listed in **Exhibit B** who will have unlimited access to the **Data**, prior to its archiving in the Public Repository.

9. "**Research Personnel**" shall mean the individuals listed in the **Research Plan** who will have limited access to and who may analyze and publish works using the **Standardized Data** and **Open Data**.

10. "**Personnel**" shall mean any **SOP4CWD** collaborator, including **Providers**, **Cornell Research Personnel**, and **Research Personnel**.

11. "**Public**" shall mean all World Wide Web users who will have unlimited digital access to **Open Data** after its archiving in the **Public Repository**.

12. "**Data Security Plan**" shall mean the plan for protecting the security of the **Data** provided under this **Agreement**, as detailed in **Exhibit C**.

13. "**Data Citation**" shall mean the formal citations for the public use of **Standardized Data** or **Open Data** in research products under this **Agreement**, as detailed in **Exhibit D**.

Term of this Agreement

14. The term of this **Agreement** shall be for one (1) year from the **Effective Date** with automatic annual renewals until the completion of activities outlined in the **Research Plan**.

15. Nothing contained in this **Agreement** shall be construed as an obligation to enter into any future agreement concerning the **Data**.

16. No right or options under any patent, copyright, trademark, mask works, or equivalent rights are granted by this **Agreement**.

Reception, Collection, and Disclosure of Data

17. **Cornell** shall, throughout the term of this **Agreement**:
    a. Receive, collect, and use **Data** only in accordance with the methods of the **Data Security Plan**, the aims of the **Research Plan**, and the **Provider**'s instructions and in good faith performance of its obligations under this **Agreement**, or where disclosure is required by law, in which case **Cornell** shall notify **Provider** of such and shall use its best efforts to limit the nature and scope of the required disclosure and shall only disclose the minimum amount of **Data** necessary to comply with the law;
    b. Receive, collect, and use **Data** in accordance with all applicable privacy, security and data protection laws, rules and regulations;
    c. Maintain the accuracy and integrity of **Data** it collects or uses, including the implementation of all reasonable physical, technical and administrative safeguards, as appropriate for the sensitivity of the information, to protect **Data** from loss, misuse, and unauthorized access, disclosure, alteration or destruction, including without limitation, implementation and enforcement of administrative,

technical and physical security policies and procedures, and training of all staff responsible for handling **Data** on the requirements set forth in the **Data Security Plan**;

d. Use encryption or equivalent measures in connection with any transfer, communication or remote access connectivity involving **Data**, as detailed in the **Data Security Plan**;
e. Notify **Provider** promptly of any unauthorized uses or disclosures of **Data** or any security incident involving **Data**, including without limitation any incident whereby **Cornell** reasonably believes to involve unauthorized access to or disclosure of **Data**, of which it becomes aware;
f. Cooperate with **Provider** to respond to any inquiries regarding practices related to the collection, use, and disclosure of **Data** in connection with this **Agreement** or any requests to access and correct **Data** in accordance with applicable law;
g. Obligate and ensure that any third-party agent or subcontractor to whom it provides the **Limited** and/or **Standardized Data** agrees in writing to the same restrictions and conditions that apply through this **Agreement** with respect to Use or Disclosure;
h. Disclose only **Standardized Data** to the **Research Personnel** for the execution of activities described in the **Research Plan**;
i. Disclose only **Open Data** to the **Public Repository** in the event that a journal requires open sharing for formal academic publication.

Ownership of Data

18. Ownership of **Data** will be retained by **Provider** and will be cited as such in all publications and research outputs listed in the **Research Plan**. For citations, see the **Publication** section below.

19. If, prior to formal archiving of the **Open Data**, the **Provider** requests that **Cornell** destroy the **Data**, **Cornell** will do so on whatever media it may exist within 15 working days of **Provider**'s written request to do so.

Access to the Data

20. Access to the **Primary Data** will be limited solely to the **Provider** and **Cornell Research** and will remain unhindered throughout the execution of the **Research Plan**.

21. Access to the **Limited Data** will be limited solely to the **Provider** and **Cornell Research Personnel** and will remain unhindered throughout the execution of the **Research Plan.**

22. Access to the **Standardized Data** will be limited to the **Provider**, **Cornell Research Personnel,** and **Research Personnel** and will remain unhindered throughout the execution of the **Research Plan**.

23. Access to the **Open Data** will be in accordance with the guidelines of the academic journal in which works resulting from the **Research Plan** are published.

