Open access • Posted Content • DOI:10.1101/819532

# Mapping the Host-Pathogen Space to Link Longitudinal and Cross-sectional Biomarker Data: Leptospira Infection in California Sea Lions (Zalophus californianus) as a Case Study — Source link

Katherine C. Prager, Michael G. Buhnerkempe, Michael G. Buhnerkempe, Denise J. Greig ...+9 more authors

**Institutions:** University of California, Los Angeles, Southern Illinois University School of Medicine, University of California, Berkeley, The Marine Mammal Center ...+6 more institutions

**Published on:** 25 Oct 2019 - bioRxiv (Cold Spring Harbor Laboratory)

**Topics:** Population

**Full Title:** Mapping the Host-Pathogen Space to Link Longitudinal and Cross-sectional Biomarker Data: *Leptospira* Infection in California Sea Lions (*Zalophus californianus*) as a Case Study

**Short Title:** Mapping the Host-Pathogen Space

K.C. Prager[1*], Michael G. Buhnerkempe[1,2], Denise J. Greig[3,4], Anthony J. Orr[5], Eric D. Jensen[6], Forrest Gomez[7], Renee L. Galloway[8], Qingzhong Wu[9], Frances M.D. Gulland[3,10], James O. Lloyd-Smith[1]

[1] Department of Ecology and Evolutionary Biology, University of California, Los Angeles, California, USA

[2] Department of Internal Medicine, Southern Illinois University School of Medicine, Springfield, Illinois, USA

[3]The Marine Mammal Center, Sausalito, California, USA

[4]California Academy of Sciences, San Francisco, California, USA

[5]Marine Mammal Laboratory, Alaska Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, Seattle, WA USA

[6]U.S. Navy Marine Mammal Program, Naval Information Warfare Center Pacific, San Diego, CA

[7]National Marine Mammal Foundation, San Diego, California, USA

[8]Centers for Disease Control and Prevention, Atlanta, Georgia, USA

[9]Hollings Marine Laboratory, National Ocean Service, Charleston, South Carolina, USA

[10]Karen Dryer Wildlife Health Center, University of California Davis, California, USA

24

25 * Corresponding author

26 E-mail: kcprager@ucla.edu (KP)

## Abstract

Confronted with the challenge of understanding population-level processes, disease ecologists and epidemiologists often simplify quantitative data into distinct physiological states (e.g. susceptible, exposed, infected, recovered). However, data defining these states often fall along a spectrum rather than into clear categories. Hence, the host-pathogen relationship is more accurately defined using quantitative data, often integrating multiple diagnostic measures, just as clinicians do to assess their patients. We use quantitative data on a bacterial infection (*Leptospira interrogans*) in California sea lions (*Zalophus californianus*) to improve both our individual-level and population-level understanding of this host-pathogen system. We create a "host-pathogen space" by mapping multiple biomarkers of infection (e.g. serum antibodies, pathogen DNA) and disease state (e.g. serum chemistry values) from 13 longitudinally sampled, severely ill individuals to visualize and characterize changes in these values through time. We describe a clear, unidirectional trajectory of disease and recovery within this host-pathogen space. Remarkably, this trajectory also captures the broad patterns in larger cross-sectional datasets of 1456 wild sea lions in all states of health. This mapping framework enables us to determine an individual's location in their time-course since initial infection, and to visualize the full range of clinical states and antibody responses induced by pathogen exposure, including severe acute disease, chronic subclinical infection, and recovery. We identify predictive relationships between biomarkers and outcomes such as survival and pathogen shedding, and in certain cases we can impute values for missing data, thus increasing the size of the useable dataset. Mapping the host-pathogen space and using quantitative biomarker data provides more nuanced approaches for understanding and

50 modeling disease dynamics in a system, yielding benefits for the clinician who needs to
51 triage patients and prevent transmission, and for the disease ecologist or epidemiologist
52 wishing to develop appropriate risk management strategies and assess health impacts on a
53 population scale.
54

55 ## Author Summary

56 A pathogen can cause a range of disease severity across different host individuals, and
57 these presentations change over the time-course from infection to recovery. These facts
58 complicate the work of epidemiologists and disease ecologists seeking to understand the
59 factors governing disease spread, often working with cross-sectional data. Recognizing
60 these facts also highlights the shortcomings of classical approaches to modeling infectious
61 disease, which typically rely on discrete and well-defined disease states. Here we show that
62 by analyzing multiple biomarkers of health and infection simultaneously, treating these
63 values as quantitative rather than binary indicators, and including a modest amount of
64 longitudinal sampling of hosts, we can create a map of the host-pathogen interaction that
65 shows the full spectrum of disease presentations and opens doors for new insights and
66 predictions. By accounting for individual variation and capturing changes through time
67 since infection, this mapping framework enables more robust interpretation of cross-
68 sectional data; e.g., to detect predictive relationships between biomarkers and key
69 outcomes such as survival, or to assess whether observed disease is associated with the
70 pathogen of interest. This approach can help epidemiologists, ecologists and clinicians to
71 better study and manage the many infectious diseases that exhibit complex relationships
72 with their hosts.

## Introduction

To gain insights into population-level trends, disease biomarker data are often reduced to binary form (e.g. presence/absence of a pathogen, antibodies or disease) for statistical analyses and parameterizing models of disease transmission. By contrast, to understand disease in an individual, the full quantitative range of available biomarker information is used to determine the precise clinical status of an individual, make treatment decisions, assess prognoses and limit transmission risk to others. While clinicians consider a clinically ill individual with a high or rising antibody titer as diagnostic of a current or recent infection [1], ecologists or epidemiologists typically classify individuals as exposed or not, and infected or not, based on a cut-off titer value [2], potentially discarding useful information contained in finer scale variations in titer magnitude. More detailed data on infection status and health can provide key information to both the clinician and ecologist that can help with accurate diagnosis (clinician) and effective system conceptualization, model construction and parameterization (ecologist). However, such data can be difficult to interpret, particularly for wildlife hosts, and all data types may not be available for each individual assessed. Hence, cases captured in clinical and surveillance data often do not fit neatly into distinct categories. Severely ill, recently infected individuals are easily identified (e.g. by high antibody titer, pathognomonic clinical signs, detection of pathogen), but are often just the tip of the iceberg. In reality, a variety of presentations may exist at each point along the timeline from infection to recovery, with individuals exhibiting a range of disease severity and antibody titers (e.g. from severely ill to apparently healthy and with very high to undetectable titers), and with both infected and uninfected individuals detected at any given combination of disease severity and antibody titer (Fig 1 and S1 Box).

Recently, efforts have been made to assess how biomarkers of disease and infection change relative to each other and over time, with the aim of identifying consistent patterns to improve our understanding of host-pathogen dynamics in human [3] [4], domestic animal [5], experimental [3], and wildlife systems [6] [7]. Longitudinal studies in which individuals are monitored through time provide key insights into how specific host-pathogen biomarkers, e.g. antibody titer, measures of disease severity, and pathogen load, change through the course of infection and recovery [7-10], with some studies showing how biomarker values may be associated with specific outcomes such as survival and transmission [3, 11]. In systems for which biomarkers show predictable temporal variation, quantitative data may provide information about an individual's stage in the infection and recovery process [3, 6, 12, 13], enhancing our understanding of population-level dynamics by providing key data for model structure and parameterization [6, 10, 12-19]. Assessment of quantitative values and multiple biomarkers can also elucidate individual within-host dynamics such as the outcome of an infection (infection chronicity, survival), how heterogeneity in antibody titer responses relates to clinical disease or symptoms, and probability of transmission to others [3, 11-13, 15, 20-22]. These findings can have direct implications on both the individual scale (e.g. triaging and treating patients, assessing prognosis and forward transmission risk) and the population scale (e.g. controlling transmission and hence outbreaks, predicting population dynamics, estimating incidence). These previous studies highlight the usefulness of including multiple data types, of understanding the nature of the relationship between multiple biomarkers of infection and disease, and of using quantitative data to better understand host-pathogen dynamics to make informed management decisions. However, although these studies explore facets of

119 this new frontier in infectious disease dynamics, none combine all facets within a single
120 study system, and few focus on disease in wildlife species.

121 We address this gap by linking longitudinal and cross-sectional data on multiple disease
122 measures from an unconventional study system: *Leptospira interrogans* serovar Pomona
123 (henceforth *"Leptospira*") infection in California sea lions (*Zalophus californianus*). This
124 system exhibits yearly, seasonal *Leptospira* outbreaks of varying magnitude, as reflected in
125 both clinical cases of *Leptospira* infection seen at marine mammal rescue and rehabilitation
126 centers [23] and in population-level serosurveys [24]. *Leptospira* is a good model for
127 examining complex manifestations of a host-pathogen relationship, as mammals infected
128 by pathogenic species within the genus *Leptospira* can exhibit a wide range of clinical
129 presentations, from fulminant clinical disease to silent infections, and while some hosts
130 may clear the infection quickly, others continue to shed the pathogen for months to years.
131 The dominant clinical signs of leptospirosis (the disease caused by *Leptospira* infection) in
132 California sea lions reflect the kidney damage inflicted by the bacteria, and clinically ill sea
133 lions present in varying stages of renal failure. The host-pathogen relationship for
134 pathogenic *Leptospira* spp. is conventionally attributed to specific *Leptospira* strain-host
135 species pairs and described dichotomously, as an acute and potentially fatal infection in
136 'accidental' host species, or as a chronic and predominantly subclinical infection in
137 'maintenance' host species [1, 25]. Yet, California sea lions show characteristics of both
138 accidental and maintenance hosts. During major outbreaks, roughly two-thirds of sea lions
139 stranding with clinical *Leptospira* infections die – typical of accidental hosts. However,
140 genetic evidence [26] and age-structured sero-epidemiology [24] suggest that *Leptospira* is
141 enzootic in the sea lion population, and furthermore subclinical chronic infections – typical

142 of maintenance hosts – occur in sea lions and are the possible mechanism for population-
143 level pathogen persistence from one outbreak to another [19, 27, 28].

144 Using longitudinal data on antibody titer, disease severity and pathogen shedding,
145 we track the temporal progression of *Leptospira* infections in California sea lions that
146 experienced either severe illness or subclinical infection. We use the relationship between
147 these different biomarkers to create a 'host-pathogen space' in which we track the
148 progression of known infected individuals through time and establish that they follow a
149 clear, unidirectional trajectory. Using this mapping approach, we then plot cross-sectional
150 data from a broader group of sea lions – either apparently healthy, wild-caught individuals,
151 or those stranding due to a broad range of health issues (i.e. not pre-selected for or against
152 leptospirosis), and use the patterns cast by the longitudinal data to interpret those in the
153 cross-sectional data. In human terms, the longitudinal data are akin to disease-specific
154 long-term monitoring of individual cases, whereas the cross-sectional data are akin to
155 prospective, random population surveillance, and unfiltered sampling of hospital patients,
156 and are therefore more representative of the overall population. We show that the
157 longitudinal data broadly capture the patterns in the cross-sectional data, suggesting
158 consistency in dynamics despite the greater set of individual presentations present in the
159 cross-sectional data. Our identification of a consistent trajectory through host-pathogen
160 space enables us to roughly situate cross-sectionally sampled individuals in their time-
161 course of infection, showing how our approach could elucidate disease dynamics in many
162 systems – from wildlife to humans – where most available data are cross-sectional. We also
163 find that patterns within the host-pathogen space provide population-level insights into the
164 range of disease experienced, duration of shedding, and associations between antibody

165 titer and infection status. This allows us to explore predictive relationships such as links
166 between disease severity and survival, and between antibody titer and shedding duration.
167 We also identify important differences between patterns in cross-sectional and
168 longitudinal data, and generate and test hypotheses regarding the source of these
169 differences, e.g., we identify renal disease from causes other than *Leptospira*.
170

171 # Results

172 ## Establishing a Host-Pathogen Trajectory with Longitudinal Data

173 We tracked the temporal progression of three important biomarkers of *Leptospira*
174 infection – anti-*Leptospira* serum antibody titer, renal compromise, and urinary leptospiral
175 DNA shedding – in 15 sea lions that were followed longitudinally from infection to clinical
176 recovery. Thirteen of these were initially severely ill and were followed for 6-12 weeks
177 (henceforth termed CLINICAL), and 2 never showed clinical signs and were followed for 3
178 years (termed SUBCLINICAL for subclinical, or SUB1 and SUB2 when referred to
179 individually; Table 1). Combined, data from the CLINICAL and SUBCLINICAL animals
180 enabled us to assess host-pathogen dynamics in animals exhibiting a range of initial clinical
181 disease. The CLINICAL animals are typical of what would be reported by hospitals or
182 rehabilitation centers for a given disease but may comprise only a small fraction of
183 infections experienced in a population. The majority of acute infections may involve no
184 evident disease, similar to the SUBCLINICAL animals, and would only be detected through
185 prospective surveillance efforts and unfiltered sampling of hospital cases.
186

**Table 1.** Description of the different data sets used in our study. Columns include the category of data "Group" and Sub-group", the "Sample Size" of unique individuals, the "Selection Criteria" used for inclusion, the "Additional Details" regarding individuals included, the "Day 0", i.e. the first day for which *Leptospira* infection related biomarkers were tracked in an individual, longitudinally monitored sea lion, and the "Length of Observation", i.e. the period of time over which data were collected.

193

| Group | Sub-group | Sample Size | Selection Criteria | Additional Details | Day 0 | Length of Observation |
|---|---|---|---|---|---|---|
| Longitudinal | CLINICAL | 13 | Presented initially with clinical signs of severe renal compromise consistent with leptospirosis*. | Survived infection and released into wild 6-12 weeks after admission to rehabilitation center. | First day anti-*Leptospira* antibody titer detected (0 – 18 days of admission) | 6 - 12 weeks |
| | SUB | 2 | Never showed clinical signs of leptospirosis. Admitted to rehabilitation center for treatment of other condition. Magnitude of the first detected anti-*Leptospira* antibody titers, and timing (October of a major *Leptospira* outbreak year in the wild sea lion population {Greig, 2005 #52}, suggest relatively recent *Leptospira* infection. | **SUB1:** No detectable anti-*Leptospira* antibodies initially, but seroconversion occurred (i.e. acquired anti-*Leptospira* antibodies) at some unknown point during rehabilitation, and in the absence of any observed clinical signs of leptospirosis. | First day anti-*Leptospira* antibody titer detected ($log_2$ titer=10 on 10/23/11, 15 months after admission). | 3 years |
| | | | | **SUB2:** Moderately high anti-*Leptospira* antibody titer at admission. No clinical signs of leptospirosis. Released into wild 3 weeks after admission. Readmitted 3 months after initial admission, still no clinical signs of leptospirosis. | First day anti-*Leptospira* antibody titer detected (($log_2$ titer=7 on 10/18/11, the day of admission) | |
| Cross-sectional | STRAND | 724 | All sea lions admitted to rehabilitation center for any cause, including leptospirosis. (i.e. not filtered by clinical signs) | | N/A | 1 day |
| | WILD | 730 | Apparently healthy, free-ranging sea lions. | | N/A | 1 day |

194
195 * Leptospirosis is the disease caused by infection with pathogenic species within the genus *Leptospira.*

196

We tracked changes in clinical disease using a 'renal index' that we derived from serum chemistry values (i.e. blood urea nitrogen, creatinine, sodium, chloride and phosphorus) associated with the compromised renal function seen in severe cases of leptospirosis [23]. Within the first 72 hours of admission to rehabilitation the severely ill animals that survived (CLINICAL) had high initial renal index values that ranged from 4.15 to 13.67, but they recovered rapidly with all scores declining into the healthy range within 15 to 61 days (median = 27 days; Fig 2A). By contrast, in the three years that they were monitored, we never detected serum chemistry evidence of renal compromise in the subclinical animals (SUB1 and SUB2; Fig 2B).

