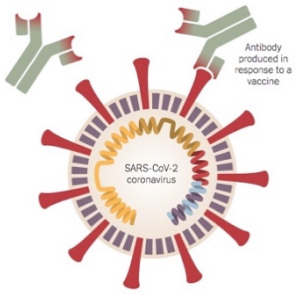
Nonhuman Primate Field Guide



Considerations for the use of primate models for SARS-CoV-2 treatments and vaccines

Prepared by the Coronavirus Vaccine & Treatment Evaluation Network, National Primate Research Centers

Supported by the Accelerating COVID-19 Therapeutic Interventions and Vaccines (ACTIV) Public-Private Partnership

Editors: Rudolf Bohm, Debra Bratt, Deborah H. Fuller, Nancy L. Haigwood, Sheri Hild, Mark Lewis, Jay Rappaport, Koen K.A. Van Rompay

December 2020

Credits: SARS-CoV-2 image from the New York Times
Photo: William F. Sutton, Oregon National Primate Research Center

Table of Contents

A. Introduction and Purpose of this Guide

SARS-CoV-2 burst onto the medical and scientific stage in late 2019, a novel coronavirus that has taken the world by storm. One of the most important activities in understanding virus infectivity, life cycle, and pathogenicity of a new virus involves the use of animal models. As investigators in academics and in the scientific industry grapple with developing effective therapies and vaccines, it is critical to understand how different animal models may provide valuable information to inform the science. This **Nonhuman Primate Field Guide** is designed as a companion to the **Small Animal Field Guide**. The purpose of the **Nonhuman Primate Field Guide** is to inform investigators of the current status of nonhuman primate (NHP) models, the availability of resources to perform studies, and to provide some details about experimental design and sampling to guide scientific planning.

Due to the highly infectious nature of SARS-CoV-2, this virus is classified as an agent requiring Biosafety Level 3 (BSL-3) for laboratory experiments. To perform experiments in animal models, including NHP models, Animal Biosafety Level 3 (ABSL-3) containment is required during the time that animals are exposed to SARS-CoV-2. At this time, regulations require that all animals, including NHPs, that are exposed to ABLS-3 agents must be euthanized at the end of the experiment, so all studies are terminal.

NHPs are often in short supply in the US, due to the specialized expertise and infrastructure required for breeding and the demand for animals to support ongoing research (NHP Evaluation and Analysis of Future Demand and Supply). Furthermore, there are only a few facilities that can perform research using NHP at ABSL3 and those have limited housing capacity. At this writing, all SARS-CoV-2 experiments that utilize NHP resources supported by the US National Institutes of Health (NIH) Office of Research Infrastructure and Programs (ORIP) must receive a programmatic priority by NIH (Notice of Limited Availability of Research NHPs). NIH expects users to submit the information requested in the COVID-19 NHP Study Information Portal so that the urgency of the proposed research and its timeline for potential impact on public health can be assessed by the NIH COVID-19 Expert Panel (the "Expert Panel"). The Expert Panel is composed of federal NIH staff with expertise in virology, immunology, therapeutic and vaccine development, NHP models, and other highly relevant subject matter areas. The Expert Panel will provide a programmatic recommendation for specific COVID research projects, which will be communicated to the National Primate Research Centers (NPRCs) or other ORIP-supported NHP facilities so the process of animal allocation and study initiation established at the Centers will be implemented appropriately.

If such studies are considered, or warranted, there are a number of critical issues and variables to consider in designing an effective study. This field guide is aimed at giving extensive information to guide all potential assessments, recognizing that each study may focus on a subset of assessments. It is focused on studies with SARS-CoV-2, but many of the areas of consideration will be similar for other infectious diseases that target the lung. These general considerations fall into the major categories of scientific, budgetary, regulatory, and logistical.

1. **Scientific.** It is critical that studies be designed using the most appropriate model for the pertinent scientific questions, and be powered so the correct number of animals are used to obtain a statistically significant outcome. There is growing expertise to guide these important questions, and this guide provides many of the details for design, statistical considerations, sampling, and some of the limitations that may help to prioritize regimen schedules.

2. **Budgetary.** NHP research requires a substantial budget. The study's size and complexity will drive costs, as well as many logistical issues noted in 4 below. All of the institutions that perform NHP work have experienced individuals that can work with you to develop a budget that accurately reflects the scientific goals and takes into consideration the personnel who will be needed to do the work.

3. **Regulatory**. All NHP studies must be peer-reviewed, and must adhere to USDA regulations, guidelines promulgated through Guide for the Care and Use of Laboratory Animals (https://grants.nih.gov/grants/olaw/guide-for-the-care-and-use-of-laboratory-animals.pdf) and PHS Policy (https://olaw.nih.gov/policies-laws/phs-policy.htm), and other regional and local regulations on research animal use. All research projects involving NHP must be approved by the local Institutional Animal Care and Use Committee (IACUC) and, at many institutions, the Institutional Biosafety Committee (IBC). As with budgeting, local experts at the research site will work with you to be sure that compliance with these regulations are met. There are also regulations that govern how samples must be handled (BSL-3 laboratories) and where and how they can be shipped.

4. **Logistical.** In order to initiate a study, the most critical logistical barrier is time, due to the need to budget accurately and adhere to regulatory requirements noted above, as well as time to identify a location that can schedule the experiment that you desire in a reasonable timeframe. Currently, COVID research has top priority at the NPRCs and potentially at other sites as well.

If an NHP study is warranted, there is much more detailed information available to investigators in the form of Standard Operating Procedures (SOPs) that will be made available for the design and conduct of studies that advance into testing. Ideally, individual studies will be comparable to other studies by the use of shared virus stocks and assays, and similar sampling procedures and time points, which will increase reproducibility and add to the growing base of knowledge.