Use of the Data

24. The **Data** will be used solely for activities described in the **Research Plan**.

25. Use of the **Data** will be consistent with **Cornell**'s policies regarding scientific integrity and ethics.

26. For **Providers**, sharing data with the **SOP4CWD** project does not limit its use to the **SOP4CWD** project. The **Provider** can simultaneously use their own **Primary Data** for internal projects or share the same data with other outside entities, in whatever manner they wish.

Data Security Procedures

27. The plan for protecting the security of the **Data** provided under this **Agreement**, as detailed in the **Data Security Plan**, will be followed by **Cornell** until such time as all originals and copies of the **Data** are returned to **Provider**, or in the case of the final use of **Open Data**, archived in the **Public Repository** should it be mandated for open science purposes by the academic journal.

Disposition of Data upon Completion of Research Project

28. Upon termination of this **Agreement**, **Cornell** shall in the case of the **Primary Data**, **Limited Data**, or **Standardized Data** return or destroy, or in the case of the **Open Data** archive in the **Public Repository**, all information provided by the **Provider** within 6 months of the completion of the project. Beyond 6 months, **Cornell** shall retain no copies of the **Data** received or collected from **Provider**, except those archived in the **Public Repository**. In the event that **Cornell** determines that returning or destroying the **Primary Data**, **Limited Data**, or **Standardized Data** is not feasible, **Cornell** shall provide to **Provider** notification of the conditions that make return or destruction infeasible. In such case, **Cornell** shall extend the protections of the **Agreement** to such **Primary Data**, **Limited Data**, or **Standardized Data** for as long as it is retained by **Cornell**. This provision shall survive termination of this **Agreement**.

29. **Cornell** may archive the **Open Data** in a **Public Repository** in order to satisfy academic journal requirements. In such an event, it is acknowledged and agreed that **Cornell** will cite each source and may grant a copyright license for the **Open Data** in the **Public Repository**. The **Open Data** will then be assigned a formal persistent identifier (e.g. doi), and the persistent identifier will be shared with the **Provider** and referenced in the final copies of publication and products.

Publication

30. It is acknowledged and agreed that **Cornell** and the **Research Personnel** may publish and/or present the results of analyses using the **Open Data** as detailed in the **Research Plan**. Project deliverables include (but are not limited to) publications, conference presentations, conference posters, webinars, medias, and other written or online materials.

31. **Cornell** will provide the **Provider** with a copy of any deliverable that utilizes the **Open Data**. For journal articles, a copy will be provided at the time it is submitted for review; for lectures or presentations, at the time the abstract is submitted and a copy of the full presentation at the time of the meeting; for online resources, when it is uploaded into the public domain.

32. **Cornell** will formally cite the **Open Data** in journal articles, presentations, or other online and written products using the following format:

    _____
    _____
    _____

    (Complete desired reference; examples are provided in **Exhibit D.**)

33. **Cornell** will include the following **Provider** co-authors in the list of authors of journal articles, presentations, or other online and written products resulting from use of the **Data**. Co-authors are those individuals that played a substantial intellectual role in the data collection, analysis, and writing of the final publication.

    _____
    _____
    _____

    (List desired **Provider** individuals to be included as co-authors.)

34. **Cornell** will include the following **Provider** individuals in the Acknowledgments section of journal articles, presentations, or other online and written products resulting from use of the **Data**. Acknowledgeable individuals or entities are those that played an instrumental role in data procurement or management, such as but not limited to the collection, processing, curation, QA/QC, standardization, or transfer of the **Data**.

    _____
    _____
    _____

    (List desired **Provider** individuals to be acknowledged.)

35. **Cornell** will include the following funding sources in the Acknowledgments section of journal articles, presentations, or other online and written products resulting from use of the **Data**.

    _____
    _____
    _____

    (List name(s) and award/grant number(s) for all entities that funded **Data** collection and curation activities up to and including transmission of **Data** to **Cornell**.)

36. **Cornell** will indicate in journal articles, presentations, or other online and written products resulting from use of the **Data** that the **Provider** and/or funder have not reviewed nor endorsed this work, and the results expressed herein do not necessarily reflect the positions of the [above-named **Provider**/funder], its directors, officers, managers, affiliates, nor its agents.

37. Neither **Cornell** nor the **Provider** shall make use of the other's name or that of any member of the other's staff for publicity or advertising purposes without the approvals defined above, or without another form of prior written approval of the other party.

Modifications to Research Plan

38. If changes in the **Research Plan** or computer environment will alter the **Research Plan** as part of this **Agreement**, **Cornell** shall provide to the **Provider** a revised **Research Plan**. No revisions shall be implemented until a fully executed amendment to this **Agreement** is in place. **Cornell** shall provide the

revised plan 30 days prior to the execution of any revisions. The **Provider** shall reply to the proposed revisions within 30 days or the revised plan will be implemented.

