Development and Validation of CystiAgent: an Agent-Based Model to Investigate Transmission and Control of *Taenia solium* in Peru

By

Ian Wright Pray

A DISSERTATION

Presented to the PhD in Epidemiology Program
School of Public Health
Oregon Health and Science University and Portland State University
in partial fulfillment of the requirements for the degree of:

Doctor of Philosophy

May 15, 2019

Abstract

Background: The pork tapeworm (*Taenia solium)* is a parasitic helminth that imposes a major health and economic burden on poor rural populations around the world. Larval infection of the human brain (neurocysticercosis) is responsible for one-third of seizure disorders in endemic areas and economic losses from tainted pork are substantial. As recognized by the World Health Organization, a key barrier for achieving control of *T. solium* is the lack of an accurate and validated simulation model with which to study transmission and evaluate available control strategies.

Methods: This dissertation research represented a three-part effort to develop and validate a *T. solium* simulation model to fill this need. First (Chapter 3), I conducted a series of field studies in rural Peru to acquire data necessary to build the model; this included a GPS tracking study of 108 free-roaming pigs, and surveys of pig husbandry and human sanitation practices. Second (Chapter 4), I developed the novel agent-based model, called "CystiAgent", and conducted sensitivity analyses to evaluate its functionality. Third (Chapter 5), I validated the model predictions against data from two large trials conducted in Peru.

Results: In the GPS tracking study, we detected a significant seasonal variation in pig roaming ranges, and generated precise estimates for the size of pig home ranges, the use of corrals to contain pigs, and open defecation practices among humans, all features that were used to develop and parameterize the CystiAgent model. Sensitivity analysis of the model in Chapter 4 revealed that the model functioned as expected, and identified key features of the model that should be addressed with future research and/or local calibration in order to reduce uncertainty in the

model. Validation of the model in Chapter 5 found that CystiAgent was able to accurately replicate baseline levels of transmission in simulated villages and performed well when control strategies were applied to the model, although intervention effectiveness was overestimated compared to the observed data in some villages.

Impact: The CystiAgent model developed here represents an important new tool for advancing control of *T. solium* in Peru and around the world. CystiAgent is the first *T. solium* transmission model to accurately represent key spatial and behavioral features of transmission, and is the first model to be validated with longitudinal data, a fact that will provide credibility and interpretability to its predictions in the future. While continued work will be needed to improve this model, this research provides an important foundation against which future models can be developed and compared. A refined CystiAgent model will have the potential to rapidly compare strategies for *T. solium* control and allow for promising strategies to be prioritized for evaluation in field trials.

Acknowledgements

In this section, I would like to acknowledge the amazing group of people that made it possible for me to complete this dissertation work. While my name will appear on the title page, the truth is that not a single word of this dissertation could have been written without the dedication and hard work of my colleagues, collaborators, staff, mentors, and predecessors – a list that could easily reach into the hundreds of people from the U.S. and Peru. Foremost among these contributors is Dr. Seth O'Neal, my research mentor, advisor, and dissertation committee chair at the OHSU-PSU School of Public Health. It would be impossible for me to overstate Dr. O'Neal's contribution to my development as an epidemiologist, scientist, and a person. He consistently prioritized my learning and development over his own research goals, not hesitating to sit down with me for hours at a time to sort through the issue of the day, and ensuring that every project I worked on contributed in some way to my training. He also recognized the importance of field work in an epidemiology training program, encouraging me to go to Peru to pursue the challenges or living and working there. Lastly and most importantly, he relentlessly believed in me and this project even when I doubted myself.

The directors and staff at the Center for Global Health in Tumbes, Peru were also instrumental in turning this ambitious project from a paper proposal into a real live study. These people include Claudio Muro, Ricardo Gamboa, Percy Vilchez, Luz Maria Moyano, and Ruth Atto, who welcomed my wife and me to Tumbes and found time to prioritize my project despite their busy schedules. Of course, if I had not been able to catch any free-roaming pigs, none of this work would have been possible, so I owe a special thanks to Juan Jose, Alex, Harold, and William, who helped me in the field with the dirty work of catching pigs and mapping human

feces. Finally, The Cysticercosis Working Group in Peru (CWGP) is a collaborative network of dedicated researchers who were instrumental in providing the scientific foundation upon which my research was built. I drew heavily upon data and prior work from Drs. Hector Garcia, Armando Gonzalez, and Robert Gillman to complete this project.

My dissertation committee – Drs. William Lambert, Wayne Wakeland, and William Pan – supported me through all phases of this project and I am extremely grateful for their contributions and dedication. From the initial study design and proposal through analysis and final writing, they found time to review drafts and answer my questions. Perhaps most importantly, they showed confidence and support for me to meet my ambitious deadlines even when it required a significant amount of blind faith.

I also would not be the scientist I am today without help from the passionate and hard-working folks at the OHSU-PSU School of Public Health. This includes Dr. Lynn Marshall, director of the Epidemiology PhD program, and countless other dedicated professors, staff, and administrators who taught me how to be an epidemiologist. I also owe thanks to my peers in the epidemiology program and our local crew of mentees in Dr. O'Neal's lab. They have been on this journey with me for the past four years, and it has been an amazing experience to learn from and grow with this incredible and passionate group of epidemiologists.

A number of institutions were responsible for the financial backing of this research, and I am grateful for their commitment to me and this project. The Fulbright Program (U.S. Department of State) is dedicated to supporting international education and funded my research experience in Peru. I was also directly supported by the National Institute of Neurological Disorders and Stroke, grant number NIH R01-NS080645, and the OHSU-PSU School of Public Health. Key datasets used in this project were funded by the NIH grant listed above and the Bill and Melinda Gates Foundation.

Finally, none of this would have been possible without the love and support of my family. My parents, John and Lee, and my sister, Sarah, always believed in me and supported my

curiosity from a young age. The person who undoubtedly sacrificed the most for this dissertation, however, and to whom I owe a lifetime of gratitude is my loving wife, Hannah. She stuck with me through the long hours, absenteeism, and ups and downs of the past four years. Without her unwavering support over these past four years, I simply would not have made it this far.

Thanks to all of you, and to anyone else who encouraged or supported me throughout these trying four years, even if I neglected to mention you here. You are all inspiring, and I hope to someday honor your contributions by passing on your passion, leadership, and mentorship to the next generation of public health researchers.

Contents

List of Tables

List of Figures

List of Abbreviations

ABC	Approximated Bayesian Computation
ABM	Agent-Based Model
ANOVA	Analysis of Variance
CEDP	Cysticercosis Elimination Demonstration Project
CoAg-ELISA	Copro-Antigen Enzyme-Linked Immunosorbent Assay
CWGP	Cysticercosis Working Group in Peru
EITB	Enzyme-Linked Immunoelectrotransfer Blot
GPS	Global Positioning System
HT	Human Taeniasis
LHS-PRCC	Latin-Hypercube Sampling Partial Rank Correlation Coefficient
LoCoH	Localized Convex Hulls
MAE	Mean Absolute Error
mAh	milliamp-hours
MDA	Mass Drug Administration
MRI	Magnetic Resonance Imaging
NCC	Neurocysticercosis
NIH	National Institutes of Health
NSM	Niclosamide
NTD	Neglected Tropical Disease
OFZ	Oxfendazole
PAHO	Pan-American Health Organization
PC	Porcine Cysticercosis
RST	Ring Strategy Trial
WHO	World Health Organization

Preface

Five years ago, in 2014, the World Health Organization (WHO) convened a meeting of cysticercosis researchers from affected countries at its headquarters in Geneva to discuss progress towards controlling the disease. At the time, a large-scale trial of cysticercosis elimination strategies funded with millions of dollars from the Bill and Melinda Gates Foundation had recently been completed with promising results, and the prospects for control and elimination of cysticercosis seemed to be growing. The study included over 100,000 people and demonstrated that with intensive application of drugs to treat human taeniasis and porcine cysticercosis, transmission could be eliminated in an endemic region of Peru.