B. NHP Models for SARS-CoV-2 Infection and Disease

Since late 2019, the SARS coronavirus SARS-CoV-2 has been circulating worldwide, resulting in the current pandemic, as disease and deaths mount due to COVID-19 disease. Multiple vaccines and therapies have been developed, and a number of these are currently in human clinical trials. Importantly, several animal models of COVID-19 have been developed that can provide

critical data on pathogenesis, efficacy, and correlates of protection. Following infection with SARS-CoV-2, Syrian hamsters, K18-ACE2 mice, and some aged African green monkeys (AGMs) develop moderate to severe pulmonary disease. Rhesus macaques (*M. mulatta*), cynomolgous macaques (*M. fascicularis*), and other macaques do not generally experience severe disease but do support robust viral replication and develop mild to moderate pulmonary disease with multiple inflammatory biomarkers, and are thus considered excellent models for testing therapeutics, vaccines and monoclonal antibodies. Notably, decades of research have led to an excellent understanding of the pathogenesis and immunity of coronaviruses and other infectious diseases in rhesus macaques. Scientists who are interested in performing experiments in NHP models are encouraged to read several review articles that have been published recently, in order to understand the potential and limitations of different models. The science continues to move rapidly as discoveries are made, so it is also wise to consult with contacts who can steer the specific study that an individual investigator may have in mind.

Some of the recent reviews are:

Muñoz-Fontela, C., Dowling, W.E., Funnell, S.G.P. *et al.* Animal models for COVID-19. *Nature* 586**,** 509–515 (2020). https://doi.org/10.1038/s41586-020-2787-6

Johansen, M.D., Irving, A., Montagutelli, X., Tate, M.D., Rudloff, I., Nold, M.F., Hansbro, N.G., Kim, R.Y., Donovan, C., Liu, G., Faiz, A., Short, K.R., Lyons, J.G., McCaughan, G.W., Gorrell, M.D., Cole, A., Moreno, C., Couteur, D., Hesselson, D., Triccas, J., Neely, G.G., Gamble, J.R., Simpson, S.J., Saunders, B.M., Oliver, B.G., Britton, W.J., Wark, P.A., Nold-Petry, C.A., Hansbro, P.M. Animal and translational models of SARS-CoV-2 infection and COVID-19. *Mucosal Immunol.* 2020 Nov;13(6):877-891. doi: 10.1038/s41385-020-00340-z. Epub 2020 Aug 20.PMID: 32820248

Hewitt, J.A., Lutz, C., Florence, W.C., Pitt, M.L.M., Rao, S., Rappaport, J., Haigwood, N.L. ACTIVating Resources for the COVID-19 Pandemic: In Vivo Models for Vaccines and Therapeutics. *Cell Host Microbe.* 2020 Nov 11;28(5):646-659. doi: 10.1016/j.chom.2020.09.016. Epub 2020 Oct 1.PMID: 33152279

The most important consideration in planning to use NHP models is whether the answer can be obtained with another species such as mouse, ferret or hamster. In many cases, small animal studies will be required prior to advancing to NHP, and some questions may not be addressible in NHP models.

With support from the Foundation for the National Institutes of Health (FNIH), Accelerating COVID-19 Therapeutic Interventions and Vaccines (ACTIV) Preclinical working group has developed an **Open Data Portal** where up-to-date summaries for both small and large animal models are available. This portal provides information about the extent of the interventions, disease manifestations, and pathology evaluated.

C. COVID-Related BSL-3 and ABSL-3 Resources for NHP Research

National Primate Research Centers

A summary of laboratory, clinical, and animal resources located at each of the seven NPRCs (California, Oregon, Southwest, Tulane, Washington, Wisconsin, Yerkes) is publicly available at the NPRC website, NHP Research Center Support Capabilities and Resources. Investigators can review the scientific and animal resources at different NPRCs and then contact them directly by email if they are interested in further details. The Coronavirus Coordinating Center at the Tulane NPRC will serve as a centralized facility that will work with all the NPRCs to assure the use of optimized protocols, best practices, and the harmonization of protocols across different Centers. The combined NPRC facilities and associated expertise comprise the Coronavirus Vaccines & Therapeutic Evaluation Network (CoVTEN). Coordination will occur through the interaction with an CoVTEN Operations Committee, with representation of each of the NPRCs as well as the Coronavirus Working Group, a consortium of of scientific and veterinary experts established by the NPRC system. In fact, the work of the Operations Committee and their subcommittees have contributed enormously to the writing of this field guide. The Coordinating Center at Tulane will further develop and host a Data Center that will permit the deposition and analysis of CoVID-19 related data from studies performed at all NPRCs.

Additional locations for BSL-3 and ABSL-3 work

Outside of the NPRCs, there are several institutions that can support COVID-19 and NHP research. The list below is not intended to be exhaustive, but rather to provide a starting point for those who are designing experiments (**bold** denotes that this location has a SARS-CoV-2 task order award in 2020). Websites denoted with an asterisk lead to home pages of investigators as contacts for those institutions.

a. ABSL3 Sites:
    i. Battelle Memorial Institute
    ii. Bioqual, Inc
    iii. Boston University National Emerging Infectious Diseases Laboratories (NEIDL)
    iv. Cornell University *
    v. George Mason University: National Center for Biodefense and Infectious Disease
    vi. Louisiana State A&M University
    vii. Medical College of Wisconsin: Center for Infectious Disease Research (CIDR)
    viii. Lovelace Biomedical
    ix. University of Chicago: Howard T. Ricketts Laboratory
    x. University of Georgia *
    xi. University of North Carolina *
    xii. University of Pennsylvania

xiii. University of Pittsburgh
xiv. University of Tennessee

b. BLS3 Sites:
   i. Geneva Foundation
   ii. Ragon Institute of MGH, MIT and Harvard
   iii. Rutgers University
   iv. Sanford Burnham Prebys
   v. Southern Research
   vi. University of Alabama
   vii. University of Louisville
   viii. University of Maryland *
   ix. University of Michigan
   x. University of Texas Medical Branch in Galveston, TX
   xi. Vanderbilt University

c. Other sites outside the US have capacity for this research. Again, some examples, but not an exhaustive list, include:
   i. Defense Science and Technology Laboratory (United Kingdom)
   ii. Erasmus University Rotterdam (Netherlands)
   iii. Fraunhofer Institute for Molecular Biology and Applied Ecology (Germany)
   iv. Public Health England (United Kingdom)

D. Nonhuman Primate Selection and Inclusion Guidelines

The purpose of this section is to provide information that will enhance standardization of pre-screening and selection of NHPs to be assigned to SARS-CoV-2 pathogenesis, vaccine and therapeutic studies performed by the ACTIV public-private partnership. Availability of NHP varies based on the species and other qualities that are required. Indian-origin rhesus macaques have been the preferred species for many NIH supported studies due to their availability in U.S. Government-funded research facilities. However, their availability is limited to breeding facilities within the US. Chinese origin rhesus may be an option, if available. Other possible choices are cynomolgus macaques, pigtailed macaques, and African Green monkeys.