Antibody titers in individual CLINICAL sea lions exhibited simple exponential decay (Fig 2C), while the SUBCLINICAL animals exhibited a more complex pattern. Visual inspection of the SUBCLINICAL data suggested a biphasic pattern with an initial rapid phase consistent with that of the CLINICAL animals, followed by much slower decay (Fig 2D). Using a simple linear regression for each individual, we calculated half-life ($t_{1/2}$) estimates in CLINICAL sea lions that ranged from 6.4 to 29.4 days with a median $t_{1/2}$ of 17.1 days (Table 2). Using piecewise linear regression we calculated first phase $t_{1/2}$ values of 26.8 and 6.1 days for SUB1 and SUB2 respectively, and second phase values of 976 and 433 days (Table 2). The fact that first phase estimates for the two SUBCLINICAL animals fall within or close to the range seen for CLINICAL suggests consistency in early phase titer kinetics, regardless of the initial disease severity, and supports the assumption that our observations captured the end of the initial stage of infection for these SUBCLINICAL animals. Furthermore, our findings are qualitatively and quantitatively consistent with a

219 pattern of initial rapid antibody decay followed by a slower decay, as seen in other systems
220 where long-term antibody titer kinetics were tracked within individuals [29, 30].

221

222 **Table 2.** Antibody titer decline rates and half-life values in days with their corresponding
223 95% confidence intervals [95% CI]. Data are reported for each individual in the CLINICAL
224 and SUB datasets as well as for the first and second phase of titer decline observed for the
225 SUB animals. Rates for the CLINICAL animals are ordered from high to low with the median
226 decline and half-life values in ***bold italics***. The titer decline rate marked with an asterisk (*)
227 was not significantly different from zero.

228

| | | Antibody Titer | |
|---|---|---|---|
| | Animal ID | Decline Rate | Half-life [95% CI] |
| CLINICAL | 1 | -0.156 | 6.4 [6, 6.9] |
| | 2 | -0.149 | 6.7 [4.4, 14] |
| | 3 | -0.099 | 10.1 [8.7, 12] |
| | 4 | -0.075 | 13.4 [9.8, 21.4] |
| | 5 | -0.065 | 15.5 [11.1, 25.9] |
| | 6 | -0.061 | 16.3 [12, 25.3] |
| | ***7*** | ***-0.058*** | ***17.1 [11.5, 33.3]*** |
| | 8 | -0.058 | 17.2 [15, 20.2] |
| | 9 | -0.058* | 17.3 [7.3, infinity] |
| | 10 | -0.049 | 20.5 [13.6, 41.7] |
| | 11 | -0.046 | 21.7 [13.1, 64.2] |
| | 12 | -0.043 | 23.2 [19.2, 29.4] |
| | 13 | -0.034 | 29.4 [16.1, 168.4] |
| SUB | 1 - 1st Phase | -0.037 | 26.8 [21.2, 36.6] |
| | 1 - 2nd Phase | -0.001 | 975.9 [546.1, 4584.1] |
| | 2 - 1st Phase | -0.164 | 6.1 [4.1, 12] |
| | 2 - 2nd Phase | -0.002 | 433.4 [327.2, 641.4] |

229

To better understand the relationship between antibody titer and renal index, and to visualize how these biomarkers change relative to each other through time, we plotted the measures against each other to create a map of the host-pathogen space (Fig 3). With increasing time since infection, the CLINICAL animals followed a clear temporal trajectory, tracing a curved path starting in the high renal index and high titer space, dropping rapidly into the low renal index space with clinical recovery, and staying within the healthy range as antibody titers continued to drop (Fig 3A). In these CLINICAL animals, initial renal index values declined rapidly relative to antibody titers, so that only the earliest data points (<14 days since admission to rehabilitation) were found in the high titer, high renal index space. After 28 days, renal index scores leveled off within the healthy range and the temporal signal was dominated by antibody titer decline. By contrast, the SUBCLINICAL animals followed a straight path, always within the healthy range, as their antibody titers declined systematically throughout the 3 years that they were monitored (Fig 3A). All initial titers were high ($\log_2$ titer range CLINICAL=10-13, SUBCLINICAL=7-10) with variation among individuals observed. CLINICAL animals provided detailed information on initial changes in disease biomarkers, yet were released back into the wild within 6 – 12 weeks, providing no long-term data. In addition, as these animals stranded some unknown number of days after initial infection, the 'upswing' of antibody titers and clinical disease were not captured. Conversely, the SUBCLINICAL animals were followed for 3 years, providing important long-term biomarker data, but little on their initial dynamics (Table 1). Ultimately the CLINICAL and SUBCLINICAL paths overlapped, demonstrating convergence of the two trajectories and, potentially, similar long-term dynamics.

252 We used PCR to detect *Leptospira* DNA shed in the urine – a measure of current
253 infection and potential transmission risk to others – and added pathogen shedding data to
254 the map of the host-pathogen space. Addition of this third disease biomarker revealed that
255 many animals continued to shed despite a rapid return to healthy renal function and
256 systematic antibody titer decline (Fig 3B). All CLINICAL sea lions tested positive at least
257 once in the first 38 days, most (11/13) continued shedding despite concurrent antibody
258 titer decline and clinical recovery, and most (10/13) were still shedding at the last
259 sampling point 4 – 12 weeks after initial admission (Fig 3B; also see [28]). Subclinical
260 shedding of at least 8 weeks was detected in SUB1 [27], indicating that initial severe clinical
261 disease is not a necessary condition for shedding of this duration. Shedding was never
262 detected in SUB2, but the first urine testing date was 38 weeks after first detection of
263 serum antibodies.
264 Altogether, our findings suggest that antibody titers act as a rough clock indicating
265 time since exposure to the pathogen, with data on disease severity and pathogen shedding
266 improving the temporal resolution of the host-pathogen trajectory.
267
268 **Using the Host-Pathogen Trajectory to Interpret Cross-sectional Data**
269 Having established a temporal host-pathogen trajectory, we used our mapping
270 approach to maximize the information gained from individuals observed only once. These
271 cross-sectional data were from stranded (STRAND) and wild-caught, free-ranging (WILD)
272 California sea lions (Table 1). When mapped, STRAND data fell along the trajectory
273 mapped by the longitudinal data, but with greater variation, i.e., they cut a broader path
274 through the host-pathogen space, and their map contained some outliers (Fig 3C). The

STRAND data contained a wider range of renal index scores (-3.6 – 17.8) and a higher maximum antibody titer ($log_2$ titer = 15) than did the longitudinal data (renal index = -1.0 – 13.7; maximum $log_2$ antibody titer = 13; Fig 3A-C), suggesting that STRAND data captured a greater overall range of sea lion-*Leptospira* host-pathogen dynamics than the smaller dataset of longitudinally followed animals. The larger size (50-fold larger than CLINICAL) and broader selection conditions (i.e. including animals so ill from leptospirosis they died quickly, as well as those compromised for other reasons) of the STRAND dataset could explain this difference. By contrast, and in keeping with our assessment of apparent health at capture, the WILD animals chiefly occupied the space defined by the SUBCLINICAL and the recovered CLINICAL animals (Fig 3D). Notably, the WILD and STRAND datasets both differed from the SUBCLINICAL and CLINICAL in that they contained substantial numbers of seronegative animals.

We analyzed the distribution of renal index scores in each group, using antibody titer levels to standardize for time since infection, in order to assess (1) whether each group had a unique renal index profile or whether the SUBCLINICAL, WILD and CLINICAL groups were merely opposite extremes within the range seen in the STRAND animals, with WILD and SUBCLINICAL at one extreme and CLINICAL at the other, and (2) whether all groups converged to the same point with time since infection (Fig 4; Table 3). Renal index distributions of WILD and SUBCLINICAL animals never differed significantly. Those of the WILD and CLINICAL animals differed significantly at each antibody titer level assessed, yet the difference between their mean renal index values decreased as titers decreased, i.e. they were converging with time since infection. While some STRAND animals exhibited markedly greater renal disease than WILD animals, at all titer levels, there is also

298 substantial overlap between these groups, suggesting that the WILD animals are similar to
299 the majority of STRAND animals not suffering from clinical leptospirosis. Altogether, for all
300 datasets and regardless of the starting point in the trajectory (i.e. the renal index value at
301 the highest titers), as antibody titers declined, so did mean renal index scores and the
302 trajectories of each of the different datasets converged towards the healthy range.
303
304 **Table 3.** Mean renal index scores and sample sizes (n) by antibody titer group for initially
305 clinical (CLINICAL), initially subclinical (SUB), stranded (STRAND) and wild-caught (WILD)
306 animals. Titer group "0" contains all seronegative animals, titer group 0* contains only non-
307 shedding (i.e. urine PCR negative) seronegative animals. P-values are for two-sided
308 bootstrap Kolmogorov-Smirnov (KS) test comparisons between animal groups of renal
309 index distributions for a given titer. P-values for WILD 11+ v. WILD are for one-sided
310 bootstrap KS test comparisons within the WILD dataset, with the null hypothesis that the
311 renal index distribution for the 11+ titer group will be greater. P-values for STRAND 0* v.
312 STRAND are for two-sided bootstrap KS test comparisons within the STRAND dataset
313 comparing renal index distributions of seronegative, non-shedding animals with the other
314 titer groups.
315

316

317

| | | Antibody Titer Group | | | | | |
|---|---|---|---|---|---|---|---|
| | | 11+ | 9-10 | 6-8 | 1-5 | 0 | 0* |
| Mean (n) | WILD | 0.65 (8) | -0.18 (8) | -0.33 (11) | -0.27 (22) | -0.32 (683) | -0.28 (562) |
| | SUB | - | -0.49 (1) | -0.11 (1) | -0.42 (27) | - | - |
| | CLINICAL | 6.6 (18) | 1.47 (26) | 0.57 (22) | - | - | - |
| | STRAND | 6.44 (181) | 3.66 (76) | 0.82 (22) | 1.56 (45) | 0.3 (473) | 0.83 (78) |
| KS test p-value | CLINICAL v. WILD | 0.001 | 0.005 | 0.002 | - | - | - |
| | SUB v. WILD | - | - | - | 0.07 | - | - |
| | STRAND v. WILD | <0.001 | <0.001 | 0.04 | <0.001 | <0.001 | <0.001 |
| | STRAND v. CLINICAL | 0.56 | 0.002 | 0.04 | - | - | - |
| | STRAND v. SUB | - | - | - | <0.001 | - | - |
| | WILD 11+ v. WILD | | 0.11 | 0.04 | 0.02 | 0.04 | - |
| | STRAND 0* v. STRAND | <0.001 | <0.001 | 0.66 | 0.39 | - | - |
| | WILD 0* v. WILD | 0.07 | 0.38 | 0.66 | 0.06 | - | - |

318

319 As with the longitudinally sampled animals, leptospiral DNA was detected in both
320 STRAND and WILD animals for a wide range of antibody titer and renal index values (Fig
321 3C&D). However, unlike the longitudinal groups, the cross-sectional data also included
322 seronegative animals (i.e. no detectable anti-*Leptospira* antibodies; Fig 3C&D; plotted
323 above $\log_2$ titer of 0). These animals presented with a range of renal index scores and,
324 intriguingly, included animals shedding leptospiral DNA (Fig 3C and 3D; see section
325 'Antibody Titer Kinetics and Shedding Duration' for further discussion of these animals).
326 The broad congruence of the cross-sectional datasets with the longitudinally
327 collected data corroborates the assumption that the CLINICAL and SUBCLINICAL animals
328 jointly define the course of infection in this space and establishes that cross-sectional data
329 can be interpreted within this temporal framework.

**Tracking the Distribution of Disease Severity.** Defining the mean and range of pathogen-induced disease severity at different times since infection enhances our ability to interpret confusing host presentations (Fig 2) and hence understand disease dynamics in a system. However, biases in data sources must be considered when interpreting these data. In our study, when all data are combined, we see that initial disease severity (i.e. renal indices when animals have high antibody titers) ranges from healthy to severely ill (Fig 3). However, by design, initial renal index values in the CLINICAL animals captured only the upper range of disease severity, while the SUBCLINICAL animals occupied only the lower healthy range. WILD animals were sampled only if apparently healthy, and their renal index scores reflected this initial assessment, mostly occupying only the healthy range even during the presumed initial stage of infection (Fig 3 and Fig 4 titer level 11+). By contrast, STRAND data, which were collected without applying selection criteria to candidate animals, showed a wide range of initial disease severity and appear to knit together the various subset datasets to which specific selection criteria were applied (e.g. WILD, CLINICAL, SUBCLINICAL; Fig 3). Of note, although most of the seropositive WILD animals fall within the healthy range of renal index values, at the highest titer values a few exceed the healthy range (Fig 3D), and the mean renal index score of those individuals at this highest titer level is greater than those of the other levels (Fig 4; Table 3), suggesting that these animals can experience some degree of initial renal compromise from which they recover.