Motivated by this success, and the recent release of a highly effective vaccine to prevent porcine cysticercosis, leaders at the 2014 meeting in Geneva seized momentum to secure commitments from five highly impacted countries – Brazil, Madagascar, Ivory Coast, Vietnam, and China – to have large-scale control programs in place by 2020. The ambitious timeline was based on WHO's "Roadmap for Implementation," which, two years earlier, in 2012, laid out a series of milestones for 17 neglected tropical diseases (NTDs). For cysticercosis, the targeted milestones included having validated strategies available for cysticercosis control by 2015, and scaling these strategies in selected countries by 2020. The selected countries were joined in this commitment by the United Nations Food and Agricultural Organization, the World Organization for Animal Health, the WHO, and the pharmaceutical industry, all of which pledged their support to assist these countries in reaching their ambitious target.

An important goal that was identified in this meeting was the development of a validated and accurate transmission model. This was a tool that could be used to compare available combinations of interventions and identify those most likely to be effective and feasible when

scaled to national programs. A preliminary version of one existing *T. solium* transmission model was presented to leaders at this meeting, and reactions were enthusiastic but cautious. If key assumptions and uncertainties could be improved, they concluded, this model would be enormously helpful in meeting the 2015 and 2020 targets for *T. solium* control.

Five years removed from that seminal meeting in Geneva, *T. solium* research has carried forward and progress has been made in our understanding of transmission and control. However, there are no national programs for cysticercosis control in operation and considerable uncertainty remains regarding which strategies will be most effective and feasible for local or regional programs. While more time-consuming than originally hoped, efforts to develop *T. solium* models for this purpose have been gaining momentum. Three novel *T. solium* transmission models have been published in the past three years that have attempted to compare and evaluate available control strategies. Each of these models adopted slightly different structures and target populations, but all have one thing in common – none have been validated against observed data. As a result, the assumptions and uncertainties originally acknowledged at the WHO meeting are still major concerns for these models, and trust in their recommendations is tenuous among field researchers and policy-makers.

Until about two years ago, neither I (Ian Pray) nor my research mentor, Seth O'Neal, were involved in this dialogue. We had other research interests, and our knowledge of infectious disease modeling was sparse, at best. Yet, here was this immense need to develop a validated *T. solium* model, and no other group seemed to have the data or experience to build a useful and accurate model. We realized that we were uniquely situated to carry the torch on this effort to build a better model. Based on our connections with the Cysticercosis Working Group in Peru, we had access to data from two of the largest prospective trials for *T. solium* control ever attempted, data that could be used to validate the model, and we had a field site in Peru that could be enlisted to carry out the experiments and observational studies needed to parameterize the model. With access to the data and logistical support necessary to carry out such a modeling

effort, the only remaining piece was a sacrificial PhD student willing to go down to Peru to collect the data and spend the time to develop the model. Obviously, that student ended up being me, and the curious question, “What would it look like if we tried to build our own model?” turned into the words and pages that follow.

Chapter 1: Introduction and Research Aims

1.1 Introduction

Cysticercosis, caused by the pork tapeworm (*Taenia solium*), imposes a major health and economic burden on low-income rural populations in Latin America, Africa, and Asia. Larval infection of the human brain (neurocysticercosis, NCC) is responsible for 1.3 million [1] and 3 million [2] cases of acquired epilepsy in Latin America and Africa, respectively, and accounts for one third of all epilepsy diagnoses in these regions [3–7]. Further, cysticercosis exacts a severe economic toll that stifles growth in poor rural areas. The costs of endemic transmission include agricultural losses from infected pork, which amount to $164 million annually in Latin America [8], and medical expenses and lost productivity from the chronic management of neurological disease, which surpass those of all other foodborne parasites [9,10].

The World Health Organization (WHO) recently called for cysticercosis control strategies to be scaled and implemented in several countries by 2020 [11], yet relatively few prospective trials that evaluate such strategies have been attempted, owing largely to the high cost and time required to do this work. This gap in research has led to a limited base of evidence from which to make critical and time-sensitive recommendations for cysticercosis control, and new approaches are urgently needed to identify and prioritize optimal strategies for control.

One tool that meets the above policy needs and would allow for the rapid and efficient evaluation of control strategies is an accurate and flexible simulation model. Such a model would allow for *in-silico* testing of a broad range of available interventions – including variations in the frequency and duration of treatment, screening, or vaccination efforts, among other options – to determine which would be most likely to succeed in different endemic settings.

While five models for cysticercosis have been previously developed [12–16], the existing models have critical gaps that limit their utility for policy-makers: 1) none account for the spatial clustering of risk that results from local roaming patterns of pigs; 2) none account for human travel and migration, which leads to rapid re-introduction of transmission after interventions have concluded; and 3) none have had their projections validated with data collected from prospective trials.

1.2 Research Aims

The objective of this dissertation research was to develop and validate an agent-based model (ABM) for *T. solium* that addresses the shortcomings of previous models outlined above. The final model, called "CystiAgent," was developed with goals of being both accurate (i.e., reflects true transmission patterns observed in nature) and flexible to allow the comparison of a wide array of control options with the potential to be applied to other endemic regions in the future. The following aims represent a comprehensive effort to design, parameterize, and validate the CystiAgent model.

Research Aim 1 (Chapter 3): investigate the spatial dynamics of *T. solium* transmission in northern Peru through evaluation of the roaming patterns of pigs and their contact with open human defecation sites. In order to develop an accurate ABM for *T. solium* transmission, it is necessary to have detailed quantitative data on the behavioral, environmental, and biological factors that interact to sustain transmission in endemic villages. One of these key factors is the distance and area pigs cover while foraging, and their frequency of contact with human feces during roaming. This research aim consisted of GPS tracking of free-roaming pigs and a survey of human defecation practices in three endemic villages of northern Peru. Findings from these activities were directly used to define parameter values in the CystiAgent model, and

also provided important insights into the spatial and behavioral dynamics of *T. solium* transmission.

Research Aim 2 (Chapter 4): conduct sensitivity analyses of the CystiAgent model in order to test its functionality and identify key areas to focus future research. CystiAgent is a complex ABM that relies on many uncertain and highly variable model parameters. Exploring the model's structure and parameterization through sensitivity analysis is an essential step in both assessing the model's function and identifying key sources of uncertainty in the model. In this aim, we subjected the model to rigorous sensitivity analyses. The process allowed us to identify a key set of parameters that would benefit from additional data to reduce uncertainty, and provided confidence that the model was ready for validation (Aim 3, Chapter 5).

Research Aim 3 (Chapter 5): validate the CystiAgent model with data from two large prospective trials. In this aim, we tested the accuracy of the final CystiAgent model by validating the model against data collected in two large prospective trials in Peru: the Ring Strategy Trial and the Cysticercosis Elimination Demonstration Project. The results of this validation provided both a definitive assessment of the model's accuracy and a starting point from which future versions of the model can be developed, re-tested, and improved.

Chapter 2: Literature Review

2.1 History of *Taenia solium*

Tapeworms are among the earliest recognized afflictions of the human body. The first written reference to tapeworm infections in humans appeared in a medical scroll from Ancient Egypt (The Ebers papyrus, circa 1500 BC), which described the condition as "snakes of the belly" and recommended treatment with pomegranate roots [17]. Later, Greek, Roman, and Arabic physicians described the morphology and symptomatology of tapeworm infections [18]. It was not until much later, however, during the early 19th century, that the connection was made between tapeworms and the cysts that were known to inhabit "measly" pork and beef. When these cysts were investigated under a microscope, they were found to harbor larval worms that resembled the adult tapeworms found in humans [19]. This connection was confirmed through a series of ethically questionable experiments conducted by a German scientist, Friedrich Küchenmeister, in the 1850s. Küchenmeister harvested larval cysts from measly pork and beef, concealed them in a raw soup, and fed the soup to convicts prior to execution. In post-mortem autopsies, he found tapeworms up to 5 feet long coiled in small intestines of the convicts [20]. In later experiments, he fed tapeworm segments (harvested from the human intestines) to pigs and cows and observed the inverse phenomenon: larval cysts developing in the musculature and organs of the animals [21]. From this, it was clear that both the beef and pork tapeworm had a two-phase life-cycle that included an adult phase in humans (a condition now known as *taeniasis*) and a cystic larval phase in their respective animal hosts (*cysticercosis*) (Fig 2.1).