Nonhuman Primate Models

The choice of species to use for NHP studies will depend upon the study design as well as animal availability. Data are emerging from ongoing studies that will assist in understanding the advantages and limitations of different species for specific experiments.

Rhesus Macaques bred in the US, either Indian-origin or Chinese-origin, are obtained from US breeding facilities. Their numbers are limited and mostly under control by the

US Government. China has breeding colonies of Chinese rhesus macaques; however, they have not imported any animals since February 2020 and do not plan to open their markets until late 2021 at the earliest.

Cynomolgus, Mauritian or Southeast Asian-origin Macaques are all imported from breeding colonies in Asia or Africa. They have been available for purchase from several importers, and arrangements can be made by contacting research sites for arrangement of delivery and quarantine at that site. Availability is limited and can take multiple months to obtain. China also has breeding colonies; however, all imports have been suspended until late 2021.

Pigtailed Macaques are extremely limited in availability. There are two US breeding colonies, one at the Washington NPRC and the other at Johns Hopkins University, and very few are available to the open market. While few are available to the open market, research studies where this species is determined to be the most appropriate model can be readily supported.

African Green Monkeys can be obtained from a US breeding facility at Wake Forest University, which is of limited size, or by using a US importer. These are bred or wild caught animals from several sites in the Caribbean. Again, their numbers are limited and it can take several months to receive them.

Nonhuman Primate Characteristics

Certain characteristics such as age and medical history may influence experimental outcomes, and investigators will also need to consider each of the following characteristics in choosing research subjects.

1. **Age.** While a majority of the early SARS-CoV2 pathogenesis studies were performed using mature and aging NHPs, a majority of the U.S. government-supported vaccine studies will be/have been performed using adult (> 3-year-old and < 15-year-old) Old World NHP; specifically, rhesus macaques of Indian or Chinese ancestry.
    - Except in rare cases, animals can be pair-housed even for infectious studies. Animals should have social housing as much as possible, and if possible, supported by the experimental design (e.g., inoculate pairs on the same day, etc).
    - Despite the fact that older humans appear to experience significantly higher mortality rates associated with SARS-CoV2 infection, the disease severity differs between studies. There are insufficient numbers of aged macaques available to enroll in vaccine or therapeutic studies, and many of these macaques are already designated for studies focused on maladies that affect aging humans (e.g., Alzheimer's, Parkinson's, etc.).

2. **Sex.**
    - Unless a specific study is evaluating the effects of COVID on the reproductive tract, balanced male/female sex ratios should not be mandated as the number of breeding age, Indian-ancestry, female macaques available for terminal studies is markedly limited. Assignment of some females and subsequent disaggregation by sex may allow researchers to detect some sex differences. However, the use of high numbers of reproductive-age female macaques will create sustainability issues within US breeding colonies and have a negative impact on future NHP research.
    - Ovariectomized and ovario-hysterectomized animals could be considered for assignment, but no studies have yet been performed to date on the possible effect these alterations may have on disease course.
    - Future vaccine/challenge studies may extend to utilizing pregnant females to discern the effect of the most promising vaccines and most relevant virus strains on the dam and fetus but are not recommended at this time due to resource limitation.
    - Castrated males could be considered for assignment, but no studies have yet been performed on the possible effect of gonadectomy on disease course.

3. **Ancestry/Genotype.**
    - Given the importance of Major Histocompatibility Complex Class I (MHC-I) genotypes for experimental studies of infectious disease agents such as SIV/SHIV, vaccine development, and transplantation research, it is suggested that all rhesus enrolled in COVID-19 studies be MHC-typed even though the role, if any, MHC plays in COVI19 infection/pathogenesis is unclear.
        - Determination of Mamu-A, -B, and -DRB haplotypes of each animal is recommended. Many NPRCs have this information or can perform this assay.
        - In the event that MHC typing is not performed pre-study, DNA from each animal should be banked so MHC type can eventually be determined.
        - MHC data available for ACTIV studies should be maintained in a centralized database that can be accessed by all consortium members (NHP Coordinating Center).
    - Indian-origin rhesus vs. Chinese-origin rhesus:
        - Utilize single nucleotide polymorphisms (SNPs) to determine Chinese admixture as part of the pre-screening process;
    - Pig-tailed and Cynomolgus Macaques
        - Determine proper genotyping to be performed: Pig-tailed (Mane) or Cynomolgus (Mafa).

4. **Medical history.** NHP should be screened prior to acquisition via medical records and diagnostic screening tests. See **Procurement, shipping and quarantine** in **Section F** below.
    - When looking at the medical records: a trained primate veterinarian should review all available health records. Animals should not have any chronic

problems such a persistent diarrhea, weight loss, behavioral issues that would preclude assignment, or other chronic or untreatable diseases.

- Animals should be serologically screened for a number of virus infections, e.g., SIV, SRV, STLV, Herpes B, Measles, and circulating members of the Coronavirus family. Any animal found to be SRV or STLV positive should be rejected and most facilities will not accept Herpes B+ animals. Anti-parasitic treatments are also standard practice at most NHP holding facilities.
- Age is important to consider as well as other high-risk co-morbidities.

5. **Behavioral history.** Behavioral history should be assessed for each candidate macaque to ensure the absence of behavioral abnormalities that would preclude them from being assigned or interfere with the study objectives.