Violation of this Agreement

39. If the **Provider** determines that the **Agreement** may have been violated, the **Provider** will inform **Cornell** of the allegations in writing and will provide **Cornell** with an opportunity to respond in writing within 15 days.

40. Should any court of competent jurisdiction later consider any provisions of this **Agreement** to be invalid, illegal, or unenforceable, such provisions shall be considered severed from this **Agreement**. All other provisions, rights, and obligations shall continue without regard to the severed provision, provided that the remaining provisions of this **Agreement** are in accordance with the intentions of the parties.

41. The validity, interpretation and performance of this **Agreement** and any dispute connected herewith shall be governed and construed in accordance with the laws of the State of New York, USA.

Final Considerations

42. This Agreement contains the entire understanding between the parties with respect to the **Data** and the uses, security, and handling described herein.

43. This **Agreement** supersedes all prior understandings whether written or oral between the parties.

44. Any modification, amendment, or waiver of the terms of this **Agreement** shall require the written approval of authorized representatives of both parties.

45. Nothing contained in this **Agreement** shall be construed as an obligation to enter into any further agreement between the parties.

46. Both **Cornell** and the **Provider** represent that they have the right to enter into this **Agreement**.

47. Both **Cornell** and the **Provider** further represent that the terms of this **Agreement** are not inconsistent with other contractual obligations to which they are bound.

Signatures

The foregoing has been agreed to and accepted by authorized representatives of each party whose signatures appear below.

Cornell:

_____
Senior Grant and Contract Officer (Print)

_____
Senior Grant and Contract Officer (Signature)

_____
Date (MM/DD/YYYY)

Provider:

_____
Representative of Provider, Title (Print)

_____
Representative of Provider, Title (Signature)

_____
Date (MM/DD/YYYY)

EXHIBIT A. Research Plan

Overview

SOP4CWD is a collaborative effort to broadly improve the efficiency and effectiveness of CWD surveillance for white-tailed deer in the eastern United States. This effort merges CWD surveillance research and data science techniques to identify optimal sampling strategies for disease detection and management evaluation. Eight state wildlife agencies and USGS are the original project partners, but additional states are welcome and encouraged to join the project.

Project Components

1. Collaborative Management

Due to the breadth and scope of SOP4CWD and the number of collaborators involved, project management requires significant planning and effort. Activities include the development of data sharing agreements, tracking progress and maintaining project momentum, project documentation, coordination of meetings, and maintenance of the Open Science Framework (OSF) online site.

The Open Science Framework (OSF)

SOP4CWD uses the Open Science Framework for pre-registering, storing, managing, and sharing project-related resources. The OSF is a scientific project management system designed to facilitate documentation, preserve workflow, materials, and data, as well as provide a means to share those elements with collaborators. The OSF was developed by the Center for Open Science, a nonprofit supported by a variety of grants from federal agencies, private foundations, and commercial entities.

The OSF has several benefits beyond the more common cloud storage services such as Drive, Google Cloud, and Dropbox. Features include strong security, versioning and logging, built-in wiki documentation, and user access controls. Automatic record keeping of project developments will aid in satisfying journal research transparency requirements for peer-reviewed publications.

Collaborators are granted access to the general SOP4CWD project resources, as well as all appropriate components they for which they are a contributor or user. Contributors will not have access to primary data files from other contributors. The SOP4CWD OSF Project is being administered by the Cornell Wildlife Health Lab (CWHL) with Dr. Rachel Abbott and Nick Hollingshead managing user access and accounts.

Data Use Agreements

All state agencies participating in SOP4CWD will contribute CWD surveillance data and program information. Large datasets are critical for the statistical modeling and analyses that will be used to generate the intended project deliverables. To facilitate data sharing and ensure responsible data management, the Cornell Wildlife Health Lab will develop data use agreements (DUAs) with collaborating state agencies.

Data Use Policy

All project participants will adhere to a standardized data sharing, use, storage, and archival policy (“data use policy”), although some aspects of the policy may be customized to jurisdiction. The data use policy will cover general topics across all jurisdictions such as data security, short- and long-term storage, access, and sharing, as

well as topics that are agency-specific, such as permissions for use, desired co-authorships, citations and acknowledgements, funding information, and publication permissions.