In many systems, disentangling disease caused by the pathogen of interest versus disease from another etiology can be difficult. In our study, while STRAND data capture the

353 full spectrum of disease and follow the trajectory defined by the longitudinal data, this
354 trajectory is shifted up the renal index axis and there are some obvious outliers (e.g. mid-
355 low antibody titer, high renal index individuals; Fig 3C). STRAND renal index score
356 distributions were significantly higher than those of almost all other datasets (i.e.
357 CLINICAL, SUBCLINICAL, and WILD) at all antibody titer levels (with the single exception
358 that renal index distributions for the highest-titer groups of STRAND and CLINICAL were
359 indistinguishable (Fig 4; Table 3)). We hypothesize that this upward shift in STRAND renal
360 index score is due to individuals experiencing renal compromise from causes other than
361 leptospirosis and that overall STRAND host-pathogen dynamics are consistent with those
362 described by the longitudinal data, i.e., animals are recovered clinically from leptospirosis
363 by the time their $\log_2$ antibody titers have declined below 9 (Fig 3A-C & Fig 4).
364 To test this idea, we analyzed the group of seronegative, non-shedding STRAND
365 animals that were presumably never infected and never exposed. Any renal compromise
366 observed in this group would be from a cause other than leptospirosis and the range of
367 their renal index values provides a reference against which to compare currently or
368 previously infected sea lions. We found that the renal index distribution of these
369 seronegative, non-shedders in the STRAND dataset (denoted 0* in Table 3) was not
370 significantly different from those of the mid and lower antibody titer STRAND groups (1-5,
371 6-8; (Fig 4; Table 3). This suggests that outliers found on the map – mid-low antibody titer
372 with high renal index (Fig 3C) – which, according to the host-pathogen trajectory described
373 in Fig 3A-B, should have fully recovered from *Leptospira*-induced renal compromise, are
374 equivalent to the seronegative non-shedding STRAND animals experiencing disease from
375 another etiology. Similar analyses of the WILD animals showed no significant difference

between seronegative, non-shedding animals and animals with titers, further supporting our assumption of apparent health of this group.

**Predicting Survival.** In hospital and rehabilitation settings, determining probability of survival can be vital for patient triage and efficient allocation of resources. Using data on renal index scores that are readily available at admission, we found a significant negative relationship between these scores and survival of animals suspected of having leptospirosis (Fig 5; OR = 0.64, 95% CI = 0.53 – 0.78, $p<0.001$). This relationship is not only informative for guiding management in a clinical setting, but the absence of high renal index values in the WILD animals suggests that *Leptospira*-associated mortality in the apparently healthy animals selected for sampling is likely low.

**Antibody Titer Kinetics and Shedding Duration.** Estimating an individual's time since infection aids in assessing infection incidence [6, 16] [4] and, in combination with data on shedding status, can enable estimation of the duration of infectivity – a value which is notoriously difficult to determine in wildlife where repeated sampling of individuals is rare. To approach this problem, we explore the hypothesis that for our system there is a single dominant pattern of antibody titer decline, regardless of initial disease severity, such that titer acts as a rough measure of time since infection. We begin by noting that animals shedding leptospiral DNA exhibited antibody titers ranging from very high to seronegative (Fig 3B-D; S2 Table). Under the working hypothesis that all animals experience similar antibody titer kinetics, low titer and seronegative shedders would be chronic shedders.

To test this idea we considered the epidemiological context of our data: during our study period, outbreaks occurred in 2008 and 2011 (Fig 6A). If antibody titer decline initially occurs rapidly as seen with the CLINICAL animals and then quite slowly as seen with the SUBCLINICAL animals (Fig 2C&D), we should see relatively few low titer animals during an outbreak year, and relatively more with each year until the next major outbreak. Similarly, observation of seronegative shedders should occur after outbreaks as titers dip below detection.

We found support for our hypothesis when we plotted data from the combined WILD and STRAND datasets. The proportion of low titer ($log_2$ titer 1-5) animals increased following major outbreaks, particularly after 2011 when incidence remained very low (Fig 6A&B). Regarding seronegative shedding, as predicted the proportion of shedding animals that were seronegative was high in 2012, immediately following the 2011 outbreak; however, intriguingly, no seronegative shedders were detected in 2013 or 2014 (Fig 6C). No seronegative shedding was detected in 2008 and 2009 either, but this is likely due to very small sample sizes of animals tested in these years (N=3-6) compared with 2010-2014 (N=73-291; S2 Table). The fact that seronegative shedding occurred the year immediately after a major outbreak, but not in the two years that followed, suggests that titer decline may occur more rapidly in some individuals than predicted by SUBCLINICAL animal titer kinetics and that shedding duration may be less than 2 years. In addition, the prevalence of longer-term chronic shedders might be quite low, resulting in low power to detect.

Although noisy, these field data align with our argument that animals that are shedding while seronegative (or low-titer seropositive) may be chronic shedders. Combined with our earlier result that antibody titers act as a rough measure of time since

421 infection, this provides an opportunity to learn more about shedding duration. Precise
422 quantitative estimates are impossible, particularly due to wide uncertainty on the slow
423 decay rate of low titers, but a lower bound on shedding duration can be computed using the
424 initial rapid decay rate. Assuming a constant titer decay rate of 0.058 $\log_2$ antibody titer
425 units/day (the median decay rate of the CLINICAL animals) and an initial titer of 11 (the
426 median initial titer of the CLINICAL animals), we conservatively estimate the approximate
427 time taken to reach a given titer level (S3 Table), e.g. we calculate that it takes at least 189
428 days to reach seronegative status. From this, we can estimate the approximate duration of
429 shedding for a PCR-positive individual with that antibody titer. However, applying similar
430 logic to the decay curves suggested by the SUBCLINICAL animals, and still assuming an
431 initial titer of 11, we obtain estimates of shedding duration for seronegative shedders that
432 are much longer, e.g. ~ 6 years. Given the important caveats stated above and the low
433 SUBCLINICAL sample size (N=2), as well as further biological caveats discussed below, the
434 true shedding duration of seronegative shedders likely falls somewhere between these two
435 estimates.
436
437 **Predicting Shedding.** Data on antibody titers (and in our case renal index, also derived
438 from serum samples) are often more readily available than those on active pathogen
439 shedding. If a clear relationship between one or both of these biomarkers and shedding can
440 be established, then shedding may be predicted when shedding data are otherwise
441 unavailable. Using a dataset including all animals (CLINICAL, SUBCLINICAL, WILD and
442 STRAND) for which shedding data were available, we performed logistic regression to
443 investigate how antibody titer and renal index were related to pathogen shedding (using

444 only the first sample date per individual). We found that antibody titer contributed
445 significantly to the final model ($p<0.001$), but that renal index ($p=0.96$) and the interaction
446 between renal index and antibody titer ($p=0.85$) did not. The probability of shedding
447 increased with increased antibody titer (Fig 7A; OR = 1.67, CI = 1.55 – 1.80). Using this
448 relationship, we predicted shedding status in STRAND animals for which shedding data
449 were missing, using only the first sample date per individual, and were able to produce a
450 more complete map of the host-pathogen space (Fig 7B&C). Using all datasets, but
451 including only the first sample collected for an animal, we estimated an overall shedding
452 prevalence of 0.22 (PCR tested and untested, n=1510) which is substantially higher than
453 the prevalence of 0.15 calculated using the raw data (PCR tested only, n=811), showing we
454 may be greatly underestimating shedding prevalence when shedding data are rare relative
455 to antibody titer data.
456

457 ## Discussion

458 We have introduced a host-pathogen mapping framework that characterizes the
459 progression of *L. interrogans* infections and clinical responses in California sea lions,
460 drawing on longitudinal data from individual animals. The usefulness of our host-pathogen
461 map for interpreting cross-sectional data arises from the overall consistency in biomarker
462 dynamics across individuals, and particularly within similar groups of individuals. In
463 longitudinally sampled animals, we found antibody titer acted as a rough clock marking
464 time since infection. Although there was heterogeneity in initial antibody titers and decline
465 rates, animals followed the same broad initial pattern of titer decline regardless of
466 pathogen shedding status and initial disease severity. Our longitudinal data were censored

– we lacked data from the earliest stages of infection in these stranded animals, especially for the SUBCLINICAL animals, and the CLINICAL animals were followed for at most 90 days – yet trajectories of antibody titer decline for both groups overlapped and converged, suggesting that ultimately they follow the same long-term dynamics. Analysis of cross-sectional data corroborated this finding, as they traced the same broad trajectory as the overlapping SUBCLINICAL and CLINICAL groups, knitting them together. In addition, consistent with our observation of an initial rapid antibody decline followed by a slow second phase of decline, cross-sectional data revealed that in the years following a major outbreak and before another one occurred, the relative proportion of low titer animals increased with time as initially high titers followed biphasic decline kinetics.

Our host-pathogen map, which shows how severely ill and subclinical infections are linked, enables us to map the complexity of the host-pathogen relationship, resolve questions about apparently anomalous presentations, and is useful in addressing a particular challenge with respect to the traditional dichotomous view of *Leptospira*-host relationships. This view describes host species as either reservoir hosts that experience little disease but can shed chronically for months to years, or accidental hosts that can become severely ill, and possibly die, but do not become chronically infected [1, 25]. Thus, sea lions present an interesting challenge to this view as they show characteristics of both presentations. For example, sea lions showed a range of clinical disease in the initial stage of infection (i.e. at high titers), and although shedding was detected in the earliest phase regardless of clinical disease, there was substantial variation in shedding duration as determined by detection of both shedding and non-shedding sea lions at each antibody titer magnitude (including seronegative). Our mapping approach resolves this tension by

490 showing that, in the case of *Leptospira* infection in sea lions, the accidental and reservoir
491 characteristics seen in individual sea lion presentations are extremes of a unifying
492 trajectory of the host-pathogen interaction, and our map shows how these classical
493 manifestations are linked both within an individual's infection as well as on a population
494 level. This approach is of particular value for wildlife disease ecology, since many host-
495 pathogen relationships are poorly characterized, and an in-depth study in a controlled
496 experimental setting is generally not possible.

497 Characterizing the temporal trajectory of the infection and recovery process and
498 establishing consistency in antibody kinetics is especially important for accurately
499 interpreting the relationship between antibody titer and leptospiral DNA shedding. This
500 enables rough estimation of shedding duration and potentially identification of chronic
501 shedders – data that are key to accurate model parameterization, but notoriously difficult
502 to collect for wildlife systems. Infectious disease theory predicts that shedding duration
503 will influence transmission dynamics and modeling efforts have shown that chronically
504 shedding individuals play a critical role in population-level pathogen persistence both in
505 general [31] and specifically in California sea lions [19]. Using our map of titer decay and
506 clinical recovery, we identify low titer and seronegative shedders as likely chronic
507 shedders, and using titer kinetics of longitudinally followed animals we obtain estimates of
508 shedding duration from animals sampled only once.

509 Quantitative analysis of biomarker data can also define relationships between
510 biomarkers and specific disease outcomes. Results from these analyses can then be applied
511 to fill gaps in incomplete datasets. This is particularly important as some desired data types
512 are more difficult to obtain, or are unavailable at particular time points, e.g., survival data

513 are only available when no longer clinically useful, and urine can be more difficult to collect
514 than serum. We predicted urinary shedding of *Leptospira* from antibody titer magnitude,
515 and by applying this relationship across our full dataset we estimated a higher shedding
516 prevalence than that in the smaller dataset for which shedding data were available.
517 However, the opposite could have been true had the group composition been different in
518 the smaller 'training' dataset. Therefore, in all cases of establishing these relationships
519 between biomarkers, the impact of group composition and the epidemiological context of
520 the data must be carefully considered. For example, in our case, seronegative shedding was
521 more common the year following a major outbreak, had only these data been used when
522 establishing the relationship between titer and shedding, shedding prevalence estimates
523 would have been much higher. In future work, increasing the amount of data included –
524 biomarker, demographic, environmental – in models defining these relationships, as in
525 Borremans et al. [6], may further improve estimates by accounting for differences among
526 individuals, and time periods.

527 By analyzing quantitative antibody titers jointly with other biomarker data, such as
528 clinical or infection status, individuals that are indistinguishable by one measure but in fact
529 are biologically distinct, may be better characterized and identified. For example,
530 subclinical infection was seen across all titer levels, but severe disease was seen almost
531 exclusively at high titers – as titers declined, only the rare outlier showed evidence of
532 clinical disease. Using a binary approach to interpreting antibody titers, the outlier
533 individuals with mid-low titer and severe renal disease may be categorized as
534 'seropositive' and miscategorized as acute *Leptospira* infections. However, through
535 comparisons with a group of individuals thought never to have been infected, we show that

536 these outliers were infected in the past (giving rise to the detected antibodies) but likely
537 exhibited renal disease from another more recent etiology. This argument is supported by
538 the lack of leptospiral DNA shedding in any of these outlier individuals. This general
539 principle is relevant to the broader field of public health, as a pervasive health issue in
540 developing countries is the similarity in disease presentations among a diverse group of
541 infectious and non-infectious etiologies (e.g. acute febrile illnesses; pneumonia; diarrheal
542 disease [32, 33]) leading to possible misdiagnosis. Quantitative analysis of multiple
543 biomarkers, as exemplified by our host-pathogen map, facilitates identification of outliers
544 and thus cryptic causes of disease. This information can aid appropriate treatment choice
545 by clinicians, management recommendations by epidemiologists, and accurate estimation
546 of health and economic burden of a particular pathogen by public health agencies [34, 35].
547 Using our map and analyses of multiple biomarkers we are able to make predictions
548 regarding survival, shedding duration, and etiology of clinical disease. We propose that our
549 approach, or a similar one, may be applied to other host-pathogen systems, but system-
550 specific modifications may be required. In some systems, antibody titer dynamics may
551 contain more heterogeneity than seen in the *Leptospira*-sea lion system [6, 15, 16, 20, 21],
552 necessitating adjustments in the construction or interpretation of the host-pathogen map
553 such as using different biomarkers, or including more of them [3, 6]. For example, working
554 in an experimental system, Torres et al. [3] used blood pathogen concentration, instead of
555 antibody titer, and multiple measures of host health to build a map of 'disease space' within
556 which individuals that became severely ill and died, and those that survived, traced
557 different pathways as they moved from time since infection. They hypothesized that using

this map they could plot cross-sectional data and infer an individual's infection time-line and predict their prognosis, but did not test this idea.

Heterogeneity in the maximum antibody titer, degree of clinical disease experienced, the shape of the titer decline curve, and the duration of detectable antibodies has been noted in a number of host-pathogen systems [8, 10, 13, 15, 20, 21], and in some cases, such heterogeneity is associated with specific characteristics of an individual's infection. Subclinical infections with other pathogens have been associated with lower maximum antibody titers [15, 21], and with shorter antibody titer persistence [15]. Therefore, it was not obvious *a priori* whether antibody kinetics of subclinical and clinical infections would be the same in California sea lions. However, our findings, and those of others, indicate that despite some heterogeneity in antibody titer magnitude, titer kinetics for clinical and subclinical infections were roughly the same for *Leptospira* in sea lions and Q fever in humans [22].