6. **Experimental history.** Given the limited supply of NHPs, reuse and sharing of animals is encouraged when it does not negatively impact animal wellbeing or experimental goals.
   - Many of the early published SARS-CoV-2 studies have reused animals from previous studies to maximize the utility of the resource and reduce the number of new animals. These animals were determined to be healthy and had intact immune system. Use of these recycled animals has slowed the depletion of the current available supplies of research animals. If recycled animals are available, they should be prescreened for any preexisting SARS-CoV-2-specific immunity and receive a veterinary health check to determine if the animals can enter a new study. Use of animals with background immune responses is not recommended.
   - We recognize that previous research use may influence animal assignment. The veterinarians and researchers reviewing records prior to assignment are best positioned to determine whether or not previous manipulations may influence study outcomes. Criteria that may disqualify assignment include:
     - History of inguinal/axial lymph node biopsies if extensive biopsies are needed for the study
     - History of experimental vaccination utilizing viral vectors (e.g., cytomegalovirus (CMV), adenovirus-associated virus (AAV), rhesus rhadinovirus (RRV), adenovirus, etc.) that could interfere with other vaccines
     - History of antibody or plasma administration, which could impact resistance to Coronavirus infection or possible anti-antibody reactions;
     - History of ongoing or chronic steroid therapy that could impede immune responses to vaccines or dampen pathogenesis
     - History of repeated bronchoalveolar lavage that could enhance susceptibility to infection due to residual inflammation;
     - History of experimental infection with pulmonary pathogens that could interfere with SARS-CoV-2 infection
     - History of previous large blood collections that could limit additional collections during the study

- Prior treatment with immunomodulatory agents may have long lasting impacts beyond washout periods of the substance itself, and thus these animals are not recommended

E. Viral Stocks and Inoculation Procedures

1. **Standardized stocks.** There is a compelling rationale for the use of shared SARS-CoV-2 viral stocks that have been sequenced, titered, and validated for infectivity in vitro and in vivo prior to embarking upon infectivity studies in NHPs. Standardized and highly characterized viral stocks have been generated through a NIAID contract with BEI resources and will be provided for all CoVTEN studies run through the NPRCs. Emerging SARS-CoV-2 variants are continuously under evaluation for consideration as additional stocks to use in small animal or NHP models.

2. **Inoculation.** SARS-CoV-2 challenge strains may be inoculated via the conjunctival (100ul/eye), nasal (IN) (0.5mL/nostril), intratracheal (IT) (up to 4mL), or oral (1mL) routes, or in combination. Doses used have been up to 2.5 X 10^6 PFU (plaque-forming units) but generally, most SARS-CoV-2 challenge studies employ a combination of the IN and IT routes. Challenge dose should be discussed at the time of initiation to match with other studies to assure that it is appropriate to the experimental design.

    a. Consider volume of inoculum. Use of more than 0.5ml per nostril is not recommended. Volume IT is normally between 1-4 ml.

    b. Consider holding the animal in an upright position post inoculation for a period of time to ensure delivery of the inoculum.

3. **Verification of titer.** Ideally, virus diluted for inoculation will be back-titrated using a qualified and standardized focus-forming or plaque assay.

4. **Confirmation of infection.** Infection of animals for therapeutic studies and controls will be confirmed via validated and standardized qPCR methods implemented at each NPRC.

F. Study Design Considerations

Initiating a study in NHP can seem like a daunting undertaking, especially with the added pressure to work with an unfamiliar disease like COVID-19. In addition, the current cost of an individual NHP can be over $10,000, and most studies can easily cost over $500,000 to complete. Use of a small animal model should be considered as a first step to demonstrate proof-of-concept of the strategy before advancing the approach to a NHP model which, due to its similarity to humans, can provide valuable insight into optimal dosing and routes of delivery

for future human clinical trials as well as preliminary safety data. The purpose of this section is to guide researchers in their study design and point out considerations that are integral to study success.

1. **Project planning is crucial to study design.** One must understand the broad spectrum of services, regulatory approvals and cost to complete the desired research, as noted in the Introduction.

2. **Statistical considerations.** To decide the appropriate number of animals needed to power your study, consult with a biostatistician, making sure to include sufficient infected vs. mock-infected groups (see **Section G**, below).

3. **Additional screening**. After selecting animals based on your desired characteristics, the receiving institution should expect to receive a physical exam record and TB test result for each animal prior to shipment to your proposed research facility, if that is necessary. Close coordination with the proposed research facility is essential and will reduce many potential delays.

4. **Procurement, shipping and quarantine.** Procurement of animals at the desired location is typically arranged by the facility personnel, who should be able to arrange the animal screening, shipping and quarantine. Upon arrival to the NHP ABSL2/3, allow (30-90 days) for primates to undergo quarantine and to perform all CDC required screening, i.e., TB testing and to acclimate to their new surroundings. The stress of the transport and transition can impact hormones and the immune system; allowing for acclimatization will reduce this impact. Quarantine is extended if animals become ill or if diagnostic tests reveal health concerns; keep this in mind when procuring animals as it can affect study start dates or final available animal numbers.

5. **Prior to virus exposure.** Prior to SARS-CoV-2 exposure, NHP are maintained in at least Animal Biosafety Level 2 (ABSL2) due to risk of Macacine herpesvirus type 1 (MHV1), previously called herpes B.; all sampling will need to be acquired under anesthesia.

6. **Standard study designs.** for respiratory disease in macaques last between 5-16 days. The SARS-CoV-2 virus has been detected in samples from the upper and lower respiratory tracts, and in some peripheral and systemic samples. Studies can be shortened or lengthened depending on your experimental design. SARS-CoV-2 is normally detectable in the lungs until 8-10 days post challenge, and it can be detected in some animals for up to 28 days in the nasal cavity. However, productive infection, as detected by the presence of subgenomic (sg) RNA, has been detected in the lungs at least until day 7 and in the nasal passages by day 14, and may persist longer. Peak virus replication is seen between days 2-5 and peak lung pathology is seen between days 4 and 10.

7. **Biological samples and imaging for challenge studies.** These include body weight, bronchoalveolar lavage (BAL), blood, swabs, physical examination and cage side observation, lymph node biopsy, radiography and/or positron emission tomography-(PET-CT) scan, feces or rectal swabs. Details on recommended sampling procedures and timing are in **Section H**, below.

G. Statistical Analysis Plan

1. **Objectives for vaccine studies**

Main objective. To determine if individual COVID-19 vaccine products are efficacious in NHP models of SARS-CoV-2 infection. The exposure variables are vaccine product and dose. The measures for efficacy are in the domains of:

a) immunogenicity, including antibody responses, innate and inflammatory responses, and cellular immune responses; and
b) efficacy, including clinical scores, pathologic scores and viral loads.

Secondary objective. To determine which pre-infection immune variables, across all challenged animals, are most predictive of full protection against SARS-CoV-2 or attenuated infection. The exposure variables are post-vaccine, pre-infection measures of virus specific antibody and cellular immune responses while outcomes are clinical scores, pathologic scores and viral loads.