The data use policy will cover data-related activities that occur after the data has been collected in each participating jurisdiction. The project will assume that all data contributed to the project has been collected in accordance with the policies and best practices of the source entity, but the SOP4CWD project itself does not govern those policies.

Multistate CWD Surveillance Database

The Cornell Wildlife Health Lab will develop and manage the multistate CWD surveillance database. The database will contain standardized CWD surveillance data provided by the collaborating state agencies, as well as supporting related program information.

Objective 1.1. Create a collaborative and supportive environment for collaborators

Tasks

- ☐ Develop an Open Science Framework (OSF) project site with appropriate data permissions and project guidelines tailored to the needs of specific collaborators and jurisdictions
- ☐ Maintain project management documentation and collaborator communications on OSF
- ☐ Identify lead collaborators for project components and communications with state agencies
- ☐ Communicate project progress regularly with collaborators

Objective 1.2. Facilitate data sharing

Tasks

- ☐ Develop data use agreements between the CWHL and collaborating state agencies
- ☐ Create a data use policy for all project participants
- ☐ Maintain high quality data management and security

Objective 1.3. Develop a multistate CWD surveillance database

Tasks

- ☐ Develop data contribution guidelines and instructions
- ☐ Assist state agencies in preparing, standardizing, and processing data for sharing
- ☐ Acquire CWD surveillance data, supplemental datasets, surveillance program information and other related resources
- ☐ Develop additional data resources (including geospatial datasets)
- ☐ Develop database of contributed data and program details
- ☐ Standardize surveillance data across states
- ☐ Maintain accurate documentation of data processing
- ☐ Develop synthetic datasets (“Open Data”) for future upload to the Public Repository

Collaborators

RACHEL ABBOTT (lead), Krysten Schuler, Nick Hollingshead, Brenda Hanley

2. Developing a Common Vocabulary & Framing the Issues

Each state agency is facing a multitude of challenges related to CWD surveillance and response. Some challenges may be unique to a state given its agency structure, policies, stakeholder culture, and the physical landscape.

However, states are facing many similar challenges and exploring parallel solutions. To facilitate communication between states and find commonality, the SOP4CWD collaborators will review the terminology and conceptual frameworks used by all states, with the goal of developing consistency in internal communication, project deliverables, and messaging.

Terms that were identified or used during the first annual SOP4CWD workshop include: leading edge, spark, new infection, core, buffer, prevalence, targeted removal, as well as many others. While some of these terms may be considered universal, others suffer from ambiguity, imprecision, misunderstanding, and/or are jargon. For instance, the term "buffer" is a geographic information system (GIS) software term referring to a spatial data processing method or analysis. A more appropriate related term, with more precision should be used to describe areas of increased concern for disease introduction due to nearby disease occurrences.

Objective 2.1. Develop a regionally consistent vocabulary for CWD surveillance and response

Tasks

- ☐ Develop a list of terms currently used by state agencies and the public
- ☐ Select appropriate terms with input from collaborators
- ☐ Define terms and clarify appropriate situational uses of terms
- ☐ Define a spatial and temporal scale that is meaningful to management
- ☐ Define the disease detection limit that is meaningful to management
- ☐ Define an acceptable probability of first detection that is meaningful to management
- ☐ Define the threshold prevalence where managers agree that the disease is "endemic"
- ☐ Provide document of terms (and their appropriate uses) to collaborators

Measures of success

- ☐ Glossary of terms and working definitions
- ☐ Participation and agreement from states on the use of the standardized terms and definitions
- ☐ A peer-reviewed publication

Collaborators

SONJA CHRISTENSEN (lead), Rachel Abbott, Nick Hollingshead, Brenda Hanley, Krysten Schuler

3. CWD Ecology & Epidemiology

State agencies have posed a range of questions related to the ecology and epidemiology of CWD. Finding answers to these questions is critical in identifying best management methods to prevent initial infections of CWD or curtail the spread in previously infected areas.