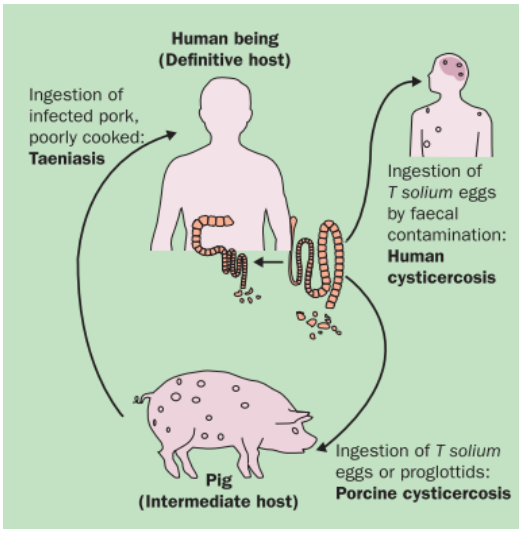


Fig 2.1. Life-cycle of *Taenia solium.* *Garcia et al.* [22].

A key discovery about the biology of the pork tapeworm, however, had yet to be made. Adult-onset epilepsy had been well-described in Roman times, but its etiology was unknown for much of human history. Autopsies in the 1500s revealed that many patients suffering from late-onset epilepsy were found to have cystic vesicles invading their brains. These were later identified as the same cystic larvae that could be found in pork, and the condition became known as *neurocysticercosis (NCC)* [23]. But where were these debilitating cystic infections coming from? The question remained a mystery until the 1930s, when physicians in London noted an unusually high incidence of epilepsy among British troops returning from overseas duty in India. Since few of these men were found to harbor intestinal tapeworms, it was concluded that they were contracting NCC simply through exposure to a contaminated environment. Just like pigs that had larval cysts spotting their organs, these British soldiers were suffering from larval infections in their brains caused by ingesting *T. solium* eggs in the fecally contaminated food and water of India [24,25].

While the above findings cemented our understanding of the basic components of the *T. solium* life-cycle, they represented just the beginning of a new line of research and discovery that is ongoing. Today, epidemiologic research aims to investigate the multi-dimensional causes of

persistent endemic *T. solium* transmission, and seeks to identify improved methods for controlling transmission and preventing current and future generations from continuing to suffer from this debilitating disease. In the following pages, I will explore the landscape of the current research on *T. solium* by first reviewing the biology and transmission of *T. solium*, and then describing its epidemiology and control. I will conclude by introducing infectious disease modeling as a powerful new tool that has the potential to drive the cysticercosis research agenda forward – towards more effective approaches to control and the ultimate goals of disease elimination and eradication.

2.1 Biology and transmission of *T. solium*

2.1.1 Taxonomy

The pork tapeworm (*T. solium*) is member of the cestode class (*Cestoda*), a class of segmented flatworms that parasitize vertebrate hosts. More specifically, the pork tapeworm is a member of the *Taeniidae* family, a unique family of cestodes that have two life-stages – an adult tapeworm that invades the small intestines of a primary mammalian host, and a metacestode (i.e., larval cyst) that colonizes tissue of an intermediate host. Among cestodes, taeniids are the family with greatest medical and veterinary relevance, with membership that includes two species of pork tapeworms (*T. solium* and *T. asiatica*), the beef tapeworm (*T. saginata*), and a canine tapeworm that causes cystic hydatid disease in humans (*Echinococcus granulosus*). Only the beef and pork tapeworms rely on humans as definitive hosts of the adult-stage worms, and only *T. solium* has the potential to cause NCC[†], the most clinically significant effect of taeniid transmission.

[†] *T. asiatica* has been proposed as an additional source of cysticercosis (and NCC) in humans, but further research is needed to confirm this. See [171] for more.

2.1.2 *T. solium* taeniasis

Infection by an adult *T. solium* tapeworm (taeniasis) is caused by consuming raw pork that contains *T. solium* larval cysts. When cysts reach the human intestines, the larvae evaginate from a protective bladder, and fix themselves to the upper intestinal wall (duodenum) of the human host. Cestode tapeworms are equipped with a head (known as the "scolex") that serves as a holdfast organ for this purpose (Fig 2.2A). Cestodes of the *Cyclophilladae* order (including those of the *Taeniidae* family) have a scolex characterized by four suckers, while the scolex of *T. solium* is further equipped with a ring of hooks (called a "rostellum") to aid in attaching itself to the host. The cestode tail ("strobila") is made up of a chain of identical segments called proglottids that function as hermaphroditic self-fertilizing reproductive units (Fig 2.2B). Each proglottid is equipped with sperm, ovaries, and a uterus that fills with thousands of eggs as the proglottid reaches maturity. Reproductively mature proglottids at the tail (called "gravid proglottids") contain approximately 50-60,000 fertile eggs and detach from the strobila in groups of 2 to 5, exiting the gut through the host's feces [26]. *T. solium* tapeworms achieve maturity approximately 2 months after initial infection [27], and reach a length of 2 to 4 meters, or 700-1000 proglottid segments [22]. The natural duration of *T. solium* tapeworm infection is not definitively known, but is now believed to be 2-5 years [22,28], much shorter than the reported 20-25 year lifespan of *T. saginata* infection [29]. This abbreviated lifespan of *T. solium* is supported by both age-specific prevalence data [30,31], and from Yoshino [27], who, in 1935, consumed three *T. solium* cysts and noted the passage of proglottid segments in stool that began 2 months after ingestion and persisted for 2 years and 3 months. The clinical features of *T. solium* taeniasis are mild and often go unnoticed, but may include abdominal pain, nausea, or diarrhea. While some tapeworm carriers are able to identify proglottid segments in their stool, this is rare due to the small size and intermittent release of proglottids. As a result, diagnosis of *T. solium* taeniasis is rare outside of organized screening programs.

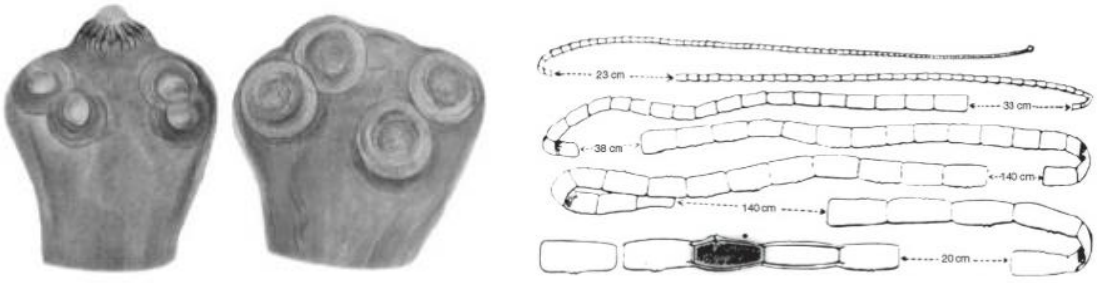


Fig 2.2. A. Scolex of *T. solium* (left) and *T. saginata* (right). Note the rostellum with hooks that can be used to identify *T. solium.* **B. Proglottid segments that make up the strobila of the tapeworm.** Larger gravid proglottids are located in the tail. *Grove* [19].

2.1.3 Porcine cysticercosis

When *T. solium* tapeworm carriers defecate outdoors, large quantities of *T. solium* eggs are released into the environment (as many as 100,000 to 300,000 eggs per day have been suggested [26,32]). *T. solium* eggs consist of an oncosphere contained within a durable outer shell (embryophore) that provides resistance from environmental conditions. The eggs (Fig 2.3A) are microscopic (30-40μm in diameter) and are indistinguishable from eggs of other *Taenia* spp. When pigs consume these eggs, either through direct coprophagia [33] or through other environmental exposures that have not yet been studied (e.g., contaminated soil or water), the oncospheres hatch in the pig gut, enter the bloodstream, and are widely disseminated within the pig. Two months after ingestion, oncospheres grow into cysts ("cysticerci") that are oval-shaped and measure approximately 4mm by 6mm (Fig 2.3B). Once established in the tissue of their intermediate host, these larval cysts, which appear as fluid-filled bladders that contain a larval worm, remain alive and viable for 1-2 years [26,34]. It is unknown whether these cysts grow preferentially in certain pig organs, but they have been commonly found in muscle (Fig 2.3C), liver, heart, and brain (Fig 2.3D) tissue of pigs [35,36].