Secondary objective. To determine which clinical variables, across all challenged animals, are most predictive of post-infection boosting of memory immune responses. The exposure variables are post-vaccine, pre-infection immune variables as well clinical scores, pathologic scores and viral loads during infection. Outcome variables are post-infection measures of innate and inflammatory responses, virus specific antibody and cellular immune responses.

Secondary objective. To identify which arm(s) of the immune system mediate protection against infection, and or limit infection severity. Mathematical models will be used to couple non-linear immune and infection interactions according to various experimental stage. First, the dynamics of innate responses, virus-specific T cell and humoral responses will be modeled after vaccination including generation of cellular memory and antibody subsets. Next, the dynamics of viral load, clinical pathology, clinical severity, inflammatory markers and immune subsets will be captured with competing models during the post-infection challenge. The goal will be to identify components of the immune response that are most likely to be responsible for limiting viral load in nasal and lung compartments, clinical pathology in lungs and systemic inflammation; and to identify how infected cells boost immunologic memory.

2. **Statistical Methods.** These are examples, and specific analyses will depend upon the exact experimental design. Consultation with statisticians is highly recommended.

    a. **Antibody responses.** Antibody responses include the output from neutralization assays, pseudovirus assays, and binding assays (ELISA). The statistical analysis will be the **Mann-Whitney test** comparing continuous data on antibody titers between vaccination groups. Responses will be compared separately at various sampling timepoints and also according to an area under the curve measure.

    b. **Innate & inflammatory responses.** Innate and inflammatory responses include outcomes from cytokine/chemokine multiplex assays and immunophenotyping assays. The outcomes are individual cytokines and cell types of interest, and will be assessed using cell counts from flow cytometry tests as well as levels of individual analytes in plasma or other fluids. One set of hypotheses will be assessed through comparing cell counts at different times for different treatment groups. The statistical analysis will be Mann-Whitney test comparing ordinal count data.

    c. **Multiplex analyses**. More complex multi-dimensional data obtained from mesoscale discovery, bulk RNA-seq and single cell RNA-seq will require an initial data reduction step prior to statistical analysis such as principal component analysis. The final output will be assessed on an ordinal scale and will also be amenable to **Mann-Whitney testing**.

    d. **Cellular immune responses.** Like innate and inflammatory responses, frequency of specific cells as a result of vaccination or infection will be of interest and assessed though flow cytometry. In addition, ELISPOT will give rise to spot counts that represent the number of cells that produce interferon gamma and other key inflammatory mediators, as a result of stimulation. **Mann-Whitney testing** will be performed to compare the cell frequency between different vaccination groups and at different timepoints after infection.

    e. **Clinical laboratory test.** Outcomes from clinical laboratory tests will include measurements on analytes using clinical chemistry analyzers from blood serum and plasma. Of specific interest will be measures of renal, hepatic and bone marrow function, as well as systemic metrics of inflammation. The comparison of continuous variables will be performed using **Mann-Whitney test** or **unpaired student t-test** if the variable can be viewed as normally distributed.

    f. **Pathology.** Histopathologic scores (ordinal data) of tissues from treatment groups will be compared by using **Mann-Whitney test.**

    g. **Clinical scores.** Clinical scores of infection severity that are assessed with ordinal scores and metrics such as oxygen requirement and weight loss (continuous

variables) will be compared by using **Mann-Whitney test.** In NHPs, the clinical scores are typically low.

h. **Viral load.** Viral load will be deconstructed into several continuous kinetic features including peak viral load, area under the curve and duration of shedding. This continuous variable will be compared using the **Mann-Whitney test.** Log-transformation of peak viral loads often gives a normal distribution, so then a t-test can be done.

i. **Mathematical modeling.** The general approach for mathematical modeling is to compare dozens of ordinary differential equation-based models with differing implicit hypotheses for their ability to recapitulate longitudinal data using non-linear mixed effects methods. Models are rewarded for fit to data and penalized for undue complexity using Akaike Information Criteria. The model that meets these criteria has the highest likelihood of explaining the observed experimental data. Given the complex and comprehensive nature of the proposed data, we are planning for models of unprecedented complexity with equations intended to recapitulate observed viral load, T cell, B cell, innate cell, cytokine, inflammation level and pathology metrics.

3. **Power calculations for vaccine studies.** For vaccines, primary efficacy is defined upon detecting difference in peak viral loads in the vaccine and placebo groups, either in nasal and BAL for both subgenomic RNA and total RNA. We assume a normally distributed $\log_{10}$ peak viral loads with standard deviations in both vaccine and placebo groups, that is estimated from the data on rhesus macaques in Corbett et al. (2020). The power of detecting a difference on $\log_{10}$ peak viral load is given in Table 1. In addition, we show in Table 2 the minimum number of animals needed for achieving 0.8 power at given peak viral load reduction.

Table 1 Power for detecting reduction in peak log viral load with different sample sizes.

Reduction in $\log_{10}$ peak viral load		0.5	1	1.5	2
# Animals in each vaccine/placebo group	10	0.18	0.54	0.87	0.98
	20	0.32	0.85	0.99	1.00

Table 2 Minimum number of animals needed for each group for 0.8 power.

Reduction in $\log_{10}$ peak viral load	0.5	1	1.5	2
# Animals needed in each group for 0.8 power	67	18	9	6

4. **Power calculations for therapeutic agents.** For therapeutic agents, the primary efficacy is determined through comparison of clinical scores among two treatment groups. We

assume a normally distributed $\log_{10}$ clinical score with standard deviations 0.26, in both vaccine and placebo groups, that is estimated from the data on vehicle-treated rhesus macaques in Williamson et al. (2020). The power of detecting a difference in clinical score is given in Table 3. In addition, we show in Table 4 the minimum number of animals needed for achieving 0.8 power at given clinical score reduction.

Table 3 Power for detecting clinical score reduction with different sample sizes.

clinical score ratio in treatment vs. vehicle		0.8	0.6	0.4	0.2
# Animals in each group	10	0.12	0.42	0.88	1.00
	20	0.20	0.73	1.00	1.00

H. Sample Collection Procedures

1. **Anesthesia and physical examinations.** Use of different anesthetics may impact biological parameters including heart rate, blood pressure, body temperature, and respiratory rate. Physical examinations should be performed at a minimum frequency and when anesthetized for sample collection and imaging procedures.