Objective 3.1. Understand disease progression over space and time

Tasks

- ☐ Explore historic spatiotemporal patterns of disease distribution and spread
- ☐ Use time series analysis of the growth rate and identify intrinsic factors that drive expansion or reduction of disease growth
- ☐ Develop a model that assesses disease growth by its demographic and spatial factors

Measures of success

- ☐ Peer-reviewed publication

Objective 3.2. Identification of "One and Dones" versus "Endemics"

Tasks

- ☐ Identify factors naturally associated with the "one-and-done" cases
- ☐ Identify factors naturally associated with an endemic disease
- ☐ Compare and contrast factors that naturally amplify or attenuate the disease
- ☐ Systematic review of factors that appear to give rise to "one and dones"
- ☐ Systematic review of factors that appear to give rise to "establishments"
- ☐ Develop a qualitative framework in which agencies can deem a CWD infection "one and done"

Measures of success

- ☐ Peer-reviewed publication

Collaborators

BRENDA HANLEY (lead), Michelle Carstensen (first author), Ani Belsare, Sonja Christensen, Nick Hollingshead, Dan Walsh, Dan Storm, David Walter, Jonathan Cook, Krysten Schuler

4. Preventing CWD Introduction

Preventing introduction of CWD to new areas is a critical concern for state agencies. CWD prevention measures are intended to inform the public of regulations and best practices to stop the introduction of contaminated materials that could transport prions and reduce the contact between potentially infected deer.

Agent modeling and other established approaches can inform state agencies about the potential effectiveness of different management strategies in preventing CWD infections from arising or becoming established in naïve habitats and metapopulations.

Objective 4.1. Understand the potential effects of disease prevention strategies

Tasks

- ☐ Assess the hypothetical effectiveness of certain management strategies (such as baiting, culling, extra seasons, or targeted removals) in preventing the spread of CWD into naïve areas
- ☐ Systematic review of management effectiveness given situational combinations of extrinsic and intrinsic disease drivers

Measures of success

- ☐ Peer-reviewed publication

Collaborators

ANI BELSARE (lead), David Williams, Dan Walsh, Brenda Hanley

5. Increasing Detection Power

In addition to "who," state agencies would benefit from data-driven recommendations on where, when, and how to sample. These recommendations could identify geographical changes in endemic areas (e.g., spread), assess changes in prevalence, identify high-risk locations susceptible to new infections, or discover novel outbreaks. Initial work to identify optimal surveillance recommendations will involve the spatial mapping of current information. At present, several states have used risk weightings from Wisconsin as proxy, but it is unclear if these weightings are applicable in other areas.

Sampling targets, timings, and locations will ultimately arise from information garnered from interface between a *Risk Probability Surface* (a spatial map of risk probabilities) and a *Disease Probability Surface* (a spatial map of disease prevalence probabilities). Sampling targets, timings, and locations will include a measure of confidence.

Objective 5.1. Identify the baseline distribution of the disease in different age/sex segments

Tasks

- ☐ Understand the annual prevalence in the spatial distribution of CWD positive and negative cases in each age/sex segment

Measures of success

- ☐ A map showing the change in CWD occurrence in the raw data in each age/sex segment across space and through time

Objective 5.2. Understand the inherent demographic risk in different age/sex segments

Tasks

- ☐ Compute the risk weightings for each age/sex segment in each partnering region using the Bayesian approach
- ☐ Compute the risk of individuals in each age/sex segment in each partnering region using the Agent Based Model
- ☐ Compare the types of risk measurements in each partnering region to identify the superior method
- ☐ Compare the risk weightings to Wisconsin (WI) to identify whether WI is a sufficient proxy for each partnering region
- ☐ Use the superior method to calculate max P(risk) for each age/sex segment in each partnering region
- ☐ Identify robustness of the risk predictions for each age/sex segment given bias in the data in each partnering region

Measures of success

- ☐ Systematic review of Risk/ABM with WI data as a proxy measure for other areas
- ☐ A map that incorporates regional and local data into supperior risk measurements

Objective 5.3. Understand probability of disease detection in different age/sex segments

Tasks

- ☐ Gather the geospatial dataset describing probability of detection across space and time
- ☐ Compute a detection probability surface for each age/sex segment
- ☐ Calculate the max P(detection) for each age/sex segment
- ☐ Identify robustness of detection predictions for each age/segment given bias in the data

Measures of success

- ☐ Online interactive map that incorporates regional and local data into the detection surface

Objective 5.4. Understand disease prevalence in different age/sex segments across space and time

Tasks

- ☐ Develop a region-wide contour map of disease prevalence
- ☐ Estimate prevalence in age/sex segments across the landscape and through time
- ☐ Pinpoint the locations where agencies are likely to discover novel infections or vectors of spread

- ☐ Identify the robustness of map predictions given bias in the sampling data

Measures of success

- ☐ Peer-reviewed publication
- ☐ A map of disease prevalence through space and time

Objective 5.5. Understand human-amplified risk across space and time

Tasks

- ☐ Assessment of the risk in established and novel areas when little or no sampling has been conducted
- ☐ Leverage processor, taxidermy, and carcass movement data into risk model

Objective 5.6. Overlay the spatial surfaces

Tasks

- ☐ Conduct geospatial analysis to identify high risk areas to discover novel infections

Measures of success

- ☐ A map of the overlap between probability surfaces across space and through time

Collaborators

NICK HOLLINGSHEAD (lead), Dan Walsh, Jon Cook, Ani Belsare, Krysten Schuler, and Brenda Hanley

6. Appropriate Resource Allocation

Agencies wish to identify the optimal distribution of sampling effort between the monitoring of established infections and the search for novel outbreaks. A combinatorial optimization algorithm will be adapted to best allocate surveillance resources among these managerial challenges.