Similarly other studies have examined whether chronic infections might be associated with different titer dynamics, e.g. higher maximum titers and longer persistence have been seen for chronic cases of Q fever [13, 20]. However, if anything, our data show the opposite trend. Instead of higher titers and greater antibody persistence, some of our hypothesized chronic carriers had no detectable antibodies. While our two subclinical animals – neither of which shed beyond the first several months of infection – had detectable antibody titers for years. Long-term persistence of antibodies in the SUBCLINICAL animals may be due to their captive status and its impact on overall condition and immune function. Alternatively, the long duration of detectable antibodies in

580 the SUBCLINICAL animals and the lack of detectable titers in some of our shedders may
581 reflect the full range of expected variation in titer decay rates and hence titer persistence.
582 The antibody titer decline that we detected in our sea lions, despite continued
583 infection in some, may be due to pathogen-specific differences in the underlying host-
584 pathogen interaction. For example, with some infections, including Q fever, the pathogen
585 continues to circulate in the blood in chronically infected individuals, [36], stimulating the
586 immune system to continue to produce antibodies and resulting in persistently high titers.
587 Conversely, in chronic *Leptospira* infections, once leptospires have colonized the kidneys
588 they appear to evade the host immune system [37, 38], which would explain the observed
589 antibody titer decline in our system, despite chronic renal infection and shedding.
590 We believe our host-pathogen mapping approach yields many benefits, however the
591 following caveats – some specific to our system, some more generalizable – must be
592 considered. Individuals in any study population will experience differences in
593 environmental exposures and conditions and we know that this, and age specific
594 differences, can lead to variation in biomarker data. For example, adaptive immunity can be
595 influenced by many factors including age, nutritional status and pathogen exposure history
596 [39-41]. In addition, several idiosyncrasies in our study may have affected our findings.
597 Estimates from the two SUBCLINICAL animals may not precisely reflect population level
598 trends given the small sample size of the SUBCLINICAL group. Our observations were
599 censored, as data from the CLINICAL animals were limited to the early phase of disease and
600 recovery, there were only a few data points from this early phase of infection in the
601 SUBCLINICAL animals, and the very earliest phase from infection to initial illness was
602 unobserved for all animals. The initial infecting dose of *Leptospira* is unknown in all cases

603 and may have varied substantially, potentially impacting immune response and disease
604 severity. Both SUBCLINICAL and CLINICAL animals experienced potentially
605 immunomodulatory conditions, specific to their time in captivity, that animals in the wild
606 would not have. For example, the SUBCLINICAL animals remained in captivity where they
607 were neutered and maintained in excellent body condition in a controlled, predator-free
608 environment. Such conditions may have increased their reserves and their capacity to
609 invest in a costly immune response [42], resulting in differences in long-term antibody titer
610 kinetics relative to free-ranging animals, i.e. a slower decline and more persistent antibody
611 titers. Similarly, CLINICAL animals received medical treatment and supportive care which
612 may have affected survival, increased their reserves and boosted their immune potential.
613 Together, these factors may help explain why estimates of shedding duration for
614 seronegative shedders, based on SUBCLINICAL antibody titer kinetics (i.e. roughly 6 years
615 to become seronegative), differed from patterns of seronegative shedding detected in the
616 wild after a major outbreak (i.e. decline to seronegative within 2 years). However, despite
617 the unique circumstances of our SUBCLINICAL and CLINICAL animals, the combined
618 longitudinal datasets describe a multiphase titer decline consistent with that found in other
619 studies [29, 43], and overall patterns seen in the longitudinal data were consistent with
620 those in the cross-sectional data and likely reflective of the entire sea lion population.
621 Approaches that integrate biomarker kinetics to interpret cross-sectional data can
622 be useful to clinicians and ecologists alike, and bridge perspectives from these often
623 separate worlds. Clinicians tend to focus on individual health, while ecologists focus more
624 on quantifying the natural process and understanding disease dynamics at the population
625 scale. Using this integrated approach, clinicians can improve individual patient survival

626 through more accurate patient triage and efficient allocation of resources, and can reduce
627 transmission risk to others. Treatments that are expensive, of limited availability, or time
628 intensive (e.g. dialysis) may be reserved for those individuals with the most severe disease
629 and the lowest probability of survival in the absence of such therapy. Conversely, in a
630 wildlife rehabilitation setting, limited resources might be directed towards those with the
631 highest probability of survival. Stricter, but possibly more expensive, measures to prevent
632 transmission can be efficiently directed at those individuals with the highest probability of
633 shedding. Ecologists can better conceptualize model structure using estimates of shedding
634 prevalence and duration and can better describe population level transmission dynamics.
635 For example, Buhnerkempe et al. [19] found that the addition of a chronic shedder
636 compartment to the traditional SIR (susceptible, infected, recovered/resistant) model was
637 necessary to accurately describe California sea lion-*Leptospira* dynamics and to capture
638 long-term patterns of pathogen persistence. Once model structure has been determined,
639 survival probabilities based on quantitative data (e.g. health, antibody titers) will influence
640 the duration spent in various model compartments and thus how they contribute to
641 onward transmission or herd immunity.

642 Historically, many of the principles of disease ecology were developed with
643 childhood diseases such as measles in mind, and these acute infections have much crisper
644 life histories for which the relatively simple SIR models can be used to capture the
645 dynamics [44] (See Fig 1 and S1 Box). As the field addresses more complex host-pathogen
646 relationships, these old assumptions break down and other models and approaches are
647 needed. Models need to include greater complexity such as longer and variable infectious
648 periods [19, 31], quantitative data and antibody titer kinetics [6, 15, 18], and multiple

649 biomarkers of disease [6] [4]. Models based on quantitative data and that integrate
650 antibody titer kinetics have been found to result in better estimates of model parameters
651 (e.g. transmission rate, basic reproductive rate) and improved model performance and
652 predictive capability (e.g. force of infection, incidence of infection), especially when only
653 cross-sectional data were available [6, 10, 12-14, 16-18]. Similarly, models integrating
654 quantitative serologic data may provide better estimates of incidence, especially relative to
655 estimates based on reported rates of illness, as such reports miss subclinical infections [12]
656 [6, 15-17]. We propose that our host-pathogen map provides a framework with which to
657 visualize quantitative data from multiple biomarkers, determine the relationships between
658 them, and identify the temporal trajectory of infection and recovery as reflected in changes
659 in biomarker levels through time. This approach is especially useful for elucidating
660 pathogen dynamics in wildlife systems, which typically rely on cross-sectional data.
661 Ultimately this approach can clarify the biology of more complex host-pathogen systems,
662 and enable the design of more appropriate dynamical models and statistical methods.
663

664 ## Materials and Methods

665 ### Study Animals

666 An overview can be found in Table 1**.**

667 **Wild-caught California sea lions.** Urine (n=637) and serum (n=732) samples were
668 collected from anesthetized or physically restrained unique sea lions (n=730; 2 animals
669 were recaptured and resampled) caught between September 2008 and November 2014
670 from three regions – southern California (San Miguel and San Nicolas Islands), central
671 California (Año Nuevo Island, Monterey and San Francisco's Pier 39) and northern Oregon

672 (Astoria, OR). All urine collection occurred under anesthesia. To minimize anesthetic risk,
673 only apparently healthy animals were captured and sampled. Estimated ages ranged from 1
674 to 5 years. These animals represent a cross-sectional sampling of the apparently healthy,
675 wild, free-ranging population and we refer to them as "WILD".
676 **Stranded California sea lions.** Urine (n=166) and serum (n=797) samples were collected
677 from 724 unique California sea lions that stranded along the central and northern
678 California coast and were admitted to a marine mammal rehabilitation center (The Marine
679 Mammal Center, Sausalito, CA) between 2008 and 2014. Animals stranded due to illness or
680 injury of all kinds, including, but not limited to leptospirosis, domoic acid toxicity, trauma,
681 neoplasia, pneumonia and malnutrition. To match the age range for the wild-caught
682 animals, only animals between the ages of 1 and 5 years were included in the study. These
683 animals represent a cross-sectional sampling of the ill and injured sea lion population and
684 we refer to them as "STRAND". Only STRAND animals with both Microscopic Agglutination
685 Testing (MAT) and chemistry results from within 14 days of each other were included in
686 the study (>95% were from within 24-hours of each other). Urine PCR results were
687 included only if these results were from urine collected within 24 hours of serum collection
688 for chemistry analysis. This was to ensure that data on current infection (PCR) was from
689 the same time point as data on clinical status (serum chemistry), which can change
690 substantially quite rapidly. Serum antibody titers have a much slower rate of change,
691 therefore we allowed serum MAT results to be within 14 days of PCR and serum chemistry
692 sample collection dates.
693 **CLINICAL**: In 2010 and 2011, The Marine Mammal Center rescued and initiated treatment
694 on 91 subadult, juvenile and yearling sea lions that stranded due to severe leptospirosis. Of

these, 66 died, typically within days of stranding (median = 4 days, interquartile range =2-7). We tracked the progression of host response (serum chemistry values, anti-*Leptospira* antibody titers) and active infection (leptospiral DNA shedding in urine) in 13 sea lions that survived to be released, and we refer to them as "CLINICAL" animals. Animals were diagnosed with leptospirosis using a combination of clinical observation, serum chemistry data and necropsy data [45]. Animals were longitudinally sampled starting on Day 0 (their first day in rehabilitation; serum only) and then approximately every 14 days thereafter (serum and urine) until an individual's release back into the wild 6-12 weeks later, as described in Prager et al. [28]. These animals were not included in analyses of the larger STRAND dataset except when specifically noted.

**SUBCLINICAL**: Two animals, which we will refer to individually as SUB1 and SUB2, collectively as SUBCLINICAL, never showed clinical signs of disease the entire period during which they were monitored as determined by clinical observation, complete physical examinations, complete blood counts and serum chemistry data. SUB1 stranded as a yearling male (i.e. between 1 and 2 years) in southern California in June 2010 (as described in Prager et al. [27]) with a flipper injury that precluded release back into the wild, and initially had no detectable anti-*Leptospira* antibodies. SUB1 seroconverted (i.e. acquired anti-*Leptospira* antibodies) in rehabilitation, with no observed clinical signs, at some unknown time-point within a 15-month period and was shedding leptospiral DNA 62 days after the first detected anti-*Leptospira* antibodies [27]. SUB1 was adopted by the U.S. Navy Marine Mammal Program (MMP) July 12, 2012, and samples were provided for monitoring every 1 to 9 months for a total of 44 months from the initial date that antibodies were detected. Because the date of infection was unknown, for our analyses we

718 used the date of the first detected titer as Day 0. SUB2 stranded as a juvenile male (i.e.
719 between 2-4 years), moderately underweight with a neck laceration, and was brought to
720 The Marine Mammal Center October 22, 2011 for treatment. He showed no clinical signs of
721 leptospirosis, but had an initial, moderately high anti-*Leptospira* antibody titer ($\log_2$ titer =
722 10). SUB2 was released back into the wild 3 weeks later but stranded again 3 months after
723 the initial stranding event with flipper injuries and still no clinical signs of leptospirosis. He
724 was never re-released and was adopted by the Navy July 11, 2012. Samples were provided
725 for monitoring every 1-7 months for 44 months from Day 0, defined as the date of the first
726 detected antibody titer, which was also the day of first stranding. It is possible that this
727 animal experienced clinical disease prior to being monitored, however he never showed
728 clinical signs consistent with leptospirosis while in rehabilitation or while at the U.S. Navy
729 MMP. For both SUB1 and SUB2, the magnitude of the first detected anti-*Leptospira*
730 antibody titers, and timing (autumn of 2011, during a major *Leptospira* outbreak in the
731 wild sea lion population; Fig 2C&D and Fig 6A [23, 45]), suggest that exposure to *Leptospira*
732 was recent – i.e. within weeks or months.

733 Samples collected from stranded animals (STRAND, CLINICAL and SUBCLINICAL)
734 for this study were collected during routine clinical care at the rehabilitation centers and
735 under their approved National Oceanic and Atmospheric Administration (NOAA), NMFS-
736 Southwest Region Stranding Agreements under the authority of the Marine Mammal
737 Protection Act. Samples collected from SUB1 and SUB2 at the U.S. Navy MMP were collected
738 during their routine care and under U.S. Code, Title 10, USC 7524. The MMP houses and
739 cares for a population of California sea lions in San Diego Bay (CA, USA). MMP is accredited
740 by AAALAC International and adheres to the national standards of the U.S. Public Health

741 Service Policy on the Humane Care and Use of Laboratory Animals and the Animal Welfare
742 Act. During their clinical care, stranded animals received a variety of treatments which may
743 have included, but were not limited to, subcutaneous fluids, antimicrobials, sedatives, and
744 gastro-intestinal protectants.
745
746 **Sample Analysis**
747 Serum agglutination testing (MAT) was performed at the California Animal Health
748 and Food Safety (CAHFS) laboratory, Davis, California, or at the Centers for Disease Control
749 and Prevention (CDC), Atlanta, Georgia, using live cultured *Leptospira* spp. (reference
750 strains) to measure the serum anti-*Leptospira* antibody titers. Samples run at CAHFS were
751 run against a 6 serovar panel and samples run at the CDC were run against a 2 or 19
752 serovar panel (as described in Prager et al. [28]). We only report MAT titer results against
753 *L. interrogans* serovar Pomona as historically this is the strain that elicits the highest MAT
754 titer in the majority of California sea lions tested [24] and it is the only serovar isolated
755 from this species to date [27, 46]. Serum samples were tested at doubling dilutions starting
756 from 1:100, and agglutination was read using dark field microscopy. Endpoint titers were
757 reported as the highest dilution that agglutinated at least 50% of the cells for the strain
758 tested [47]. Titer results were log transformed for ease of interpretation using the
759 following formula: $\log_2(\text{titer}/100) + 1$, thus a titer of 1:100 = 1, 1:200 = 2, etc. Titers
760 reported as <1:100 were set equal to 0 on both the log and regular scale. Throughout the
761 paper "antibody titer" refers to this log transformed titer value. All animals with a
762 detectable titer (i.e. ≥1:100) were considered seropositive and assumed to have been
763 infected with *Leptospira* at some point.