2. **Telemetry.** Many natural processes and measurements can be affected by anesthesia, e.g., respiration rates, temperature, blood pressure. Telemetry implants may be considered prior to the start of the study if the expense, impact of surgical implantation, recovery time, and possible adverse outcomes do not outweigh the benefit of being able to collect these types of data.

3. **Baseline samples.** These should include serum, PBMC, plasma, BAL, and possible control biopsy tissues. Time for recovery is required if any surgery is performed prior to the challenge. Surgeries may include biopsies, telemetry implant, etc. Challenge day sampling should be limited and should not involve any collections from the lungs or upper respiratory tract as this can confound results.

4. **Clinical monitoring.** Animals must be closely monitored for health issues throughout the study period. Cage side observations (visual) twice daily (activity, food intake, respiratory rate/effort, discharges, etc.), making sure that trained personnel who understand the clinical scoring scale are performing the observations to reduce inter-operator error. A defined clinical scoring algorithm will help to provide consistency and should be agreed upon and approved by the IACUC prior to the initiation of infection.

5. **Imaging (Radiography, PET, CT, PET/CT).** Imaging of the lung is recommended, if available. However, published studies have seen only limited and diffuse lung lesions in most animals, and thus other measures of disease will be critical.

6. **Blood collection.** The determination of experimental blood samples collected must account for limitations on maximum volumes that can be collected from NHP. Blood volumes must be within IACUC parameters for the study locale. Recommend sampling on day 1 (optional) 3, 5, 7, 10, 14, 21 (may consider early euthanasia) and necropsy for tissue collection.

7. **Bronchoalveolar lavage and other swabs.** Bronchoalveolar lavage **(**BALs) can be performed more or less frequently depending on the goals of the study/research questions and the willingness to risk potential confounds in tissue readouts and disease in the lung. Limiting the volume used in the collection should be considered. Volumes can be smaller, <10ml, if checking for virus levels. Larger volumes may be required if isolation of cells, cytokine, chemokines or other proteins are needed. Considerations with the collection of bronchoalveolar lavage (BAL) should include:

    a. The use of a bronchoscope to collect BAL versus collection performed without a bronchoscope and potential impacts to the sample collected.

    b. BAL collections should be avoided on the day of challenge.

    c. The volume of fluid instilled for the BAL and consideration of potential impacts on the lung tissue and composition of BAL fluid samples, short and long term. The volume of BAL collected should also be assessed based on the type of analysis planned and the size of the animal. Smaller volumes are required to assess levels of virus in the lung and larger volumes may be needed to assess cells or chemokine/cytokine levels.

    d. BAL is not an innocuous procedure and has to be scheduled and carried out in a way to minimize adverse outcomes. The timing of BAL and inoculation will be something discussed with NPRC personnel during study design.

8. **Viral swab collection.** The selection of appropriate swabs (both size and material) are critical for collection of experimental viral swabs during SARS-CoV-2 research. Swabs should be synthetic tip (nylon, rayon or polyester) with a plastic or metal handle (not wood) the style of the swab may be flocked, spun or compressed. The collection of both nasal and oro-pharyngeal swabs should be considered and may be done daily. Saliva collection and nasal wash should be considered. Samples should be frozen immediately or treated with RNA preservative solution if planned to use for PCR testing.

9. **Necropsy samples/sharing** (See **Pathology, Section J).** Necropsy for tissue collection is usually scheduled between days 14-21 post inoculation depending on the goals of the research. Considerations should be made to do these earlier if tissues are required for virus detection or evaluation of histology at the peak of pathology.

I. Assays

1. Virologic assays

a. Tulane National Primate Research Center's "Virus Isolation and Characterization Core" may be contacted to perform sample viral loads and serum neutralization assays for all CoVTEN projects and may also be available for other locales, which may have their own testing sites.

b. Viral quantification should be performed using assays specific for viral genomes (the "genomic" assay) as well as subgenomic transcripts or specific for only the subgenomic mRNA. The subgenomic mRNA assay is indicative of active viral replication and has been used to show the efficacy of several vaccines/therapeutics in NHPs.

c. Viral replication (in vitro virus growth assay to measure replication-competent virus) in non-swab/BAL tissue samples should also be quantified when needed using validated and standardized procedures.

2. Considerations for antibody screening

a. Timepoints: Evaluation of binding and neutralizing antibody should include baseline levels and timepoints before and after each dose.

b. Samples: The primary sample type is serum or plasma. Mucosal antibody responses are generally measured in BAL but may also include analysis in nasal secretions and saliva.

c. Binding antibody assays include ELISA using Spike protein for capture antigen but can also employ other viral antigens depending on the vaccine, therapeutic or study design.

d. Neutralizing antibody assays can be performed in a variety of ways, with different advantages and disadvantages, as noted below.
   - Whole virus neutralizing antibody assays, requiring BSL-3 level containment, utilize whole virus. It is important to characterize the stock by sequencing, determining the viral stock titer ($TCID_{50}$/unit volume), as well as particle-to-infectivity ratio (as inactive particles may adversely affect the ability of antibodies to neutralize a specific viral stock). Furthermore, the specific virus used in such assays should be in agreement with the purpose of the assays and the questions being asked. For example, serum from animals immunized against the Washington strain Spike could be tested against the homologous virus, and potentially one or more new variants to determine the breadth of protection.

Such assays should also be conducted using parallel monoclonal antibody standards that are well characterized to assure the rigor and reproducibility and further aid in the interpretation of results.

- The pseudovirus neutralization assay can be performed at BSL-2 enhanced levels using purified antibodies and plasma, where the spike protein is packaged into heterologous virions.
- Another assay that can be performed in BSL-2 enhanced levels is a surrogate neutralization assay, where the antibody competes with the interaction of Spike protein bound to an indicator tag, with the human ACE-2 receptor bound to the surface of an assay plate.
- Plaque reduction neuralization titer assays (PRNT) and fluorescence neutralizing titer assays (NT; https://www.nature.com/articles/s41467-020-17892-0) utilize infectious virus and need to be performed in BSL-3 biocontainment. The PRNT assays has the advantage that it permits measurement of antibody-antigen interactions where the immunogen (e.g. Spike or RBD) is in its native conformation and in the context of a virus particle. Further, in the case of PRNT, different live virus variants can be tested without the need to re-engineer a fluorescent indicator virus, psuedovirus or surrogate neutralization assay.