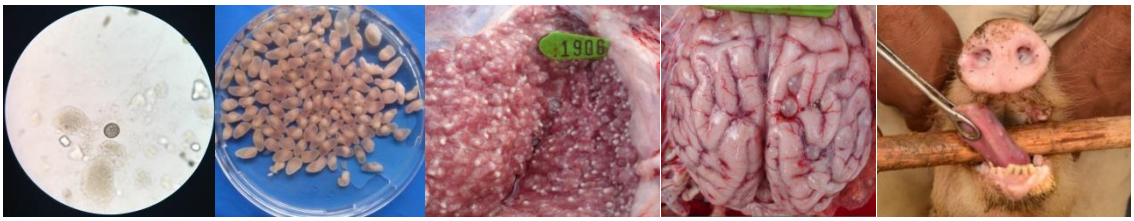


Fig 2.3. A. *T. solium* eggs (10X). **B.** *T. solium* metacestodes (larval cysts) removed from an infected pig. **C.** Pig muscle heavily infected with *T. solium* cysts. **D.** Pig brain infected with *T. solium* cysts. **E.** Pig tongue being inspected for cyst infection. *Photo credit: IWP and Center for Global Health Tumbes.*

Cyst infection in pigs has traditionally been detected by local farmers through palpating the tongue of the pig to feel for the fluid-filled cysts (Fig 2.3E) [37]. This method is still used as a non-invasive diagnostic in many epidemiological surveys [38–41], but only pigs with extremely high cyst burdens (>100 cysts) can be reliably diagnosed in this way [42–44]. Other similar methods that can diagnose these heavily infected pigs include ultrasonography [43], eyelid examination [35], and partial carcass dissection [36]. Most infected pigs, however, are found with only a small number of cysts (often fewer than 10 cysts [43]), and serological methods or full carcass dissection are needed to detect these light infections. The enzyme-linked immuno-electrotransfer blot (EITB), which evaluates pig sera for antibodies against *T. solium*, is the most commonly used diagnostic for this purpose [45]. While EITB is highly sensitive, pervasive exposure to *T. solium* eggs in the environment, dynamic immune responses over time, and the persistence of maternally transferred antibodies lead to significant false positivity and, therefore, limited utility of this assay for the diagnosis of infection in individual pigs [46,47]. Despite this, EITB remains an important tool for monitoring the effectiveness of population-level interventions, as the presence of seropositive pigs serves as a good indicator for the presence of *T. solium* eggs in the environment, untreated tapeworm carriers, and ongoing transmission in the community [48–50].

2.1.4 Human cysticercosis and NCC

Cysticercosis in humans is characterized by infection of *T. solium* larval cysts in human tissue, and occurs when humans ingest *T. solium* eggs. While human cysticercosis is a serious health concern, it does not contribute to transmission of *T. solium,* and is not a public health hazard like human taeniasis and porcine cysticercosis. The mechanism by which humans ingest eggs is not definitively known, but is thought to include fecal-oral spread through close familial contacts or auto-infection (hand-to-mouth self-infection of a tapeworm carrier) [51]. Fecally contaminated food and water have also been proposed as sources of infection, but have not been confirmed. After passing through the gut lining into the bloodstream, eggs develop into cysts that may lodge in any organ system, but are particularly pathogenic when invading the brain and central nervous system – a condition known as NCC (Fig 2.4). Due to the ability of cysts to evade immune detection in the brain, symptoms of NCC may not occur until many years after initial infection and depend on the location and intensity of infection [22]. Epileptic seizures occur in 50-80% of patients detected with cysts in the brain parenchyma [52], while headache, stroke, and intracranial hypertension are also common [22]. Death may occur from seizures or if cysts invade the ventricles or sub-arachnoid space of the brain. Diagnosis of NCC is often made when new onset of seizures occurs, and can be confirmed by serologic assays in combination with neuroimaging [53], although the sensitivity of these methods for detecting patients with a single cyst or calcified lesions from past infections is low [54]. Treatment and clinical management of NCC is complex and will not be covered in detail here. Briefly, surgical excision is considered for some cystic infections, but, more commonly, anti-parasitic drugs are taken to kill cysts while corticosteroids are used to control inflammation results from cyst death. Life-long management of chronic epilepsy and related symptoms is often required [22].

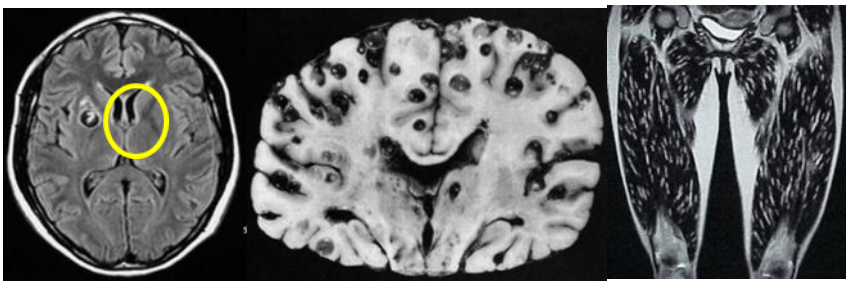


Fig 2.4. **A.** MRI of a single cyst in the parenchyma [55]. **B.** Cross-section of a brain with massive cerebral cysticercosis [8]. **C.** Massive cyst infection of muscle tissue [22].

2.3 Epidemiology and control of *T. solium*

2.3.1 Human cysticercosis and NCC

Taenia solium transmission is endemic across much of the developing world, including significant portions of Latin America, Africa, and eastern Asia (Fig 2.5). The extent of disability caused by NCC is substantial in these affected regions, where NCC is responsible for an estimated one-third of seizures and epilepsy [3–7]. Overall, as many as 1.3 million people in Latin America [1] and 3 million people in Africa [2] have active epilepsy that is caused by NCC. Beyond this, neuroimaging and seroprevalence studies strongly suggest that these active symptomatic cases are just the tip of the iceberg. Up to 20% of people living in endemic villages of Latin America have calcifications in their brains from prior cyst infections [5,7,56–58], and the seroprevalence among humans in most endemic villages is between 10 and 25% [7,30,59–62]. While the majority of individuals who are seropositive or have brain calcifications never present with neurological symptoms [56], such high levels suggest that infection is common and exposure to *T. solium* eggs in the environment of endemic rural villages is widespread.

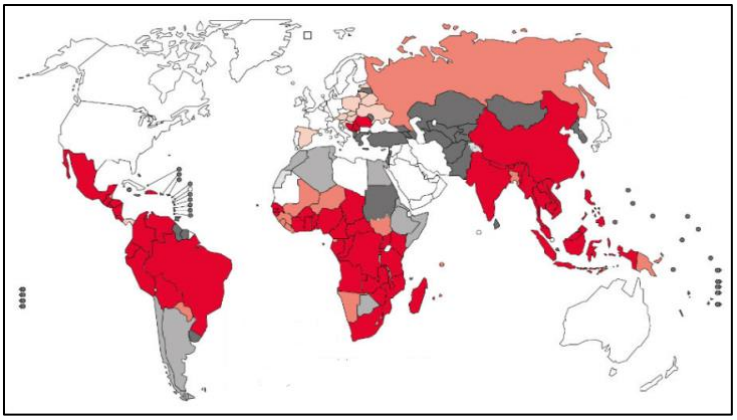


Fig 2.5. Countries where *T. solium* is known to be endemic (red), suspected to be endemic (light red), has isolated transmission (pink), possible transmission but no data (light grey), unknown (dark grey), and is not endemic (white). *WHO* [63].