764 Serum chemistry analyses of wild-caught sea lions and stranded sea lions from The
765 Marine Mammal Center were performed on an ACE® Clinical Chemistry System (Alfa
766 Wassermann, Inc., West Caldwell, New Jersey, USA), those of SUB1 were performed initially
767 on either a VetTest® 8008 Chemistry Analyzer (IDEXX Laboratories, Inc., Westbrook,
768 Maine, USA) or a Cobas 800 modular analyzer (Roche Diagnostics, Indianapolis, Indiana,
769 USA). Once SUB1 and SUB2 were at the MMP, serum chemistry analyses were performed
770 on a Roche Cobas 8000 system (Roche Molecular Systems, Pleasanton, CA, USA) by the
771 Naval Medical Center in San Diego, CA.
772 We assessed leptospiral DNA shedding in urine using real-time polymerase chain
773 reaction (PCR) as described in Wu et al., [48]. Because urine was collected under
774 anesthesia, and anesthesia can pose a health risk to compromised animals, only apparently
775 healthy wild-caught animals were caught and sampled, and of the STRAND animals, urine
776 was collected only from those undergoing anesthesia for other reasons or during necropsy.
777 Individuals shedding leptospiral DNA were considered infected and infectious as the
778 primary mode of transmission of *Leptospira* is through either direct or indirect contact
779 with leptospires shed in the urine of infected individuals [49].
780
781 **Data Analyses**
782 **Antibody Titer Kinetics.** For data from each CLINICAL individual, we used linear
783 regression to characterize how the $\log_2$-transformed antibody titer declined with time. We
784 calculated the rate of antibody titer decline as the slope of the regression line and the t ½ as
785 the negative inverse of the slope. Using these data we determined the median titer decay
786 rate and $t_{1/2}$ for the CLINICAL animals.

787 Visual inspection of the SUBCLINICAL data suggested a biphasic pattern of titer
788 decline (Fig 2D) during the time that they were monitored, echoing findings from earlier
789 work [29, 30]. Thus we used piecewise linear regression to estimate the titer decline rate
790 and $t_{1/2}$ of the first and second phases separately for SUB1 and SUB2. For each animal, we
791 estimated the specific day that determined the change point of the regression by fitting
792 models over a range of possible change points from 10 to 500 days and using the day that
793 yielded the model with the minimum mean-squared error.

794 **Shedding Duration.** We used the median antibody titer decline rate ($r$ = 0.058 $\log_2$
795 antibody titer units/day) of the CLINICAL animals (Table 2) as well as their median initial
796 antibody titer ($t_i$ = 11) to calculate an approximate estimate of the lower bound of shedding
797 duration ($D$) in days for each observed titer level $t$ in PCR-positive sea lions using the
798 following equation:

799 $$D = \frac{t_i - t}{r}$$

800 This is a lower bound because it ignores further shedding after the date of observation, and
801 it neglects any shedding that occurred at titer levels > $t_i$.

802 Similarly, using SUBCLINICAL animal biphasic decay rates, we calculated an approximate
803 estimate of the duration of shedding ($D_s$) if an animal started at an initial titer equivalent to
804 the median initial antibody titer of the CLINICAL animals ($t_i$ = 11) and continued shedding
805 until the animal became seronegative. Using the following equation we used the
806 SUBCLINICAL specific decay rates (Table 2; $r_1$ and $r_2$) to estimate the durations of the initial
807 and secondary phases for each animal, and the SUBCLINICAL specific titer at which the

808 phase switch occurred (SUB1 $t_s$ = 2; SUB2 $t_s$ = 5) to mark the switch from initial to second
809 phase decay rates:

810 $$D_s = \frac{t_i - t_s}{r_1} + \frac{t_s}{r_2}$$

811 **Renal Index.** Blood urea nitrogen, creatinine, sodium, chloride and phosphorus are serum
812 chemistry values known to change with leptospirosis-induced renal compromise [23]. We
813 used principal components analysis (PCA) to derive a single measure of renal function from
814 these five serum chemistry values, which we termed the renal index. PCA was performed in
815 R using the command "prcomp" in the program "stats"[50]. BUN was $\log_{10}$ transformed and
816 each individual serum chemistry measure was scaled to have unit variance prior to
817 analysis. The dataset used included all longitudinal data (CLINICAL and SUBCLINICAL), as
818 well as all WILD animals that were both seronegative for anti-*Leptospira* antibodies and
819 negative for urinary leptospiral DNA shedding (i.e. the 0* group from Table 3). CLINICAL
820 and SUBCLINICAL animals were included to capture the range of clinical compromise in
821 infected animals from initial infection through recovery, and the subset of WILD animals
822 was included to anchor the analysis with a group of apparently healthy, uninfected,
823 unexposed animals. The first principal component (PC1) explained 54.8% of the variation
824 in the data, and had factor loadings consistent with clinical reports of leptospirosis-induced
825 disease (i.e. indicating elevated blood urea nitrogen, creatinine, sodium, chloride and
826 phosphorus [23]). Therefore we used PC1 as the renal index to assess clinical severity of
827 leptospirosis. Similar PCA results were found using just cross-sectional data (STRAND). To
828 establish the range of values corresponding to healthy renal function, we computed the
829 95% interquantile range of renal index values (i.e. PC1) experienced by the apparently

healthy WILD animals (-1.72 to 1.74). As values increased above this range, so did the degree of renal compromise.

Using the linear coefficients associated with each serum chemistry variable for PC1 from this analysis of longitudinal and WILD data, we calculated renal index values for all animals with serum chemistry results in the STRAND and WILD datasets.

**Predicting Survival and Shedding.** We used logistic regression to assess renal score as a predictor of survival in stranded animals at admission. We used the first sample available from all animals in the STRAND and CLINICAL groups with samples collected within 72 hours of admission (n=103) and that were categorized as leptospirosis cases based on clinical signs and serum chemistry values (BUN > 100 mg/dl, creatinine > 2mg/dl). Because we sought to assess the usefulness of this prediction as a tool for triaging animals in a rehabilitation center, antibody titer data were not included as only serum chemistry results would be available at this time.

In a separate analysis, we used multivariate logistic regression to assess predictors of leptospiral DNA shedding. Candidate predictors included serum anti-*Leptospira* antibody titer, renal index scores, and the interaction between these two variables. The dataset included all study animals for which we had PCR results, but only the first PCR result per individual. We used the "anova" command in the R package "stats" [50] to perform backward stepwise selection and the likelihood ratio method to include only variables that contributed significantly at the 0.05 level to the final model. Our final model included only antibody titer, so we then used the relationship between shedding and titer to predict the shedding status of the untested animals. To do this, we calculated the expected number of

853 shedders amongst the untested animals at each observed titer level using the total number
854 of untested animals and the probability of shedding at that titer level. We then randomly
855 selected this expected number of animals from amongst the untested animals at that titer
856 level and assigned them a positive shedding status. We performed logistic regression in R
857 using the "glm" command in the package "stats" [50].
858 **Comparing Renal Index Distributions.** We used the Kolmogorov-Smirnov (KS) test to
859 assess differences in distributions of renal index scores between groups of sea lions
860 (CLINICAL, SUBCLINICAL, STRAND, and WILD) and within groups by anti-*Leptospira*
861 antibody titer level. Because distributions were not continuous, we used the bootstrap
862 Kolmogorov-Smirnov test "ks.boot" in the "Matching" package in R [51]. To achieve
863 sufficient sample sizes, titers were collapsed into five groups, based on the titer kinetics
864 observed in longitudinally followed animals (CLINICAL and SUBCLINICAL). The highest
865 grouping included titers ≥ 11, consistent with the majority of the initial titers in CLINICAL
866 animals (11/13 had titers ≥ 11). All CLINICAL animals were in the healthy renal index
867 range by the time they had a titer of 8 and were released by the time their titer declined to
868 6, so these levels were used to define the ranges of the next two groupings: titers 9-10 to
869 capture animals recovering from clinical disease, and 6-8 to capture the recovered,
870 subclinical phase as seen in CLINICAL animals. Titer group 1-5 captured the longer-term
871 subclinical phase, as seen in SUBCLINICAL animals. Titer group 0 captured seronegative
872 animals.
873 **95% Confidence Intervals (CI).** 95% CI in Fig 6 were calculated in R using binom.confint
874 in the package "binom" using the Pearson-Klopper method [52]. 95% CI for Table 2 were

875 calculated using normal approximations based on linear regressions for antibody titer
876 kinetics.

877 **Figures.** All figures were made in R. Logistic regressions were plotted using logi.hist.plot in
878 the package "popbio" [53], all other figures were made using ggplot in the package
879 "ggplot2" [54].

880 **Ethics Statement**

881 All California sea lion samples were collected under authority of Marine Mammal
882 Protection Act Permits No. 932-1905-00/MA-009526 and No. 932-1489-10 issued by the
883 National Marine Fisheries Service (NMFS), NMFS Permit Numbers 17115-03, 16087-03,
884 and 13430. The sample collection protocol was approved by the Institutional Animal Care
885 and Use Committees (IACUC) of The Marine Mammal Center (Sausalito, CA; protocol #
886 2008-3), the University of California Los Angeles (ARC # 2012-035-12), and the Marine
887 Mammal Laboratory (Alaska Northwest 2013-1 and 2013-5). The Marine Mammal Center
888 and the Marine Mammal Laboratory adhere to the national standards of the U.S. Public
889 Health Service Policy on the Humane Care and Use of Laboratory Animals and the USDA
890 Animal Welfare Act. UCLA and the U.S Navy Marine Mammal Program are accredited by
891 AAALAC International and adhere to the national standards of the U.S. Public Health
892 Service Policy on the Humane Care and Use of Laboratory Animals and the USDA Animal
893 Welfare Act.

894

895 ## Acknowledgements

896 We would like to thank the volunteers, veterinarians, biologists and staff from The Marine
897 Mammal Center (Sausalito, CA), The Marine Mammal Care Center Los Angeles, the National

898 Marine Mammal Laboratory Alaska Fisheries Service, Oregon and Washington
899 Departments of Fish and Game, at the U.S. Navy Marine Mammal Program, the Año Nuevo
900 State Park, and the University of California Santa Cruz's Año Nuevo Reserve for their
901 logistical support of this study and assistance with sample and data collection. In particular
902 we would like to thank Carlos Rios, Lauren Palmer, Jeff Harris, Sharon Melin, Robert
903 DeLong, Julia Burco, Patrick Robinson, Celeste Parry, Guy Oliver, and Pat Morris. We would
904 also like to thank the Lloyd-Smith Laboratory at UCLA for useful discussions about the
905 manuscript. This work was supported by the National Science Foundation Awards OCE-
906 1335657 (https://www.nsf.gov/funding/pgm_list.jsp?org=OCE; JL-S, KCP) and DEB-
907 1557022 (https://www.nsf.gov/funding/pgm_list.jsp?org=DEB; JL-S and KCP), the John H.
908 Prescott Marine Mammal Rescue Assistance Grant Program
909 (https://www.fisheries.noaa.gov/grant/john-h-prescott-marine-mammal-rescue-
910 assistance-grant-program; JL-S), the Hellman Family Foundation
911 (https://www.apo.ucla.edu/faculty-career-development/hellman-fellowship/hellman; JL-
912 S), the US Department of Defense Strategic Environmental Research and Development Program
913 Award RC-2635 (https://www.serdp-estcp.org/Funding-Opportunities/SERDP-Solicitations; JL-
914 S, KCP), the De Logi Chair in Biological Sciences (https://www.apo.ucla.edu/academic-
915 listings/endowed-chairs; JL-S), and the Research and Policy for Infectious Disease
916 Dynamics (RAPIDD) program of the Science and Technology Directory, Department of
917 Homeland Security, and Fogarty International Center, National Institutes of Health
918 (https://www.fic.nih.gov/About/Staff/Pages/epidemiology-population.aspx JL-S, KCP,
919 MGB). The funders had no role in study design, data collection and analysis, decision to
920 publish, or preparation of the manuscript.

921

922

# References

1. Faine S, Adler B, Bolin C, Perolat P. *Leptospira* and leptospirosis. 2 ed ed. Melbourne, Australia: MediSci; 1999.

2. Gilbert AT, Fooks AR, Hayman DT, Horton DL, Muller T, Plowright R, et al. Deciphering Serology to Understand the Ecology of Infectious Diseases in Wildlife. Ecohealth. 2013. Epub 2013/08/07. doi: 10.1007/s10393-013-0856-0. PubMed PMID: 23918033.

3. Torres BY, Oliveira JH, Thomas Tate A, Rath P, Cumnock K, Schneider DS. Tracking Resilience to Infections by Mapping Disease Space. PLoS Biol. 2016;14(4):e1002436. doi: 10.1371/journal.pbio.1002436. PubMed PMID: 27088359.

4. Rydevik G, Innocent GT, Marion G, Davidson RS, White PC, Billinis C, et al. Using Combined Diagnostic Test Results to Hindcast Trends of Infection from Cross-Sectional Data. Plos Comput Biol. 2016;12(7):e1004901. Epub 2016/07/08. doi: 10.1371/journal.pcbi.1004901. PubMed PMID: 27384712; PubMed Central PMCID: PMCPMC4934910.

5. Strelioff CC, Vijaykrishna D, Riley S, Guan Y, Peiris JSM, Lloyd-Smith JO. Inferring patterns of influenza transmission in swine from multiple streams of surveillance data. Proceedings of the Royal Society B-Biological Sciences. 2013;280(1762). doi: 10.1098/rspb.2013.0872. PubMed PMID: WOS:000319385100020.

6. Borremans B, Hens N, Beutels P, Leirs H, Reijniers J. Estimating Time of Infection Using Prior Serological and Individual Information Can Greatly Improve Incidence Estimation of Human and Wildlife Infections. Plos Comput Biol. 2016;12(5):e1004882. PubMed PMID: Medline:27177244.