3. **Cellular immunology.** Cell-mediated immune (CMI) responses are generally measured in the blood and a wide range of analyses are possible within a standard blood draw. For certain assays, Peripheral Blood Mononuclear Cell (PBMC) can be frozen and batch-analyzed together. It is also possible to analyze mucosal cellular immune responses cells collected from the BAL although generally the yield is low, thus limiting the types and number of assays that can be performed. Analysis of mucosal responses usually must be done on fresh cells.

    a. **ELISpot detection of virus-specific T cell responses.** Enzyme linked ImmunoSpot (ELISpot) is a straightforward technique to quantify virus-specific cellular immune responses by way of enumerating cells that produced a particular molecule (generally IFN$-\gamma$) in response to stimulation with peptides representing portion of the SARS-CoV-2 proteome. If used in SARS-CoV-2 projects at the NPRCs, ELISpot will be performed at the project site, rather than in a centralized core lab. An SOP has been established to harmonize both the preparation of cells for ELISpot as well as the ELISpot procedure itself as best as possible between the sites, but some potential for variation does exist. These include:
        - Variation due to automated spot quantification, which involves both a specific instrument for imaging and counting spots, and a technician to scrutinize the counting of the spots.
        - Inherent differences in instruments peripheral to but important in the assay, such as incubators used for overnight incubations.

    b. **Intracellular Cytokine Staining (ICS) detection of virus-specific T cell responses.** Intracellular cytokine staining (ICS) is based on the same concept as ELISpot, the

production of particular molecules (i.e., cytokines/chemokines) in response to stimulation with peptides representing portion of the SARS-CoV-2 proteome. Similar to ELISpot, ICS will be performed at the project site, rather than a centralized core lab. ICS is more complex than ELISpot and thus is more subject to inherent variables that can impact performance. These include:

- Inherent differences in instruments peripheral to but important in the assay, such as incubators used for overnight incubations. This variable can be quite important for ICS.
- Inherent differences in flow cytometers used to collect ICS data. All of the sites have high end instruments fully capable of performing ICS, but even small differences between instruments can impact data output.

4. **Innate and adaptive immunophenotyping.** There are a number of advantages to performing these analyses. Flow cytometry is a powerful means to obtain high-level resolution on the frequency and phenotype of key immune subsets. These panels can serve to complement multiplex-based assays to gain an understanding of the cellular source of cytokines and chemokines; they can also provide quantitative overviews of immune cells in tissues to complement immunohistochemistry and immunofluorescence-based strategies. Finally, immunophenotyping can provide real-time information on inflammatory dynamics following infection and intervention.

   a. **Panel choice for multiparameter flow cytometry.**

   - **Configuration of the instrument.** The choice of panels and suggested fluorophores will depend on configuration of the flow cytometer (lasers and filters) available at each primate center.

   - **Sample type.** Choice of specific markers in the innate, adaptive, and hybrid panels will depend on sample type. For instance, while CD45 is important as a positive gate for immune cells in the BAL this marker is not critical for samples from blood and lymph nodes.

   - **Viability Dye.** Need will vary by sample type. This step is optional for fresh whole blood samples, yet critical for BAL, and for enzymatically digested, cryopreserved samples

   - **Whole blood versus PBMCs.** Whole blood is an ideal sample to delineate granulocyte and neutrophil populations. Due to high frequency of neutrophils, 100-150 µl whole blood is sufficient.

   - **Enzymatically digested tissues.** Enzymatic digestion for liberation of immune cells from tissues (Collagenase, Trypsin, Dispase) impacts surface expression of

certain antigens. Consult with experts when running flow panels of enzymatically digested tissues.

- **Volume of whole blood for phenotyping infrequent populations.** At least 150-200 ul whole blood is recommended to obtain sufficient events for populations like plasmacytoid dendritic cells (pDCs) and myeloid dendritic cells (mDCs) which are infrequent in the periphery and tend to vary in frequency between animals.

b. Tools.

- One useful tool is the flow cytometry panel builder: https://www.thermofisher.com/us/en/home/life-science/cell-analysis/flow-cytometry/antibodies-for-flow-cytometry/flow-cytometry-panel-builder.html.html

- When considering use of a novel marker, refer to the Reactivity Database on the NHP Reagent Resource Center for information on cross-reactive clones. Consult with experts on need for Fc block. There is a considerable body of research describing the identification of optimized antibodies for use in identifying NHP surface markers for leukocytes.

5. **Soluble immunomodulators/biomarkers measurement in soluble samples.** This section described practices to standardize measurement of immunomodulators and biomarkers in fluids collected from NHPs. These can be used for pre-screening macaques for vaccine studies, or to detect changes in soluble immunomodulators (e.g. cytokines, chemokines, growth factors) following SARS- CoV-2 infection or SARS-CoV-2 vaccines/therapies studies. While different platforms have been identified capable of measuring soluble immunomodulators in bio-fluids, the Luminex platform has been selected based on the availability of this technique across the NPRCs, thus allowing for harmonization of safety operations procedures/protocols across NPRCs.

   a. Luminex platform overview. The Luminex platform is a bead-based multiplexed immunoassay system in a microplate format. Each assay can measure the level of soluble proteins in a homogenous liquid sample, offering simultaneous detection of approximately 40-60 different immunomodulators (multiplex analysis). The Luminex platform is particularly useful if the biological sample is limiting in that it reduces the volume of needed sample. In addition, the simultaneous measuring of multiple immunomodulators associated with a certain infection/disease (e.g., SARS-CoV-2/COVID-19), may offer a more accurate understanding of the multiple pathways affected by it, and of the effects of therapies and vaccines. Fluids can be measured including whole blood, plasma, serum, tissue homogenates/cell suspensions, and cerebrospinal and bronchoalveolar fluids among others.

b. Samples volume required and required biosafety level.*

- Serum (50 µl/well)/ BSL2
- Plasma (50 µl /well)/ BSL2
- Bronchoalveolar lavage fluid (50 µl/well) / BSL-3
- Cerebrospinal fluid (50 µl/well)/ BSL2
- Tissue homogenates/Cell suspensions (50 µl/well)/BSL-2 or BSL-3 depending on tissue.