In addition to the substantial burden of cysticercosis in the developing world, chronic imported cases of NCC are now a significant cause of morbidity in the U.S. due to increasing travel and immigration from endemic areas [22]. NCC-related hospitalizations in the U.S. now exceed those of all other neglected tropical diseases (NTDs) combined [64], and, in U.S. cities with large immigrant populations, as many as 10% of seizure-related emergency department visits are attributable to NCC [65]. Risk factors for human cysticercosis and NCC vary considerably by region, but typically include behaviors or environmental conditions that contribute to fecal-oral transmission of *T. solium* eggs. These include poor sanitation caused by insufficient use of latrines [66–68], deficient personal hygiene [60,69], personal history of taeniasis (which could lead to cysticercosis via auto-infection) [30,58,60,69], and family members or neighbors with taeniasis [30,70].

2.3.2 Taeniasis

While NCC has the greater clinical significance due to its debilitating health effects, taeniasis is of primary concern when considering epidemiology and control of *T. solium* due to its role as the reservoir and source of new cases of human and porcine cysticercosis. In endemic rural villages of Latin America, where the majority of surveys have been conducted, the prevalence of *T. solium* taeniasis is typically between 1-3% [31,59,60,71–74], but has been

observed as high as 6% [30,75]. Less common reports from Africa and Asia suggest that the prevalence in these regions may be closer to 5% [69,76] and reach 10-15% in some hyper-endemic pig-farming communities of Asia [77,78]. Across all endemic rural regions, cases of taeniasis have been observed to cluster in the same household [31,59,60,76,79], likely due to shared diets of infected pork, and are more common among people who report raising pigs and eating pork regularly [59,60,77]. In most regions, taeniasis risk does not concentrate among any particular sex or age group. The exception to this is in Southeast Asia, where men are more likely to have taeniasis due to cultural beliefs about raw pork consumptions, which some view as a ritualistic rite-of-passage reserved for men [80,81].

2.3.3 Porcine cysticercosis

Cystic larval infection in pigs (porcine cysticercosis) is relevant both because it perpetuates the life cycle of *T. solium* (though causing taeniasis when ingested), and because it represents a significant source of economic loss for poor rural farmers. A report from 25 years ago estimated that famers in Latin America lose $164 million annually to infected pork [8], and the number is likely much higher now . In endemic rural villages, where over half of households in a village may raise pigs for supplemental income, economic losses from discarded meat sometimes surpass a household's monthly income [82].

The economic burden of porcine cysticercosis is exacerbated by its high prevalence in endemic rural communities. Seroprevalence (by EITB) among pigs typically ranges from 30-60% in endemic villages [30,44,83–85], while 10-20% of pigs have viable cyst infection confirmed by carcass dissection [36,75]. This sharp contrast between the low prevalence of human taeniasis (1-3%) and the high prevalence of pig infection highlights the astonishing biotic potential of *T. solium* tapeworms, which, as noted, may release 100,000 or more eggs per day in the feces of their human hosts, allowing a single active tapeworm carrier to potentially cause widespread infection among pigs in a village. Environmental forces such as wind, water, flies [86], and dung beetles [87] have been proposed as mechanisms that could widely disperse *T. solium* eggs and

explain such pervasive pig infection, while roaming patterns of free-ranging pigs could also be responsible [88,89]. Despite this evidence of dispersion, the single greatest risk factor for pigs is their proximity to a tapeworm carrier – either belonging to or living within 50-100 meters of a tapeworm carrier is associated with a 5-10x increased risk of infection among pigs [74,75,90]. These conflicting spatial patterns suggest that pigs close to tapeworm carriers are likely foraging directly on contaminated feces deposited through open outdoor defecation around the household, while distant pigs may be infected through other less concentrated environmental reservoirs.

2.3.4 Prevention and control

Elimination of *T. solium* was achieved in the developed world by improving sanitation and reducing domestic pig raising. Unfortunately, these structural improvements are not likely to reach the developing world for some time, and other more immediate approaches must be considered. While the past 30 years of *T. solium* research have seen monumental strides in the development of new diagnostics, improved drugs, and a vaccine for pigs, prospective trials of these interventions in endemic communities have been few and far between. Most efforts to date have focused on either chemotherapy of human and/or pigs, or health promotion education. While no intervention has yet to achieve sustained interruption in transmission, incremental progress has been made as new strategies learn from past successes and failures. A summary of strategies that have been tested to date are described below.

Treatment of human taeniasis: The first large-scale efforts to control *T. solium* took place in Latin America in the 1980s and 90s, and attempted mass drug administration (MDA) of the human population with niclosamide or praziquantel [48,71,91,92]. Both drugs are efficacious against taeniasis in humans (single-dose efficacy of 77% [93]), but niclosamide is now the preferred choice due to a lower incidence of adverse events [94]. These early studies were instrumental in showing the potential for disease control through chemotherapeutic interventions, yet only saw minor reductions in transmission, largely because they did not address the porcine reservoir. A less common approach to MDA has been to screen the human population for

taeniasis, and provide focused treatment to positive individuals. This approach allows for follow-up and re-treatment of identified cases, improving the cure rate of the drug [93]. In such screening efforts, the most common test currently available is the enzyme-linked immunosorbent assay for coproantigen detection (Co-Ag ELISA). This stool-based assay is 98% sensitive and 99% specific to active *Taenia* spp. infections [72], but does not distinguish *T. solium* taeniasis from other tapeworm infections of the *Taenia* genus [95]. Despite evidence that this "screen-and-treat" approach is at least as effective as presumptive treatment in population-based interventions [94], it is rarely attempted due to the operational costs of collecting and processing stool samples.

Treatment of porcine cysticercosis: Cyst infection in pigs can be successfully treated with a variety of efficacious anti-helminthic drugs, but most common is a single oral dose of the drug oxfendazole. Most viable cysts die within 4 weeks of treatment, and pork is generally cyst-free and safe to consume within 3 months [34,96]. Oxfendazole has also been shown to provide pigs with immunity against re-infection for up to 3 months after treatment [97]. Due to the persistent reservoir of human taeniasis and high turnover of pigs in endemic communities, presumptive treatment of the pig population in isolation is not likely to be successful [98] and is typically used in combination with human interventions.

Vaccination of pigs**:** A highly effective vaccine was recently developed to prevent cyst infection in pigs. In field trials, the "TSOL18" vaccine achieved 99.9% [99] and 100% [100] protection against new cyst infection. The utility of this vaccine for widespread control, however, has been questioned due to the operational challenges of administering multiple booster doses in areas that are both remote and have high rates of pig turnover [80].

Combined approaches: Simultaneous treatment of humans and pigs has the potential to attack the *T. solium* reservoir in both hosts, and has been proposed as a possible means of achieving widespread elimination or eradication. This approach has been attempted four times to date –three in Peru, and once in Lao. In the first Peruvian trial, a single round of mass treatment was applied to humans and pigs and produced only a small and temporary reduction in prevalence

[49]. The Lao study was a small pilot, but observed a significant reduction in transmission after administering pig vaccination combined with human and pig MDA [101]. The second Peruvian trial was by far the largest prospective *T. solium* trial attempted to date, and represented the first effort to demonstrate the feasibility of *T. solium* elimination across an entire region [94]. The final strategy reached a population of 80 thousand people and included vaccination of all pigs in the region combined with multiple rounds of human and pig MDA. While this effort was highly effective at interrupting transmission, it was extremely costly (~$20 million), and pre-intervention levels of endemicity returned shortly after its conclusion. The last Peruvian trial was recently completed and unpublished at the time of this review [contact Seth O'Neal]. This trial included treatment in humans and pigs every 6 months for 2 years and resulted in substantial reductions in transmission. It is, however, unlikely that that this control was sustained after the program ended.

Targeted approaches: In response to the high cost and short-term gains of mass-applied interventions, recent efforts have focused on lower-cost targeted approaches that may be more sustainable and feasible for local governments. One such approach, called "ring strategy," was recently applied in Peru [50]. Ring strategy relies on the knowledge that human and pig infections are geographically clustered in endemic villages, and seeks to apply treatment to humans that live within 100 meters of a heavily infected pig (as detected through tongue inspection). In a two-year trial of ring strategy in 16 villages of northern Peru [not yet published, contact Seth O'Neal], this approach resulted in a significant reduction in transmission, while improving community participation compared to previous mass interventions [50]. Outside of ring strategy, only one other targeted approach has been attempted at the population-level. This was a campaign to focus treatment for taeniasis on school-aged children, aiming to reduce costs and improve treatment compliance by integrating the campaign into existing government de-worming programs [102]. Similar to previously described human-only approaches, this intervention achieved short-term reductions in the number of tapeworm carriers, but did not achieve sustained control, likely due to not addressing infection in the pig population.