Objective 6.1. Hone sampling strategies in novel areas

Tasks

- ☐ Ascertain the level of sampling needed to attain the desired prevalence information without oversampling novel areas
- ☐ Develop importance sampling given the synthetic probability distribution of prevalence to hone sampling in novel areas
- ☐ Identify robustness of sampling predictions in novel areas given biased data
- ☐ Assess whether mandatory sampling (census) alleviates sampling bias

Measures of success

- ☐ Software to see changes in probability of novel disease detection by sampling method
- ☐ Tool to know when enough information exists (e.g. to stop searching in a location)

Objective 6.2. Hone sampling strategies in established areas

Tasks

- ☐ Determine whether differing sampling type (census vs. sample) affects the probability of detection
- ☐ Determine whether differing sampling methods affect the probability of disease detection
- ☐ Identify how sample type (census vs. sample) affects our understanding of disease prevalence
- ☐ Identify how sample methods affect our understanding of disease prevalence

- Determine whether a rotating sampling scheme could affect cost and knowledge of the disease prevalence
- Identify whether combinations of sample types and/or methods can achieve satisfactory probability of detection and/or understanding of disease

Measures of success

- Software to see changes in probability of detection in established areas by sampling method

Objective 6.3. Allocating finite sampling resources between novel and established areas

Tasks

- Write a combinatorial optimization algorithm to find the best partition of resources among novel and established disease areas to maximize the probability of detection while minimizing the overall cost
- Explore the efficacy of this partition given a financial constraint (e.g. a testing cap)
- Provide equivalent surveillance options given pragmatic spatial and temporal logistical factors

Measures of success

- Software to identify the optimal allocation of resources and equivalent sampling alternatives
- Peer-reviewed publication

Collaborators

BRENDA HANLEY (lead), Ani Belsare, David Walter, David Williams, Nick Hollingshead, James Kelly, Dan Grove

7. Post-Detection Response

Many agencies have not yet detected CWD or are currently discovering novel CWD infections within their jurisdictions. Immediate detection response is crucial for planning activities that may aid in halting the initial spread to contain or eliminate the disease.

The regional data will be leveraged to identify which anthropogenic activities may be considered for prevention. As well, the data will be used to identify anthropogenic activities that function to attenuate or amplify the spread of disease after the first case has been found in a new area.

Objective 7.1. Use what we learned to halt or stall the spread of CWD

Tasks

- Use the agent-based model to identify strategies that prevent an initial infection
- Identify the “extent of influence” of a new detection
- Determine possible management alternatives for culling (e.g. where, when, and which age/sex segment) to halt or reduce CWD spread

Measures of success

- Systematic review of management scenarios that function to prevent an initial infection
- Tool that provides culling guidance in certain areas

Collaborators

Ani Belsare (lead), David Walter, David Williams, Nick Hollingshead, James Kelly, Dan Grove

8. Long-Term Monitoring Post-Establishment

Several collaborators joined the project with preexisting CWD infections in their jurisdictions. In some cases, initial responses have depleted resources to manage CWD in the long term. Agencies in such situations seek data driven guidance on how long-term sampling can influence our understanding of the disease.

Objective 8.1. Explore the best long-term strategies

Tasks

- ☐ Identify how differing long-term sampling strategies impact understanding of disease prevalence
- ☐ Identify the parameter space of CWD, and see if those parameter ranges change over time
- ☐ Identify risk of CWD for genotypes in established and new infection areas

Measures of success

- ☐ Tool that specifies allocation of surveillance resources in the long term

Collaborators

ANI BELSARE (lead), David Walter, Brenda Hanley, Nick Hollingshead, James Kelly, Dan Grove, David Williams

9. Understanding Risk at an Individual Deer Scale

Some agencies are funding specific components of the SOP4CWD but are amenable to sharing modeling outputs with project collaborators. For example, Tennessee is supporting the construction of a model designed to communicate the risk of infection of an individual deer directly to hunters. Given overwhelming historic public participation in CWD testing – participation of which is financially unsustainable for the state - TN sees that this model will benefit their management activities by 1) allaying hunter fear of consuming a CWD positive animal, 2) allowing hunters to gauge whether to pay to process their meat prior to receiving test results, and 3) to reduce voluntary submission and testing of deer that have low probabilities of disease.