**Note: Biosafety level for specimens may vary by NPRC or other specific location.*

c. Handling of samples and their effects on readout. Collection, processing and decontamination of tested samples should follow the direction of the SOPs. Readouts may be affected by cycles of freeze and thaw as this can cause protein degradation and this should be avoided. EDTA-coated/low binding tubes are preferred for this assay. In order to harmonize results across NPRCs, it is recommended that baseline samples (prior the start date of the study) for each animal be collected, so that experimental timepoints can be normalized to baseline values.

d. Luminex decontamination for transferring from BSL3 to BSL2 (if required). If transferring Luminex assay plates from BSL-3 to BSL-2 to acquire data, as may be the case for BAL specimens, a decontamination procedure needs to be performed after final washes and before adding the final buffer indicated in the commercial kit protocol. Such a protocol will need to be validated such that the virus is demonstrated to be inactivated, yet the detection of cytokines is not impaired. This would not be a concern normally for plasma or serum samples as it has not yet been possible to culture virus from these fluids. These studies can be performed at BSL-2 with the approval of the Institutional Biosafety Committee.

e. Luminex disadvantages and alternatives. While the Luminex platform offers many advantages, some immunomodulators (e.g., cytokines) remain under the limit of detection.** Alternatively, Mesoscale technology may be used instead to detect those markers. Another disadvantage is that for certain fluids such as BALs, sometimes there is bead aggregation that prevents the accurate counting of certain beads. Additionally, some NPRCs may require inactivation of samples prior to use in the Luminex assay, this may also reduce the activity of some analytes.

*** Note: Certain cytokines may not be successfully measured by the Luminex platform include IP-10/CXCL10, IL-1β and IL-10 and IFN-γ. Selection of specific kits should be done in close consultation with experts at the NPRCs.*

f. Mesoscale platform overview. As some of the cytokines may not be sufficiently detectable and quantifiable by Luminex, some investigators may wish to perform Mesoscale assay, a platform developed by Meso Scale Discovery. MSD has s large

collection of kits available validated for the detection of NHP cytokines, chemokines, and metabolic indicators. Kits are available to measure single analytes or multiple analytes in combination. A 24 plex panel is available for use with NHP that is well validated. The MSD platform utilizes a single detection method with 10 individual detection spots in each well on a 96 well plate. For this reason, fewer analytes can be measured within a single well relative to Luminex, and as such, more wells, plates, and larger sample volumes are needed, particularly for the larger panels such as the 24 plex panel. The MSD platform has higher sensitivity for certain cytokines, and is available at a subset of the NPRCs.

6. Genome-wide quantification of gene expression (RNA-seq)

a. Advantages of performing bulk RNA-seq assays.

- This assay provides genome-wide quantification of gene expression levels.
- By comparing gene expression levels between baseline samples and samples collected post-treatment or post-vaccination, the transcriptional host response to these events can be determined.
- Results from this assay can be integrated with other assays such as Luminex (cytokine/chemokine levels), immunophenotyping, and single cell RNA-seq assays to provide a comprehensive overview of the host immune response to vaccination or therapeutics.

b. Harmonization of RNA-seq assay.

- In order to directly compare results between studies, two NPRCs are designated sites for performing bulk RNA-seq assays if studies are performed there, and these services could be made available to other sites if needed.
- The Yerkes NPRC Genomics Core is charged with performing bulk RNA-seq assays on BAL cells samples.
- The Washington NPRC's NHP Genomics Core is charged with performing bulk RNA-seq assays on other sample types including whole blood, solid tissues, and isolated cells.
- Each genomic core has standardized SOPs for sample processing and data generation.
- Methods used for data processing and data analysis are harmonized between the two sites.

c. Bulk RNA-seq study design considerations.

- Sample size calculation is essential to ensure sufficient statistical power for detecting anticipated effects. Therefore, power and sample size estimations should be performed for each bulk RNA-seq study.
- Baseline samples – pre-vaccination or pre-treatment samples – are required.
- Appropriate time points for sample collection and sample-type are study-specific and dependent on the biological questions the study is designed to answer.

- As described in the procedures for sample collection (Section H), samples for bulk RNA-seq assays must be collected, preserved and stored using protocols that ensure the integrity of the RNA.

J. Pathology

Timing of necropsy is part of the study design. Tissues collected and tissue processing at necropsy should be determined prior to the start of the study. The number of animals that can be done per day is dependent on the number of tissues collected, collection methods, treatments, etc. The necropsy must be performed under ABSL-3 conditions, which increases the time to completion. NHPs are a valuable resource and maximal use of these animals is required. Sharing of tissues and tissue derived data can reduce the total numbers of animals used in research and can also advance the research of other scientists. Investigators are encouraged to contact the centralized databases from all of the NPRCs for specific tissue requests (Biomaterials Query System requires registration. To request access, contact support@nhprc.org).

a. The lungs are the main target organ in NHP models, and carefully consideration should be given as to how they are allocated for assays vs histopathology. Any lungs or lung lobes allocated for histopathology should be carefully infused with formalin according to validated methods since this is primarily an interstitial lung disease and lesions cannot be accurately assessed in collapsed lung tissue. The lung lobes primarily affected will vary depending on method of virus administration. Some lobes may have no or few lesions. Understanding how the virus administration route and technique influences virus distribution in the model is particularly important. Lung scoring strategies enable more consistent data for comparisons. The severity of the lung histology lesions in this model is highly variable and random therefore sufficient tissue needs to be evaluated to accurately assess the severity of the lesions in an individual animal

b. Soft palate, nasopharynx and oropharynx are important target organs and need to be harvested carefully at necropsy to maintain integrity and orientation.

c. The collection of the nasal turbinates is required in many protocols. Although there are a variety of techniques out there for successful collection, a technique should be used that keeps the turbinates intact.

d. Tissue fixation protocols need to balance health and safety requirements for virus neutralization with fixation times for tissue analysis using immunohistochemistry or in situ hybridization if required.

e. The lesions can be scored using a histology scoring system that is being developed. This scoring currently takes into account the following features:

i. Endothelialitis/ vasculitis (only in early timepoints approx. day 3)
ii. Hyaline membranes/ type 1 pneumocyte injury (only in early timepoints approx. day 3)
iii. Septal inflammation
iv. Septal fibrosis (generally day 7 onwards)
v. Type 2 hyperplasia
vi. Bronchial associated lymphoid tissue (BALT)
vii. Pleuritis
viii. Microvascular thrombosis (rare)
ix. Perivascular inflammation
x. Peribronchial and peribronchiolar inflammation