Other approaches – education, sanitation, and pig husbandry: While chemotherapeutic interventions have comprised the majority of interventions tested to date, primary prevention approaches to *T. solium* control have also been attempted. These health promotion interventions have typically included education about disease risk factors, personal hygiene, or pig husbandry. Results from these studies have been encouraging, as all have shown strong acquisition of knowledge [103,104], and some have translated into improved health behaviors and even small reductions in prevalence of human and pig disease [105,106]. In contrast to the success of health education, efforts to construct latrines and improve community sanitation have not been as successful [107], likely due to deeply entrenched gender-norms and cultural practices surrounding defecation [108]. Similarly, efforts to encourage the corralling of pigs have not been successful. Studies have found that confinement of pigs is typically intermittent in poor endemic areas, likely due to the added requirement of providing feed for pigs that are confined [60,109,110]. Outside of these strategies, improvements in the inspection and processing of pork have been proposed as key interventions for the control of *T. solium* [11], yet neither have been implemented or evaluated systematically.

2.4 Infectious disease modeling as a tool for *T. solium* control

2.4.1 The need for a *T. solium* transmission model

Two key gaps in the current landscape of *T. solium* research are the inability to evaluate potential control strategies due the high cost of prospective trials, and a lack of knowledge surrounding the biological, behavioral, and environmental mechanisms of transmission. These gaps are caused by a lack of research funding and a shortage of dedicated researchers – barriers that are not uncommon in the area of neglected tropical disease (NTD) research. In the face of these tight budgets and limited resources, NTD research has recently turned to infectious disease modeling to identify more efficient and cost-effective approaches to disease control. As a result

of this push, the past 20 years have seen unprecedented advancements in the modeling of NTDs [111,112]. Efforts to develop models for *T. solium*, however, have been slow to develop.

Fortunately, it appears that there is renewed interested in moving this agenda forward [11,113,114], due, in part to recent work showing the potential for *T. solium* to be eliminated or eradicated [94]. On the heels of these efforts, we now find ourselves at an important juncture in *T. solium* research. For the first time, we have the knowledge, data, and technological capacity to build an accurate and definitive *T. solium* model. At the same time, questions about which control strategies to deploy in endemic regions across the globe have never been more urgent from both a policy and human health perspective.

2.4.2 Limitation of existing *T. solium* transmission models

Responding to the need and opportunity to develop a *T. solium* transmission model, five models have been published to date [12–16]. These models range from early primitive models that aimed to capture the basic life-cycle and disease states [12,13] to more advanced agent-based [14] and population-level deterministic models [15,16] that have been used to compare available control strategies and make recommendations for policy-makers. One of these models, EPICYST [15], was recently presented to WHO and collaborating health ministries with the goal of meeting WHO's 2020 goal to identify validated *T. solium* control strategies [11].

Despite use of these models in policy-making, a report by WHO stressed the considerable uncertainty surrounding the models, and cautioned that outputs from the current models should be considered preliminary [11]. WHO has highlighted three key concerns that limit the accuracy and utility of existing models. First, model validation, or comparing model predictions to observed outcomes, is widely considered a pre-requisite for real-world use of a model, as it provides assurance that the structure, parameterization, and predictions of the model are correct [115]. None of the existing *T. solium* models have been validated with data from field interventions. Second, all existing models rely on the assumption that modeled populations mix homogenously and individuals in the population share uniform risk of infection. Numerous studies have shown

that *T. solium* transmission violates this "homogenous mixing" assumption by observing clusters of infection among certain high-risk households [74,75,90,116], and showing heterogeneities in pig roaming that lead to non-uniform infection risk among pigs [88,89]. Failure to account for key heterogeneities within populations often leads to overestimation of intervention effectiveness [117] and an inability to evaluate promising strategies that target high-risk clusters [50,118]. Finally, all existing models assume closed and static populations. Continuous travel of humans between villages and regions is known to cause rapid re-introduction of pathogens into previously controlled areas [94,119,120]. No current *T. solium* model accounts for travel or re-introduction. This is likely to result in unrealistic predictions for achieving control and elimination milestones [121].

2.4.3 CystiAgent: a novel agent-based model for *T. solium*

In order to address the above limitations and meet the need for an accurate and flexible *T. solium* transmission model, the objective of this dissertation was to develop and validate a novel agent-based model (ABM) called "CystiAgent." CystiAgent is the first *T. solium* model to incorporate an agent-based (i.e., individual-level) structure that includes a spatial platform to account for pig movement and non-uniform risk distribution, a key shortcoming of existing models. CystiAgent is also the first to include an open population structure that will provide an accurate reflection of human travel patterns and the potential for disease re-introduction. With respect to the evaluation of control strategies, the CystiAgent prototype is equipped to simulate the standard suite of interventions tested in prior models. Due to the novel spatial structure of CystiAgent, it is also the first model with the capacity to evaluate geographically targeted strategies (e.g., ring screening/treatment) that have shown promise in field trials [50]. Finally, predictions made by CystiAgent were tested against data from two of the largest prospective *T. solium* interventions ever attempted, a step that provided an important evaluation of model accuracy. No prior *T. solium* models have been validated in this way.

Chapter 3: Seasonal patterns in risk factors for *Taenia solium* transmission: a GPS tracking study of pigs and open human defecation in northern Peru

Ian W. Pray[1]*, Claudio Muro[2], Ricardo Gamboa[2], Percy Vilchez[2], Wayne Wakeland[3], William Pan[4], William E. Lambert[1], Hector H. Garcia[2,5], and Seth E. O'Neal[1,2], for the Cysticercosis Working Group in Peru.

[1]School of Public Health, Oregon Health & Science University and Portland State University, Portland, Oregon, United States

[2]Center for Global Health Tumbes, Universidad Peruana Cayetano Heredia, Tumbes, Peru

[3]Systems Science Program, Portland State University, Portland, Oregon, United States

[4]Global Health Institute, Duke University, Durham, North Carolina, United States

[5]School of Sciences, Department of Microbiology, Universidad Peruana Cayetano Heredia, Lima, Peru

*Corresponding author

3.1 Abstract

3.1.1 Background

Taenia solium (cysticercosis) is a parasitic cestode that is endemic in rural populations where open defecation is common and pigs are allowed to graze on human feces. The purpose of this study was to examine the roaming patterns of free-range pigs, and identify areas where *T. solium* transmission could occur via contact with human feces. We did this by using GPS trackers to log the movement of 108 pigs in three villages of northern Peru. Pigs were tracked for 6 days each, and tracking was repeated in the rainy and dry seasons. Maps of pig ranges were analyzed for size, distance from home, land-type, and contact with human defecation sites, which were assessed in a community-wide defecation survey.

3.1.2 Results

Consistent with prior GPS studies and spatial analyses, we found that the majority of pigs remained close to home during the tracking period and had contact with human feces in their home areas – pigs spent a median of 79% (IQR: 61-90%) of their active roaming time within 50 meters of their homes, and 60% of contacts with open defecation within 100 meters of home. Extended away-from-home roaming was predominately observed during the rainy season; and, overall, home range areas were 61% larger during the rainy season compared to the dry season (95% CI: 41-73%). Both home range size and contact with open defecation sites showed substantial variation between villages, and contact with open defecation sites was more frequent among pigs with larger home ranges, and pigs living in higher density areas of their village.

3.1.3 Conclusions

Our study builds upon prior work showing that pigs predominately roam and have contact with human feces within 50-100 meters of the home, and that *T. solium* transmission is most likely to occur in these concentrated areas of contact. This finding, therefore, supports

control strategies that target treatment resources to these areas of increased transmission. Our finding of a seasonal trend in roaming ranges may be useful for control programs relying on pig interventions, and in the field of transmission modeling, which require precise estimates of pig behavior and risk.