946 7. Chantrey J, Dale T, Jones D, Begon M, Fenton A. The drivers of squirrelpox virus
947 dynamics in its grey squirrel reservoir host. Epidemics-Neth. 2019:100352. Epub
948 2019/07/23. doi: 10.1016/j.epidem.2019.100352. PubMed PMID: 31327730.
949 8. Strid MA, Engberg J, Larsen LB, Begtrup K, Molbak K, Krogfelt KA. Antibody
950 responses to *Campylobacter* infections determined by an enzyme-linked immunosorbent
951 assay: 2-year follow-up study of 210 patients. Clin Diagn Lab Immun. 2001;8(2):314-9.
952 PubMed PMID: Medline:11238214.
953 9. Teunis PF, van Eijkeren JC, de Graaf WF, Marinovic AB, Kretzschmar ME. Linking the
954 seroresponse to infection to within-host heterogeneity in antibody production. Epidemics-
955 Neth. 2016;16:33-9. doi: 10.1016/j.epidem.2016.04.001. PubMed PMID: 27663789.
956 10. Wielders CCH, Teunis PFM, Hermans MHA, van der Hoek W, Schneeberger PM.
957 Kinetics of antibody response to *Coxiella bumetii* infection (Q fever): Estimation of the
958 seroresponse onset from antibody levels. Epidemics-Neth. 2015;13:37-43. doi:
959 10.1016/j.epidem.2015.07.001. PubMed PMID: WOS:000365890900005.
960 11. Charleston B, Bankowski BM, Gubbins S, Chase-Topping ME, Schley D, Howey R, et
961 al. Relationship Between Clinical Signs and Transmission of an Infectious Disease and the
962 Implications for Control. Science. 2011;332(6030):726-9. doi: 10.1126/science.1199884.
963 12. Teunis PFM, van Eijkeren JCH, Ang CW, van Duynhoven YTHP, Simonsen JB, Strid
964 MA, et al. Biomarker dynamics: estimating infection rates from serological data. Stat Med.
965 2012;31(20):2240-8. doi: Doi 10.1002/Sim.5322. PubMed PMID: ISI:000307648900007.
966 13. Teunis PFM, Schimmer B, Notermans DW, Leenders ACAP, Wever PC, Kretzschmar
967 MEE, et al. Time-course of antibody responses against *Coxiella burnetii* following acute Q

968 fever. Epidemiol Infect. 2013;141(1):62-73. doi: Doi 10.1017/S0950268812000404.
969 PubMed PMID: ISI:000312037600008.
970 14. Bollaerts K, Aerts M, Shkedy Z, Faes C, Van der Stede Y, Beutels P, et al. Estimating
971 the population prevalence and force of infection directly from antibody titres. Statistical
972 Modelling. 2012;12(5):441-62. doi: 10.1177/1471082x12457495. PubMed PMID:
973 WOS:000308652200003.
974 15. Buchy P, Vong S, Chu S, Garcia J-M, Hien TT, Hien VM, et al. Kinetics of Neutralizing
975 Antibodies in Patients Naturally Infected by H5N1 Virus. Plos One. 2010;5(5):e10864. doi:
976 10.1371/journal.pone.0010864.
977 16. Pepin KM, Kay SL, Golas BD, Shriner SS, Gilbert AT, Miller RS, et al. Inferring
978 infection hazard in wildlife populations by linking data across individual and population
979 scales. Ecol Lett. 2017;20(3):275-92. doi: 10.1111/ele.12732. PubMed PMID:
980 WOS:000395169300001.
981 17. Simonsen J, Mølbak K, Falkenhorst G, Krogfelt KA, Linneberg A, Teunis PFM.
982 Estimation of incidences of infectious diseases based on antibody measurements. Stat Med.
983 2009;28(14):1882-95. doi: 10.1002/sim.3592.
984 18. Vinh DN, Boni MF. Statistical identifiability and sample size calculations for serial
985 seroepidemiology. Epidemics-Neth. 2015;(0). doi:
986 http://dx.doi.org/10.1016/j.epidem.2015.02.005.
987 19. Buhnerkempe MG, Prager KC, Strelioff CC, Greig DJ, Laake JL, Melin SR, et al.
988 Detecting signals of chronic shedding to explain pathogen persistence: *Leptospira*
989 *interrogans* in California sea lions. J Anim Ecol. 2017;86(3):460-72. doi: 10.1111/1365-
990 2656.12656. PubMed PMID: 28207932.

991 20. Dupont HT, Thirion X, Raoult D. Q fever serology: cutoff determination for
992 microimmunofluorescence. Clin Diagn Lab Immun. 1994;1(2):189-96. PubMed PMID:
993 Medline:7496944.
994 21. de Graaf WF, Kretzschmar MEE, Teunis PFM, Diekmann O. A two-phase within-host
995 model for immune response and its application to serological profiles of pertussis.
996 Epidemics-Neth. 2014;9(0):1-7. doi: http://dx.doi.org/10.1016/j.epidem.2014.08.002.
997 22. Teunis PFM, Teunis PFM, Ang CW, Strid MA, De Valk H, Sadkowska-Todys M, et al.
998 *Campylobacter* seroconversion rates in selected countries in the European Union.
999 Epidemiol Infect. 2013;141(10):2051-7. doi: Doi 10.1017/S0950268812002774. PubMed
1000 PMID: ISI:000323887200006.
1001 23. Gulland FMD, Koski M, Lowenstine LJ, Colagross A, Morgan L, Spraker T.
1002 Leptospirosis in California sea lions (*Zalophus californianus*) stranded along the central
1003 California coast, 1981-1994. Journal of Wildlife Diseases. 1996;32(4):572-80. doi:
1004 10.7589/0090-3558-32.4.572. PubMed PMID: ISI:A1996VN97900002.
1005 24. Lloyd-Smith JO, Greig DJ, Hietala S, Ghneim GS, Palmer L, St Leger J, et al. Cyclical
1006 changes in seroprevalence of leptospirosis in California sea lions: endemic and epidemic
1007 disease in one host species? BMC Infectious Diseases. 2007;7:125-36. doi: 125
1008 10.1186/1471-2334-7-125. PubMed PMID: ISI:000252421600001; PubMed Central
1009 PMCID: PMC2186333.
1010 25. Zuerner RL. Host Response to *Leptospira* Infection. In: Adler B, editor. *Leptospira*
1011 and Leptospirosis. Berlin, Heidelberg: Springer Berlin Heidelberg; 2015. p. 223-50.
1012 26. Zuerner RL, Cameron CE, Raverty S, Robinson J, Colegrove KM, Norman SA, et al.
1013 Geographical dissemination of *Leptospira interrogans* serovar Pomona during seasonal

migration of California sea lions. Veterinary Microbiology. 2009;137(1-2):105-10. doi: Doi 10.1016/J.Vetmic.2008.12.017. PubMed PMID: ISI:000266578400015.

27. Prager KC, Greig DJ, Alt DP, Galloway RL, Hornsby RL, Palmer LJ, et al. Asymptomatic and chronic carriage of *Leptospira interrogans* serovar Pomona in California sea lions (*Zalophus californianus*). Veterinary Microbiology. 2013;164(1-2):177-83. Epub 2013/02/20. doi: 10.1016/j.vetmic.2013.01.032. PubMed PMID: 23419822.

28. Prager KC, Alt DP, Buhnerkempe MG, Greig DJ, Galloway RL, Wu Q, et al. Antibiotic efficacy in eliminating leptospiruria in California sea lions (*Zalophus californianus*) stranding with leptospirosis. Aquatic Mammals. 2015;41(2):203-12. doi: 10.1578/AM.41.2.2015.203 PubMed PMID: WOS:000358296900008.

29. Hammarlund E, Thomas A, Amanna IJ, Holden LA, Slayden OD, Park B, et al. Plasma cell survival in the absence of B cell memory. Nature Communications. 2017;8(1):1781. doi: 10.1038/s41467-017-01901-w.

30. Amanna IJ, Hammarlund E, Lewis MW, Slifka MK. Impact of infection or vaccination on pre-existing serological memory. Human Immunology. 2012;73(11):1082-6. doi: 10.1016/j.humimm.2012.07.328. PubMed PMID: WOS:000310717500004.

31. Kramer-Schadt S, Fern·ndez N, Eisinger D, Grimm V, Thulke H-H. Individual variations in infectiousness explain long-term disease persistence in wildlife populations. Oikos. 2009;118(2):199-208.

32. Lozano R, Naghavi M, Foreman K, Lim S, Shibuya K, Aboyans V, et al. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. The Lancet. 2012;380(9859):2095-128. doi: https://doi.org/10.1016/S0140-6736(12)61728-0.

1037 33. Murray CJL, Vos T, Lozano R, Naghavi M, Flaxman AD, Michaud C, et al. Disability-
1038 adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990–2010: a
1039 systematic analysis for the Global Burden of Disease Study 2010. The Lancet.
1040 2012;380(9859):2197-223. doi: https://doi.org/10.1016/S0140-6736(12)61689-4.
1041 34. Bogich TL, Funk S, Malcolm TR, Chhun N, Epstein JH, Chmura AA, et al. Using
1042 network theory to identify the causes of disease outbreaks of unknown origin. J R Soc
1043 Interface. 2013;10(81).
1044 35. Crump JA. Time for a comprehensive approach to the syndrome of fever in the
1045 tropics. Transactions of The Royal Society of Tropical Medicine and Hygiene.
1046 2014;108(2):61-2. doi: 10.1093/trstmh/trt120.
1047 36. Kampschreur LM, Oosterheert JJ, Koop AMC, Wegdam-Blans MCA, Delsing CE,
1048 Bleeker-Rovers CP, et al. Microbiological Challenges in the Diagnosis of Chronic Q Fever.
1049 Clinical and Vaccine Immunology. 2012;19(5):787-90. doi: 10.1128/Cvi.05724-11. PubMed
1050 PMID: WOS:000307108600022.
1051 37. Monahan AM, Callanan JJ, Nally JE. Review paper: Host-pathogen interactions in the
1052 kidney during chronic leptospirosis. Vet Pathol. 2009;46(5):792-9. Epub 2009/05/12. doi:
1053 10.1354/vp.08-VP-0265-N-REV. PubMed PMID: 19429975.
1054 38. Monahan AM, Callanan JJ, Nally JE. Proteomic analysis of *Leptospira interrogans* shed
1055 in urine of chronically infected hosts. Infect Immun. 2008;76(11):4952-8. PubMed PMID:
1056 Medline:18765721.
1057 39. Brodin P, Jojic V, Gao T, Bhattacharya S, Angel Cesar JL, Furman D, et al. Variation in
1058 the Human Immune System Is Largely Driven by Non-Heritable Influences. Cell.
1059 2015;160(1):37-47. doi: https://doi.org/10.1016/j.cell.2014.12.020.

1060 40. Carr EJ, Dooley J, Garcia-Perez JE, Lagou V, Lee JC, Wouters C, et al. The cellular
1061 composition of the human immune system is shaped by age and cohabitation. Nature
1062 Immunology. 2016;17:461. doi: 10.1038/ni.3371
1063 https://www.nature.com/articles/ni.3371#supplementary-information.
1064 41. Martinez DR, Permar SR, Fouda GG. Contrasting Adult and Infant Immune Responses
1065 to HIV Infection and Vaccination. Clin Vaccine Immunol. 2016;23(2):84-94. doi:
1066 10.1128/CVI.00565-15. PubMed PMID: PMC4744916.
1067 42. Lochmiller RL, Deerenberg C. Trade-offs in evolutionary immunology: just what is
1068 the cost of immunity? Oikos. 2000;88(1):87-98. doi: 10.1034/j.1600-0706.2000.880110.x.
1069 43. Andraud M, Lejeune O, Musoro JZ, Ogunjimi B, Beutels P, Hens N. Living on Three
1070 Time Scales: The Dynamics of Plasma Cell and Antibody Populations Illustrated for
1071 Hepatitis A Virus. Plos Comput Biol. 2012;8(3). doi: Artn E1002418
1072 Doi 10.1371/Journal.Pcbi.1002418. PubMed PMID: ISI:000302244000026.
1073 44. Anderson RM, May RM. Population biology of infectious diseases.1. Nature.
1074 1979;280(5721):361-7. PubMed PMID: ISI:A1979HE95900033.
1075 45. Greig DJ, Gulland FMD, Kreuder C. A decade of live California sea lion (*Zalophus*
1076 *californianus*) strandings along the central California coast: causes and trends, 1991-2000.
1077 Aquatic Mammals. 2005;33(1):11-22.
1078 46. Zuerner RL, Alt DP. Variable Nucleotide Tandem-Repeat analysis revealing a unique
1079 group of *Leptospira interrogans* serovar Pomona isolates associated with California sea
1080 lions. Journal of Clinical Microbiology. 2009;47(4):1202-5. doi: Doi 10.1128/Jcm.01639-08.
1081 PubMed PMID: ISI:000264797000047; PubMed Central PMCID: PMC2668322.

1082 47. Dikken H, Kmety E. Serological typing methods of leptospires. Bergan T, NOrris JR,
1083 editors. London, UK: Academic Press; 1978.
1084 48. Wu Q, Prager KC, Goldstein T, D.P. A, Galloway RL, Zuerner RL, et al. Development of
1085 a real-time PCR for the detection of pathogenic *Leptospira* spp. in California sea lions.
1086 Diseases of Aquatic Organisms. 2014;110(3):165-72. doi: 10.3354/dao02752. PubMed
1087 PMID: WOS:000342923400001.
1088 49. Ellis WA. Animal Leptospirosis. In: Adler B, editor. Current Topics in Microbiology
1089 and Immunology: *Leptospira* and Leptospirosis. Current Topics in Microbiology and
1090 Immunology. Verlag Berlin Heidelberg: Springer; 2015. p. 99-137.
1091 50. R Core Team. R: A language and environment for statistical computing. Vienna,
1092 Austria: R Foundation for Statistical Computing; 2017.
1093 51. Sekhon JS. Multivariate and Propensity Score Matching Software with Automated
1094 Balance Optimization:
1095 The Matching Package for R. Journal of Statistical Software. 2011;42(7):1-52.
1096 52. Dorai-Raj S. binom: Binomial Confidence Intervals For Several Parameterizations. R
1097 package version 1.1-1 ed2014.
1098 53. Stubben CJ, Milligan BG. Estimating and Analyzing Demographic Models Using the
1099 popbio Package in R. Journal of Statistical Software. 2007;22(11).
1100 54. Wickham H. ggplot2: Elegant Graphics for Data Analysis. New York: Springer-Verlag;
1101 2009.

1102

1103

# Supporting Information Legends

**S1 Box.** Comparison of host-pathogen interactions based on the canonical ecological model of infectious disease dynamics in which individuals can be classified into four groups: susceptible (S), exposed (E), infected/infectious (I) and recovered (R), with more complex host-pathogen interactions.

**S2 Table.** Total number of sea lions with a given $\log_2$ antibody titer by year for wild-caught (WILD), stranded (STRAND) and subclinically infected (SUB1 and SUB2) sea lions. In parentheses are the number of animals shedding leptospires for each $\log_2$ antibody titer level over the total number of PCR tested animals.

**S3 Table**. Predicted shedding duration by titer level assuming a constant titer decay rate of 0.058 log2 antibody titer units/day (the median decay rate of the CLINICAL animals) and an initial titer of 11 (the median initial titer of the CLINICAL animals).

**S4 Data.** Raw data on used for analyses. Columns include animal ID, DataType (i.e. CLINICAL, SUBCLINICAL, STRAND, WILD), AdmitYear (i.e. the year an animal was caught – WILD – or admitted for rehabilitation - CLINICAL, SUBCLINICAL, STRAND), SampleYearMAT (year serum was collected for serum MAT), LogMAT ($\log_2$ MAT result), SampleYearChem (year serum was collected for serum chemistry analysis), RenalIndex (renal index score calculated as described in the manuscript), SampleYearPCR (year urine was collected for PCR, PCR (result of PCR analysis), SurvivalData (information whether the animal survived or died during rehabilitation; wild-caught animals were released after capture, therefore survival data was unknown – NA), DaySinceAdmission (the number of days between admission to rehabilitation and date of sample collection for analysis (MAT,

1126 PCR, serum chemistry), DaysSinceFirstMAT (the number of days since sample collection for

1127 the first MAT analysis).