Unlike the risk weightings that will be used for surveillance allocations, this model will leverage demographic and spatial data to provide a probability that a specific individual deer will return a positive test result.

Objective 9.1: Understand individual risk

Tasks

- ☐ Use logistic regression to identify the risk of an individual testing positive for CWD in established and newly infected areas
- ☐ Use cluster analysis to identify the risk of an individual testing positive for CWD in established and newly infected areas
- ☐ Compare individual risk predictions from logistic regression and clustering approached to ascertain the superior predictor
- ☐ Identify robustness of risk predictions in the superior predictor given biases in the surveillance data

Measures of success

- ☐ Peer-reviewed publication
- ☐ Software capable of providing a “measure of assurance” whether a hunter-harvested deer might test positive for CWD

Collaborators

SONJA CHRISTENSEN (lead), Brenda Hanley, David Walter, Ani Belsare, Nick Hollingshead, James Kelly, Dan Grove, Dan Walsh, David Walter, Chris Jennelle

10. Understanding Risk in Areas without Testing

Objective 10.0: Understand how proximity to confirmed cases can be used to estimate underlying prevalence landscape

Tasks

- ☐ Use lat/long points as a proxy for landscape combinations (specific configurations of habitat/demographic sets) to predict positive/negative disease outcomes
- ☐ Use half of the available lat/long/outcome data in each combinatorial set to train the predictive model specific to that combination
- ☐ Test the validity of the prediction in each combination using the remaining lat/long pos/neg data.
- ☐ Do this for TN data
- ☐ Repeat using all data for all states
- ☐ Compare
- ☐ Landscape variables will include SSURgO, those data obtainable via the national landcover databases, soil databases, hydrologic databases and topos
- ☐ Positive/negatives/lat/long will be attained from James and Dan, but we will strive to derive/standardize the model framework for use beyond TN.

Measures of success

- ☐ Peer-reviewed publication

Collaborators

DAVID WALTER (lead), Nick Hollingshead, James Kelly, Dan Grove, Brenda Hanley, Krysten Schuler, Rachel Abbott

EXHIBIT B. Cornell Research Personnel

Principal Investigator: Dr. Krysten Schuler, ks833@cornell.edu, (607) 253-3900, Cornell Wildlife Health Lab, Cornell University

Project Manager: Rachel Abbott, rca74@cornell.edu, (607) 253-3224, Cornell Wildlife Health Lab, Cornell University

Data Analyst: Nicholas Hollingshead, nah88@cornell.edu, (607) 235-1356, Cornell Wildlife Health Lab, Cornell University

Mathematical Modeler: Brenda Hanley, bjh262@cornell.edu, (208) 553-4475, Cornell Wildlife Health Lab, Cornell University

EXHIBIT C. Data Security Plan

The **Cornell Research Personnel** are committed to strong data security measures and to maintaining the privacy and integrity of **Data** shared for research purposes. The **Data Security Plan** describes the specific tools and methods that will be used for the **SOP4CWD** Project.

Limiting Sensitive or Personally Identifiable Information (SPII)

Minimizing **Cornell Research Personnel** contact with **SPII** can substantially reduce the potential harm caused by an inadvertent data breach. In addition, minimizing **Cornell Research Personnel** contact with **SPII** can substantially reduce the data security measures required during data transmission and storage. The **Cornell Research Personnel** will acquire from the **Provider** the minimum necessary **Primary Data** to complete the **Research Plan**, excluding as much **SPII** as possible. The **Cornell Research Personnel** will assist the **Provider** as necessary to accomplish this.

Data De-Identification

If **SPII** is included within the **Primary Data** from the **Provider**, the **Cornell Research Personnel** will de-identify individual records by removing **SPII** from the **Primary Data** to create the **Limited Data**. The **Primary Data** will be stored in a secure location accessible only by the **Cornell Research Personnel**. The **Limited Data** will be standardized and formatted by the **Cornell Research Personnel** to create the **Standardized Data** which will be used by the **Research Personnel** as detailed in the **Research Plan**.