3.2 Background

Cysticercosis, caused by the pork tapeworm (*Taenia solium*), imposes a major health and economic burden on rural populations in Latin America, Africa, and eastern Asia [1,2]. Humans acquire the intestinal tapeworm infection (taeniasis) by consuming larval cysts that may be present in raw or undercooked pork. Adult tapeworms reside in the human gut, and may expel tens of thousands of infectious eggs each day in the host's feces [26,32], which contaminate the environment in areas where open human defecation is common. The widespread practice of free-range pig-raising in endemic areas allows pigs to consume *T. solium* eggs in human feces and develop larval cyst infection in their muscle tissue, thus proliferating the life-cycle.

The movement patterns of free-roaming pigs within endemic communities and their contact with potentially infectious human feces are key factors that influence transmission patterns. Prior studies have found that pigs raised in the same household or within 50 meters of a human with taeniasis have substantially higher rates of cyst infection [74,75,122] and antibody reactivity [90] compared to more distant pigs. This knowledge of locally acquired *T. solium* infection has led to important advancements in control in recent years. In Peru, "Ring Strategy" has led to significant disease control by offering screening and treatment for human taeniasis to people living within 100 meters of an infected pig [50].

Although the evidence for focal transmission of *T. solium* is convincing, there are significant gaps in our knowledge of transmission that have been highlighted by prior spatial studies. Namely, past studies have routinely found infected pigs living distant from known

tapeworm carriers [75,122], and ring interventions have not completely eliminated the disease [50], as would be expected if transmission were purely focal. An improved understanding of *T. solium* transmission dynamics, including elucidation of these unexplained patterns of pig infection, would have a few key impacts on the prospects for *T. solium* control. For one, it may lead to improved intervention strategies that more effectively target treatment resources to areas of transmission risk. Second, it would provide key information for the emerging field of transmission modeling. Existing models of *T. solium* transmission have been used to compare the effectiveness of available control strategies [14,15], but have not yet had sufficient data to incorporate spatial aspects of transmission. Addressing this knowledge gap requires that we investigated the behavioral and environmental factors that produce the observed spatial patterns in transmission – chief among these are the roaming patterns of pigs and their contact with human feces present in the environment due to open defecation practices.

Having previously identified these goals, we first investigated the roaming patterns of pigs in a pilot study conducted in 2015 [88]. In that study, we used GPS trackers to map the roaming ranges and contact with human feces for 37 pigs in two small villages of northern Peru. That study helped to validate the size of 100-meter rings used in Ring Strategy, but was limited by a short tracking period (48 hours), a small sample of pigs from only two villages, and tracking during the rainy season only, all factors that could have led to biased or imprecise estimates.

In the present study, we set out to further investigate the roaming patterns of pigs in this region with the goal of improving upon the limitations of our pilot study. Specifically, this study expanded to three new villages in northern Peru, included more pigs (n=108), a longer tracking period (six days), and tracking in both the rainy and dry seasons.

3.3 Methods

3.3.1 Selection of study villages and tracking seasons

Three villages in the northern Peruvian region of Piura participated in this study. We selected these villages (referred to as villages "A", "B", and "C" here) because they were generally representative of rural villages in the region, had an adequate number of households that raised free-roaming pigs, and were participating in a concurrent cysticercosis control study that provided up-to-date census information [123]. The period of GPS tracking referred to as "rainy-season" tracking took place in the study villages in April 2018, which corresponds to the end of the rainy season (December-April) and is characterized by intermittent rain and abundant wild fruits and foliage. "Dry-season" tracking took place in the same villages in August 2018, a period characterized by cool and dry weather with very little green foliage.

3.3.2 Sample size

The sample size of pigs for this study was designed to explore differences between home-range areas by season (two-sided, $\alpha = 0.05$). Our chosen sample size of 120 pigs (20 pigs per village per season) corresponded to an 80% power to detect a 35% difference in median home range by season in the full sample, and 54% seasonal difference within each village stratum. Calculations were based on mean and variance results from our pilot study in this region [88].

3.3.3 Selection of pigs

All households in participating villages were approached for inclusion in the study, and were eligible if they reported raising free-roaming pigs. At consenting households, pigs were eligible for GPS tracking if they were not regularly tied or enclosed in a corral, were at least two months old, were not pregnant or sick, and were not planned for slaughter in the next seven days. We attempted to enroll one pig from each consenting household. If multiple pigs could be captured from one household, we enrolled the pig that fulfilled an age-stratified sampling scheme. For dry season tracking, we enrolled the same pigs that participated in the rainy season when

possible. If this pig had been sold or slaughtered, we selected a pig from the same household with preference towards pigs that were the same age as the previously tracked pig.

3.3.4 GPS tracking of pigs

The GPS loggers we used for this study ("i-GotU GT-120", MobileAction Technology, New Taipei City, Taiwan) were programmed to record the GPS coordinates of a pig's location every 60 seconds. In order to last the planned 6-day roaming period at this logging frequency, we replaced the original 230 milliamp-Hours (mAh) batteries with 3.7-volt 2000mAh lithium-ion batteries (AdaFruit, New York, New York) in all devices used. After each pig was captured, the modified GPS logger was placed in a waterproof case (HPRC 1100, Plaber, Vicenza, Italy), and secured to the nape of the pig using a custom harness made of nylon webbing (Fig 3.1). All study pigs from each village were tracked over the same 6-day period. During this period, study staff returned to each enrolled household daily to check on pigs and adjust harnesses if necessary. At the end of the 6-day period, the GPS devices were removed and the spatial data were downloaded for analysis.

Fig 3.1. GPS devices placed in waterproof cases and secured to harnesses for tracking.

3.3.5 Household defecation survey

In addition to tracking pigs, we conducted household surveys to assess human defecation practices in the study villages. For this, we visited all households during the rainy season, and asked available adult residents whether their family owned a latrine/indoor bathroom or members

of their family practiced open outdoor defecation. If an outdoor area was indicated, we searched for evidence of recent defecation (e.g., feces or soiled paper) and used a handheld GPS receiver (GeoExplorer II; Trimble, Sunnyvale, CA) to record a GPS point at that location. For both latrines and outdoor defecation areas, household respondents were asked to rate their family's frequency of use between "Never", "Sometimes" or "Always". Finally, study teams logged the locations of roads, paths, and streams in the community, and inspected each for evidence of open human defecation. Study personnel were assisted in this effort by local community leaders, who guided teams to known communal defecation sites in each village.

3.3.6 Mapping and statistical analysis

All data were analyzed using R version 3.2 (The R Foundation for Statistical Computing; www.r-project.org), QGIS version 2.18 (Open Source Geospatial Foundation Project, http://qgis.osgeo.org), and Stata version 13.1 (StataCorp; College Station, TX). For spatial analyses, all spatial layers were projected with a Universal Transverse Mercator Zone 17S projection. Because obstruction of the satellite signal occurred intermittently during pig tracking, it was necessary to remove outlying points in post-processing. To do this, we removed points that were delayed > 10 seconds (suggesting signal obstruction), points for which the detected speed was greater than 3 m/s, and points with less than a 20 degree angle between the prior and succeeding GPS locations, features unlikely to be produced by natural pig movement. On average, we removed 3.1% of the total points logged for each pig due to suspected error. Additionally, in order to avoid bias due to the stress of the chase and capture of pigs, we removed the first hour and final 15 minutes of tracking time, as well as points that were recorded before, during, and after any necessary harness adjustments.

In order to create maps that represented the *active* foraging time for pigs, when they are most likely to consume human feces, we further restricted the GPS points included in the analysis by two factors. First, we excluded points taken between 10pm and 4am, a time in which most range maps showed inactivity for pigs, and second, we excluded points for which the GPS

coordinates did not change from the preceding point, suggesting inactivity. We validated this method of selecting for active foraging time by directly observing the behaviors of a subset of study pigs (n=9) in the field. For these pigs, which were each observed for 12 day-time hours, we found that removing repeat points successfully eliminated non-foraging rest-time with a sensitivity of 38% and specificity of 96%. Overall, these additional filters reduced the total number of GPS points used for each pig from an average of 7,727 total points to 4,569 active points, a 37% reduction.