1128

# Figure Captions

**Fig 1. Map showing how infected individuals move through the "host-pathogen space" in dimensions reflecting severity of clinical disease (y-axis) and time since infection (~antibody titer, x-axis).** When susceptible individuals become infected (indicated by red shading – the intensity of shading indicates the probability that an assay for current infection – e.g. PCR – would be positive), they move through this space in a trajectory towards higher titer and more severe clinical disease states and then, depending on the type of pathogen and the host-pathogen interaction, they eventually return to a state of good health and their titers decline. The shape of this trajectory and their infection state will differ based on the host-pathogen system and the degree of heterogeneity in host responses. Here we show the trajectory that would be expected based on the canonical susceptible (S), exposed (E), infected/infectious (I) and recovered (R) model of infectious disease dynamics with the position of individuals as they would pass through each of the four states indicated by the blue circles. However, individuals – represented by question marks in this figure – experiencing more complex host-pathogen interactions may fall outside of this canonical trajectory (see S1 Box for further detail).

**Fig 2. Changes in antibody titer and renal index in longitudinally sampled California sea lions.** Renal index scores (A) and $\log_2$ anti-*Leptospira* antibody titer (C) by time in individual sea lions that stranded with clinical signs of leptospirosis and were followed for 6 – 12 weeks (CLINICAL dataset). Renal index scores (B) and $\log_2$ antibody titer (D) by time in two stranded sea lions – SUB1 (square, grey line) and SUB2 (triangle, orange line) – that never showed *Leptospira*-related clinical disease and were monitored for 3 years (SUB

dataset). In panel D, regression lines, as determined by piecewise linear regression, are drawn through first and second phases of antibody titer decline for each SUB animal. For CLINICAL animals, day 0 is the day of admission to rehabilitation, for SUB1 and SUB2, day 0 is the day when anti-*Leptospira* antibodies were first detected. Grey horizontal bands in panels C and D delineate the full range of initial antibody titers in the CLINICAL animals, and in panels A and B they delineate the 95% interquantile range of renal index scores in apparently healthy, uninfected, seronegative wild-caught animals.

**Fig 3. Map of host-pathogen space.** Maps of the host-pathogen space created by plotting jittered $\log_2$ anti-*Leptospira* antibody titers (x-axis) against renal index values (y-axis). Plots created using data from the longitudinally followed animals (CLINICAL and SUBCLINICAL), color-coded by time since admission to a rehabilitation center (A) and by urinary leptospire shedding status (B). Plots created using cross-sectional data from stranded animals (STRAND; C) and wild-caught, free-ranging animals (WILD; D) color-coded by urinary leptospire shedding status. In all panels, horizontal grey bands are equivalent to those in Fig1 C&D, and the vertical grey bands are equivalent to the horizontal bands described in Fig1 A&B.

**Fig 4. Renal index score distributions by $\log_2$ antibody titer level, for each sample group**. Sample groups are as described in the methods and are wild-caught (WILD), subclinical (SUB), clinical (CLINICAL), stranded (STRAND) sea lions. Titer groups were chosen based on the titer dynamics observed in the longitudinally followed animals through time. Groups roughly match the different phases of the host-pathogen relationship

ranging from initial infection (11+), clinical recovery (9-10) and two stages of historic infection (1-5 and 6-8). Group 0 contains seronegative animals. The grey line denotes the upper boundary of the healthy range for renal index values. The upper whisker extends from the hinge to the highest value that is within 1.5 * IQR of the hinge, where IQR is the interquartile range. The lower whisker extends from the hinge to the lowest value within 1.5 * IQR of the hinge.

**Fig 5. Survival probability predicted from renal index score.** The probability of survival as a function of renal index score plotted over histograms of the number of animals surviving (top histogram) or not surviving (bottom histogram) by renal index value. Analyses were run using data from samples collected within 72 hours of admission to a rehabilitation center from animals stranding and diagnosed with leptospirosis.

**Figure 6.** Proportion of stranded animals with leptospirosis (A), proportion of seropositive animals that have low titers ($log_2$ titer 1-5) by year with 95% CI (B), and proportion of shedding animals that are seronegative by year with 95% CI (C). Total sample size for each proportion is indicated within the box. Only STRAND data included in (A) WILD and STRAND data for (B) and (C). The proportion of leptospirosis strands is highest in the two outbreak years – 2008 and 2011 – and the proportion of low titer animals increases with each year after the outbreaks. Similarly, the proportion of seronegative shedders increases after the major outbreak in 2011, but then declines to zero by 2013. The single shedder in 2013 had a $log_2$ antibody titer of 3, no animals were shedding in 2014; therefore a proportion could not be calculated. Few shedders were detected in 2008 and 2009 due to

small sample sizes of animals PCR tested for shedding (2008 N=6, 2009 N=3, 2010 N=73, 2011 N=158, 2012 N=119, 2013 N=162, 2014 N=291).

**Fig 7. Shedding probability predicted from $log_2$ anti-*Leptospira* antibody titer.** (A) The probability of shedding as a function of $log_2$ anti-*Leptospira* antibody titer plotted over histograms of the number of animals shedding (top histogram) and not shedding (bottom histogram) by antibody titer (A). STRAND data plotted using the 'host-pathogen map' framework as in Fig 3C. Data divided into those individuals that were PCR tested and for which shedding status was known (B), and those that were not PCR tested and for which shedding status was predicted (C; positive = red, negative = black).