Data Storage

The **SOP4CWD** Project will use the Center for Open Science ((*https://cos.io/our-products/osf/*) Open Science Framework (OSF) for all data storage and file sharing among **Personnel**. The OSF is a web application that supports scientific project management with features including cloud file storage, version control, user access control, and archiving. OSF Storage is the built-in file storage system in the OSF. OSF Storage uses Google Cloud for active storage and archival storage, and data at rest is encrypted. The storage location for the **SOP4CWD** OSF Project will be the United States. The **SOP4CWD** OSF Project will be setup and administered by the **Cornell Research Personnel**.

The OSF also allows storage by third party add-on services, such as Amazon S3, Box, Google Drive, OneDrive, and others. OSF links these third party storages to OSF users using WaterButler, an external data handling service (*https://github.com/CenterForOpenScience/waterbutler*). For **Providers** who do not wish for their data to be stored in OSF Storage, the use of a third-party storage add-on may be acceptable following approval by both the **Provider** and **Cornell Research Personnel**.

If, in order to complete the **Research Plan**, it is necessary to store **Data** using a different service, the service will be reviewed for security and must be approved by the **Provider**, **Cornell Research Personnel**, and **Research Personnel** on a case-by-case basis.

Data Transmission

Transmission of data to OSF Storage is encrypted with Secure Socket Layer (SSL). All file sharing among **SOP4CWD Personnel** will occur through the **SOP4CWD** OSF Project unless, in order to complete the **Research Plan**, it is necessary to transmit **Data** using a different method or service. If a method or service other than the OSF

SOP4CWD Project is used, the method or service must meet acceptable security levels and be approved by the **Provider** and **Cornell Research Personnel**. Data transfer by email or other insecure methods will not be permitted.

User Security

All **SOP4CWD Personnel** are required to use two-factor authentication for their OSF accounts.

SOP4CWD Personnel who have an ORCID account (https://orcid.org/) or are affiliated with an OSF partner institution may use these credentials to create and access their OSF account. However, **SOP4CWD Personnel** using one of these account setups must confirm that two-factor authentication is properly enabled through these account access methods

SOP4CWD Personnel who have created accounts directly in OSF or who do not have two-factor authentication through their institutional or ORCID login are required to enable two-factor authentication for their OSF account.

User Access Control

User access control will be administered in the OSF **SOP4CWD** Project by the **Cornell Research Personnel** in order to control access to the **Data**. In OSF, user access and permissions are assigned at the Project and Component levels. Components are conceptually equivalent to "sub-projects." A Component will be set up for each **Provider** for the purpose of transmitting **Data** to the **Cornell Research Personnel**. The **Primary Data** and derived **Limited Data** will be accessible only to the individual **Provider** and the **Cornell Research Personnel**. The **Standardized Data** will be stored in an OSF **SOP4CWD** Project Component accessible to all **SOP4CWD** Project **Personnel**. Access to **Data** is summarized in Table 1.

Table 1. Personnel access to data sets in the OSF SOP4CWD Project.

Data Set	Definition	*User Group Access*		
		Provider	**Cornell Research Personnel**	**Research Personnel**
Primary Data	The original data provided by the **Provider** to **Cornell Research Personnel**	Yes	Yes	No
Limited Data	The **Primary Data** excluding any **SPII** as identified by the **Provider** or **Cornell Research Personnel**	Yes	Yes	No
Standardized Data	The **Limited Data** standardized and formatted by **Cornell Research Personnel** and used per the **Research Plan**	Yes	Yes	Yes

Personal Device Security

All **SOP4CWD** Project **Personnel** are expected to maintain a secure local working environment including:

- The use of a password-locked screensaver and timeout lock
- Installation and maintenance of antivirus software
- Use of a firewall
- Maintaining up-to-date software
- Not installing or running programs from untrusted software
- Following institutional and agency IT department guidelines

EXHIBIT D. Formats for Formal Citation of Data Source

The **Provider** should determine how they would like their **Data** to be acknowledged and/or cited in publications, presentations, and other research products resulting from its use.

Open Data that may be posted in the **Public Repository**:

"FileName.txt". Agency Name. 20XX. *Title of dataset with dates*. Location of Provider.

In rare instances, **Providers** may request that their **Data** be omitted from formal archiving in the **Public Repository**. In such rare and exceptional circumstances, journals require that readers of the publication are able to maintain persistent and reasonable access to the **Standardized Data** via direct **Provider** request:

"FileName.txt". This dataset can be obtained upon request by contacting [Persistent Contact] in the [Agency Division] at [Agency Name]. 20XX. *Title of dataset with dates*. The [Persistent Contact] may be contacted by calling [persistent telephone] or by emailing [persistent email].