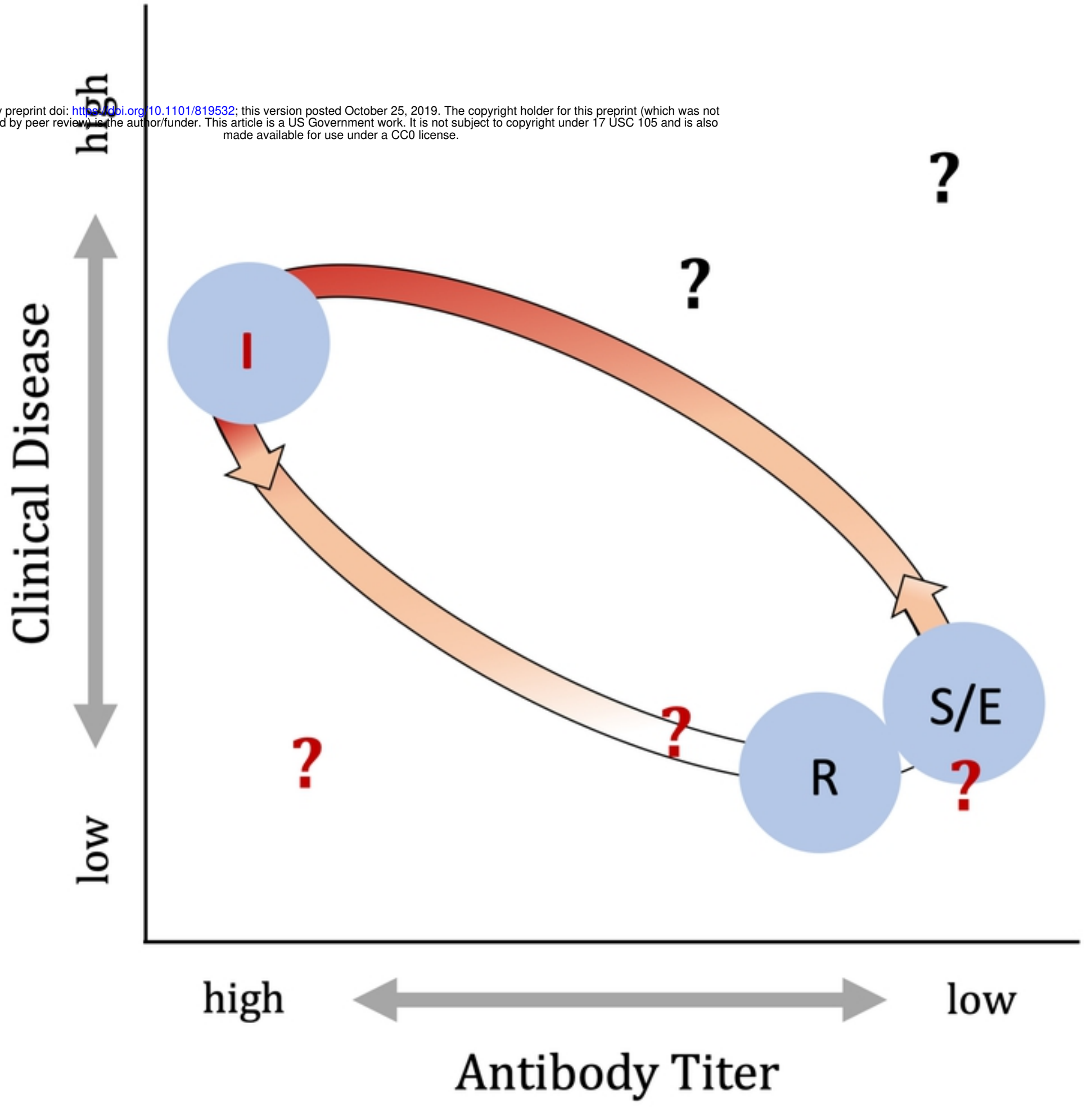


Figure 1

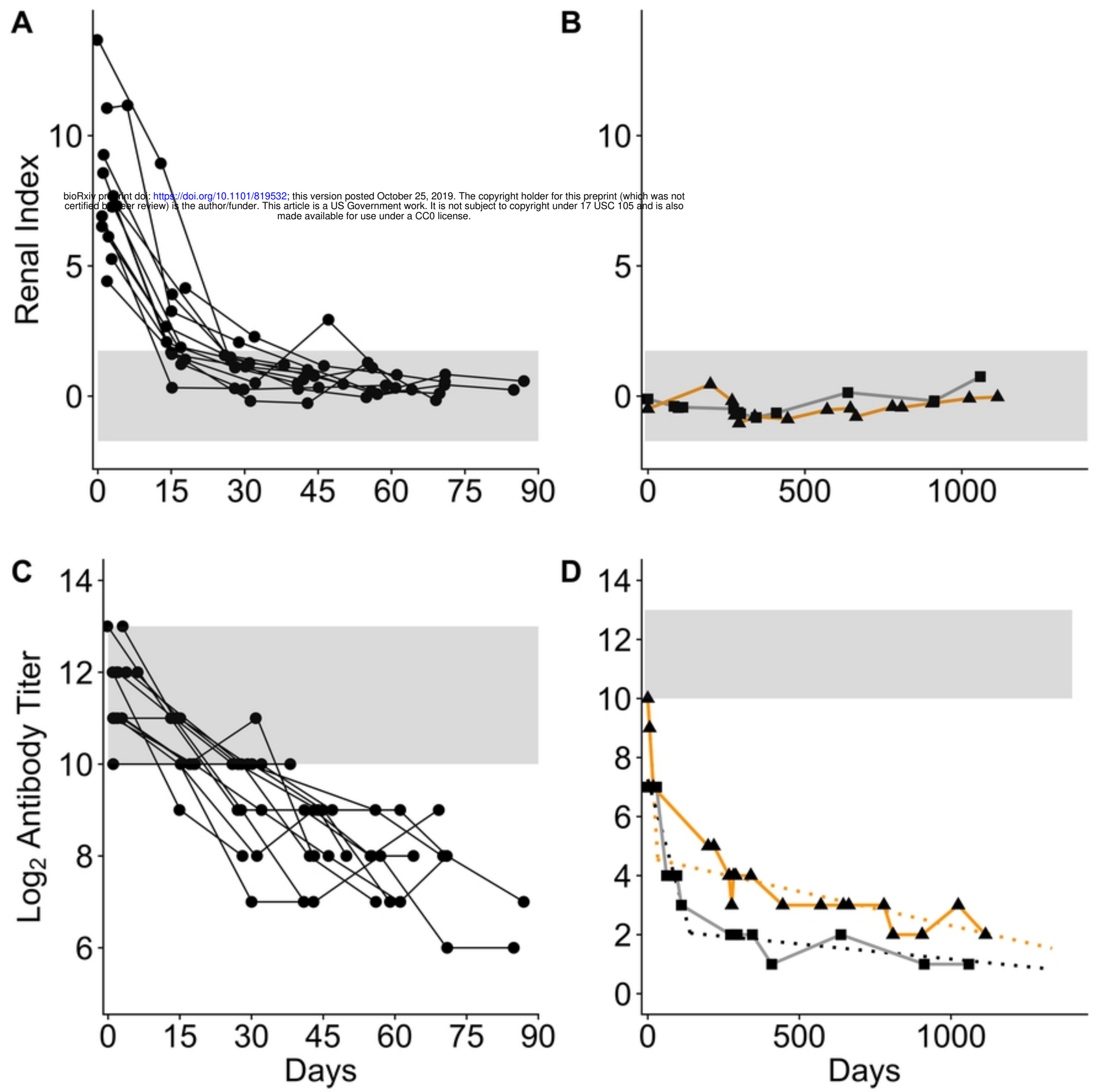


Figure 2

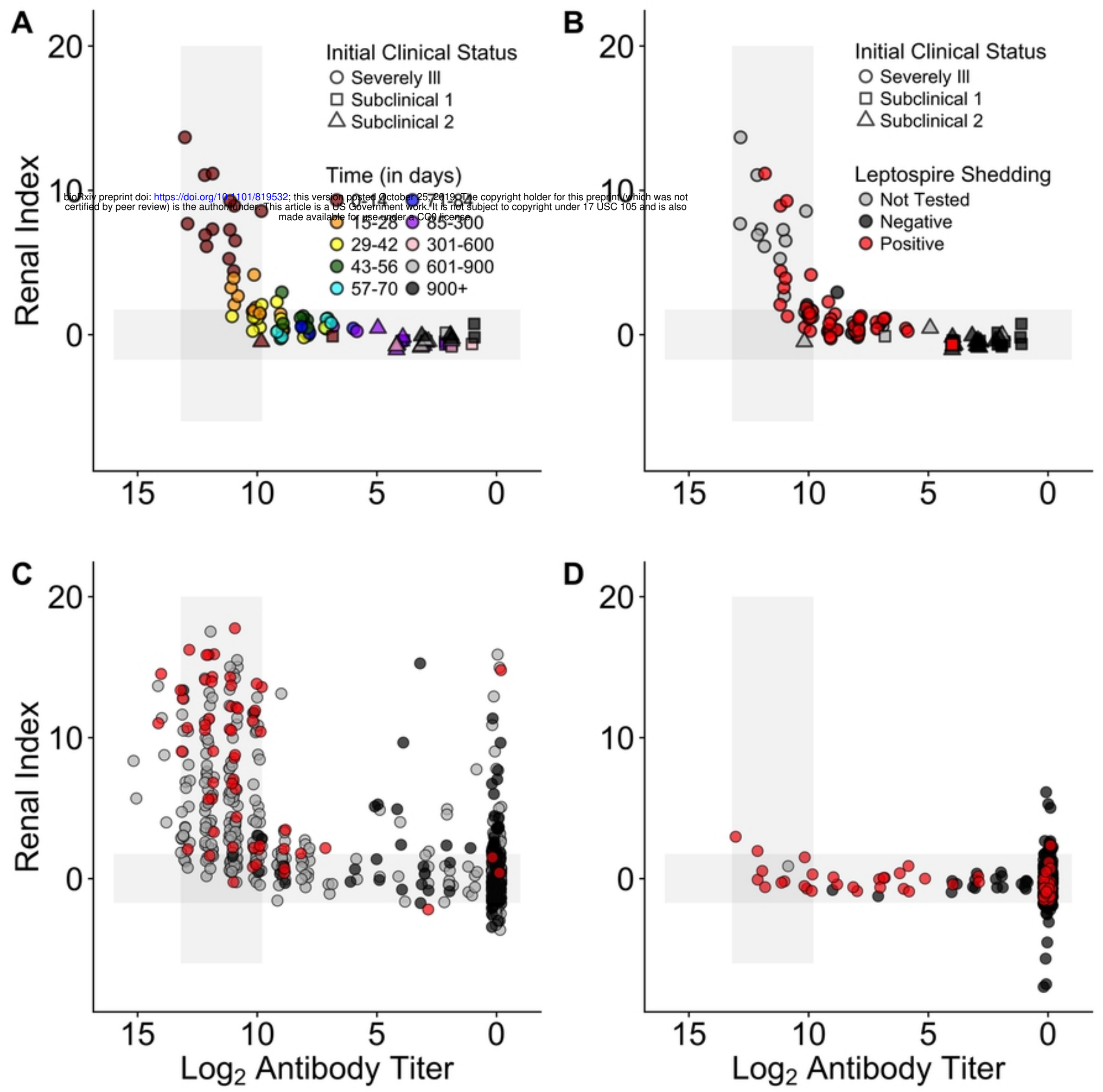


Figure 3

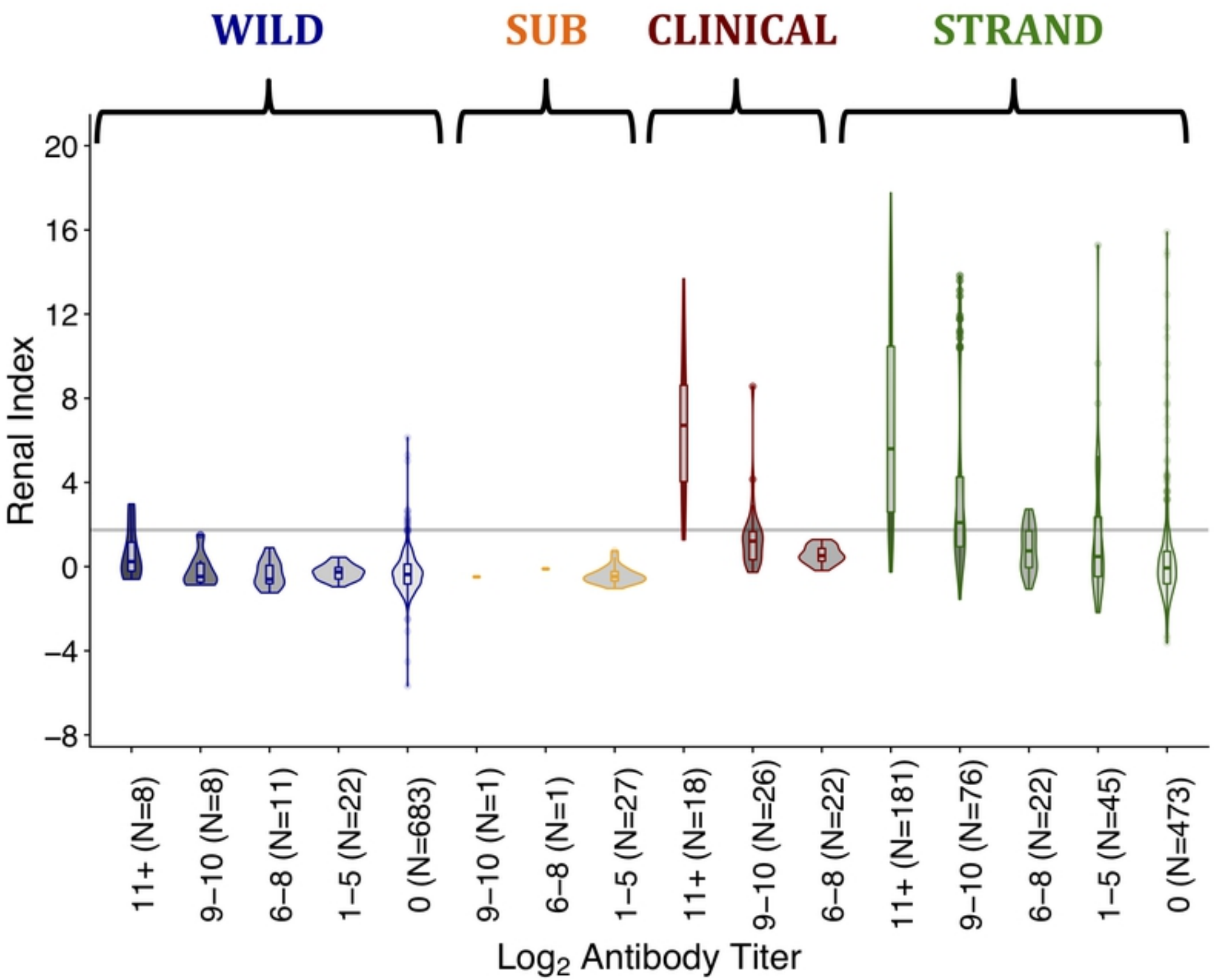
WILD
SUB
CLINICAL
STRAND
20
16
12
8
4
0
−4
−8
Renal Index
11+ (N=8)
9–10 (N=8)
6–8 (N=11)
1–5 (N=22)
0 (N=683)
9–10 (N=1)
6–8 (N=1)
1–5 (N=27)
11+ (N=18)
9–10 (N=26)
6–8 (N=22)
11+ (N=181)
9–10 (N=76)
6–8 (N=22)
1–5 (N=45)
0 (N=473)
Log2 Antibody Titer

Figure 4

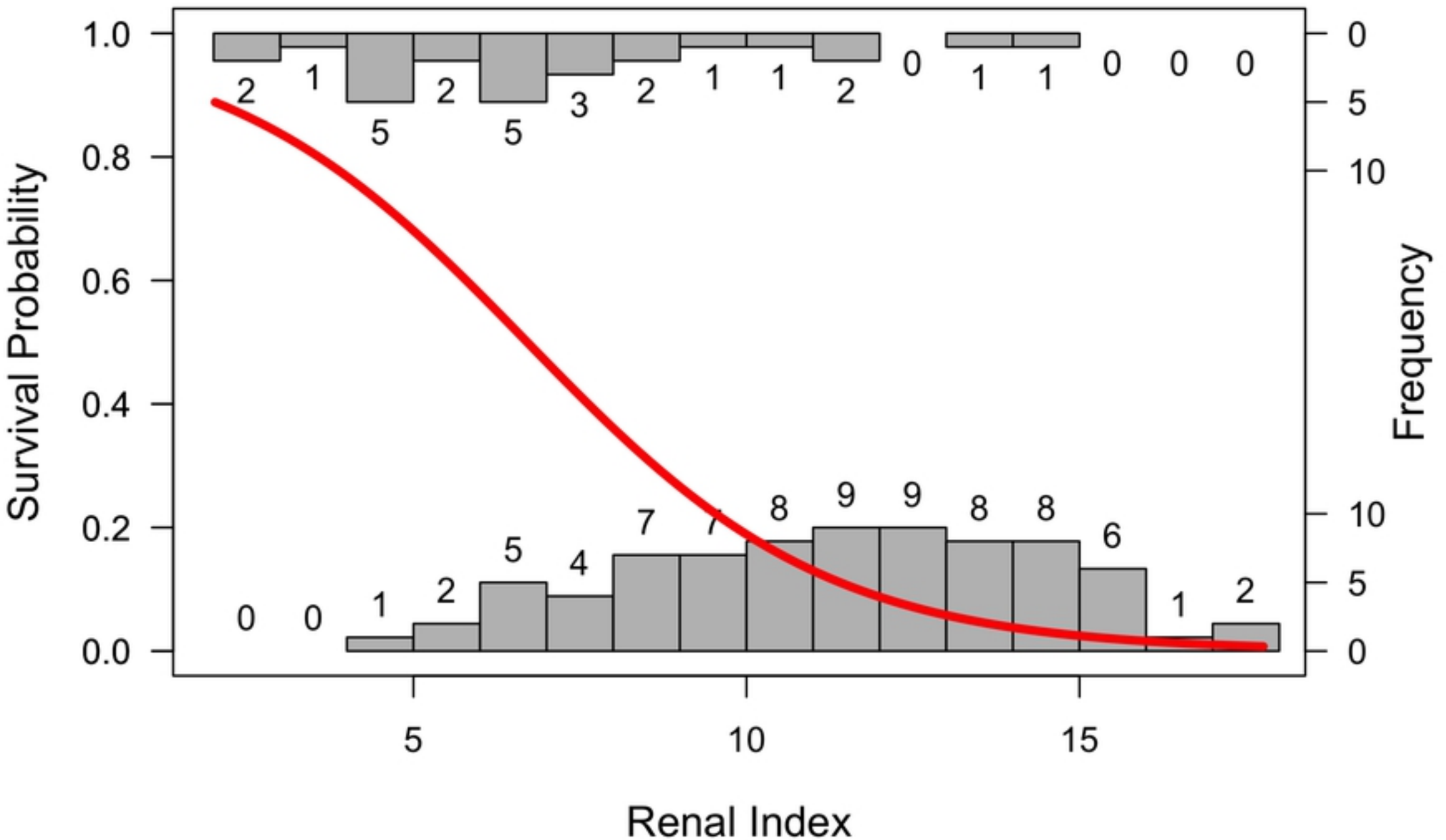
Survival Probability
1.0
0.8
0.6
0.4
0.2
0.0
Frequency
0
5
10
10
5
0
2 1 5 2 5 3 2 1 1 2 0 1 1 0 0 0
0 0 1 2 5 4 7 7 8 9 9 8 8 6 1 2
5
10
15
Renal Index

Figure 5

Figure 6

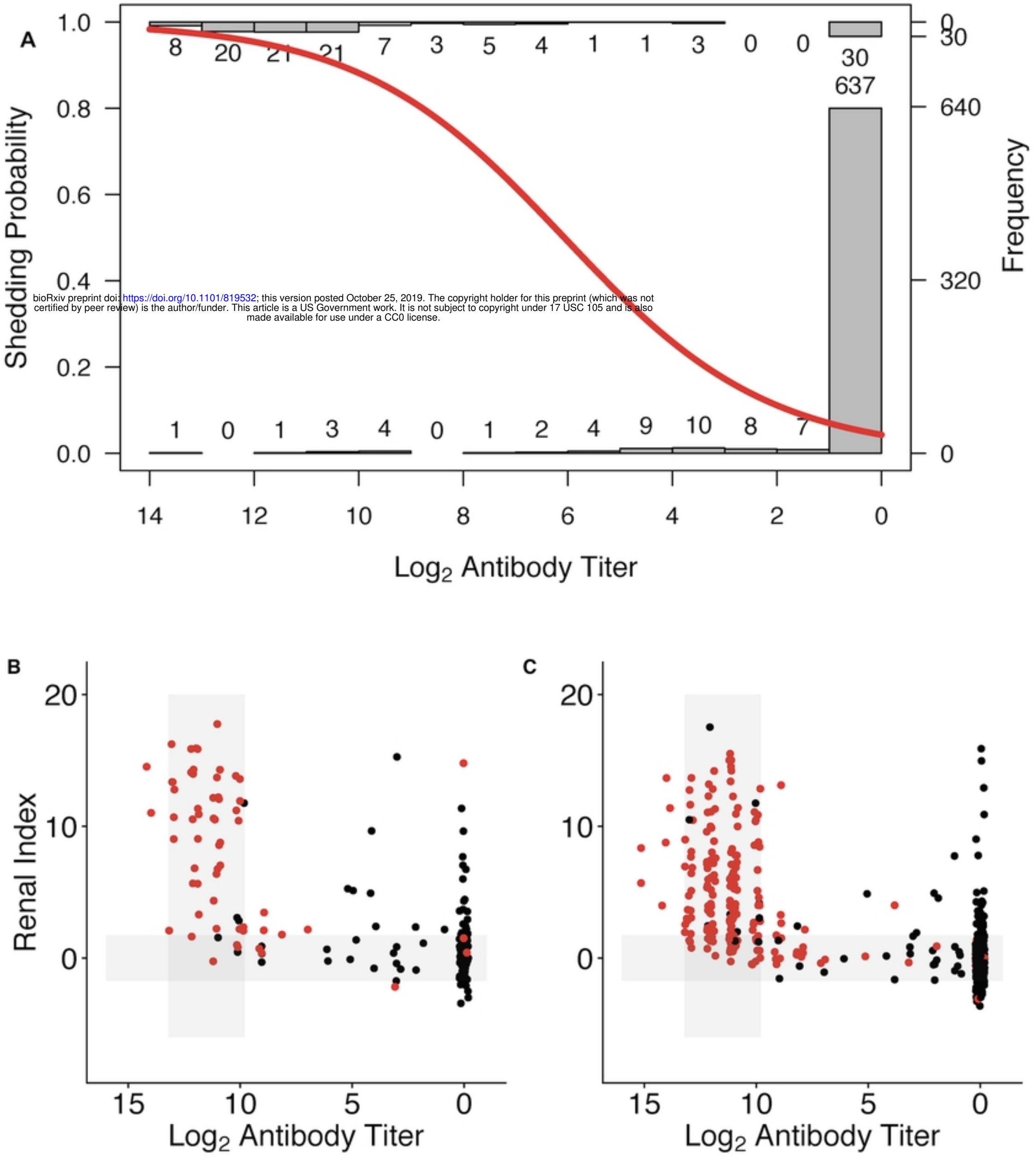


Figure 7