After obtaining final datasets for each pig, we analyzed roaming ranges using the "LoCoH" (Localized Convex Hulls) Homerange Analysis Algorithm for R [124,125]. A detailed description of the LoCoH algorithm can be found elsewhere [126]. Briefly, we used the *a*-nearest-neighbors LoCoH method (*a* for *adaptive*), which is a non-parametric mapping algorithm that creates convex polygon hulls around each GPS point based on a flexible number of nearest-neighbor points. The *a*-method uses fewer nearest-neighbor points to constructs hulls in less dense areas of the range, thus avoiding the problem of large polygons forming in sparsely occupied areas. We found that the algorithm produced optimal roaming areas when the "*auto-a"* function required a minimum of 95% of points to form polygons with 30 nearest neighbors. The output of the LoCoH algorithm produced maps of each pig's range that identified three areas based on specified isopleth cut-off values. As suggested by the algorithm developers [126], the "core range" represents the densest 50% of a pig's range, the "home range" is the densest 90%, and the "maximum range" is the area that contained 100% of the convex hulls (Fig 3.2).

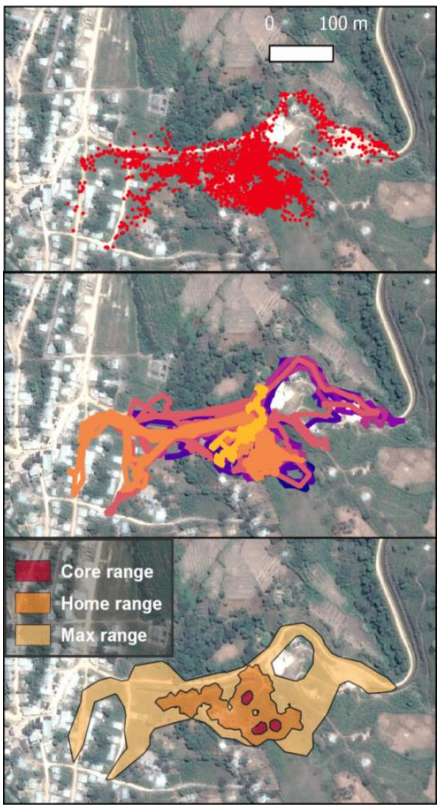


Fig 3.2. Top: map of raw GPS points from a single pig (Village B). Middle: line map of same pig's roaming pattern with each color representing a unique day of movement. Bottom: final LoCoH map of same pig's range with colors representing Core (50%), Home (90%), and Maximum (100%) range levels. *Satellite images from Google Satellite Hybrid extension for QGIS. Last update, April 05, 2017.*

In order to analyze pig roaming ranges with respect to land features and open defecation areas, we created detailed vector maps for each study village. For this, Google Earth satellite images[1] were overlaid with manually logged household and road layers to categorize village land into one of four mutually exclusive land-types: peri-domestic, roads/paths, farmland, and vegetation. Peri-domestic areas were formed by generating 20-meter buffers around household coordinates and merging the areas surrounding contiguous households and common areas (e.g., school, recreational fields, etc.); roads and paths were manually logged in the field, and enhanced with a 4-meter buffer in post-processing; farmland was assigned in post-processing by digitizing visible fence-lines that contained discernable rows of crops; all remaining areas not fitting these

[1] Images from Google Satellite Hybrid extension for QGIS. Last update: April 05, 2017. Map location: 4°38'12.84"S, 79°59'29.87"W.

categories were classified as vegetated – these remaining areas were composed of undeveloped land with sparse tree cover, bushes, and streams.

We processed LoCoH maps with respect to these base layers in order to extract a variety of roaming outcomes. These included the total area of core, home, and maximum LoCoH ranges, the proportion of tracking time spent in each land-type, the number of human defecation points within each level of a pig's range (core, home, and maximum ranges) and their corresponding land-types, and distance of each GPS point to the pig's household, which was used to determine the proportion of time spent within 50, 100, 150, and 200 meters of home.

Roaming outcomes were first analyzed descriptively, and were then analyzed for associations with pig-, household- and village-level predictors. These predictors included pig age (in months), sex, household herd size, household density (number of neighboring households within 100 meters), village of residence, and tracking season. These predictors were used to create a variety of multivariable models for pig roaming: ordinary least squares regression models for the log-area of core, home, and maximum ranges, negative binomial models for the number defecation points inside pigs' home, and maximum ranges, and a logistic regression model for the presence of at least one open defecation site within a pigs' core ranges. Predictors and interactions were retained in either model if they were significant ($p < 0.05$) when added in stepwise procedure. Because of similarities in the results of our models for core, home, and maximum ranges, only the results of the two home-range analyses will be presented here, but all models and corresponding coefficients can be found in Appendix A.

3.4 Results

3.4.1 Village and household characteristics

All three study villages are rural communities where small-holder farming is the primary economic activity and raising free-roaming pigs is common practice. Between 53% and 70% of

households reported raising pigs, and only 5% - 29% of those pig-owners reported always corralling their pigs (Table 3.1). Despite similar population sizes (range: 83-95 households), the three study villages had important differences. Village A was larger, flatter, and less densely housed than the other two villages, while Villages B and C were smaller and built on steep sloping terrain. Village B was the smallest and densest village, characterized by fewer latrines, a higher rate of open defecation, and significantly more open defecation sites.

Table 3.1. Characteristics of study villages and defecation survey.

	Village A	**Village B**	**Village C**
Human population	279	250	372
Households	95	83	83
Household density (mean # households within 100m)	**6.9**	**26.1**	**11.2**
Area	1.93 km^2	0.45 km^2	0.58 km^2
Participated	77/95 (81%)	70/83 (84%)	79/83 (95%)
Latrine prevalence	74/77 (96%)	46/70 (66%)	75/79 (95%)
Open defecation*	**13/77 (17%)**	**32/70 (46%)**	**25/79 (32%)**
Total defecation sites	30 (20%)	79 (52%)	42 (28%)
Pig owners	41/77 (53%)	45/70 (64%)	55/79 (70%)
Corral prevalence	31/41 (76%)	17/45 (38%)	18/55 (33%)
Actual corral use†	**12/41 (29%)**	**6/45 (13%)**	**3/55 (5%)**

*Some houses with latrines also reported open defecation
†Corral in "good" condition and owner reports that it is used "always"

3.4.2 Pig population

We enrolled a total of 114 pigs for GPS tracking between the two seasons. Six pigs were excluded from the analysis because of a combination of device failure (n=3), lost devices (n=2), and an owner's decision to corral the pig (n=1). This led to a final sample of 108 pigs tracked; 53 in the rainy season and 55 in the dry season. Of the 53 rainy season pigs, we were able to repeat dry season tracking for 15 pigs (28%) and track a pig from the same household for 37 pigs (70%). There were no significant differences in the sex, age, or village distribution of pigs between the rainy and dry seasons Appendix A.

Pigs included in the analysis were tracked for an average of 5.4 days (range: 2.2 to 6.6 days). The targeted 6-day tracking period was incomplete for 21 (19%) of the 108 pigs analyzed.

Reasons for incomplete tracking included pre-mature battery death or device failure (n=16), owner's decision to withdraw (n=4), and pig death (n=1, unrelated to study).

3.4.3 Household distance and defecation contact

We first analyzed the amount of time pigs spent at increasing distances from their homes. In both tracking seasons, pigs spent the majority of their active time within 50 meters of their homes (medians: 74% in rainy, 85% in dry, p=0.12, Fig 3.3A). The proportion of active roaming time spent at increasing distances decreased substantially outside of 50 meters in both seasons. The median proportions of active time spent was 8.8% and 7.8% at 50-100m, 3.9% and 1.7% at 100-150m, 2.0% and 0.5% at 150-200m, and 2.1% and 0.7% at >200m, in rainy and dry seasons, respectively.

Despite spending the majority of total time very close to households, distances at which contact with human defecation sites occurred followed a slightly different pattern (Fig 3.3B). In both seasons, the majority of contact between pig ranges and defecation sites occurred between 50-100m of the household (means: 1.66 in rainy, 1.43 in dry, p=0.58), while the contact rate at increasingly distant areas from the household was similar between rainy and dry seasons, but disproportionately elevated compared to the total time pigs spent at those distances.

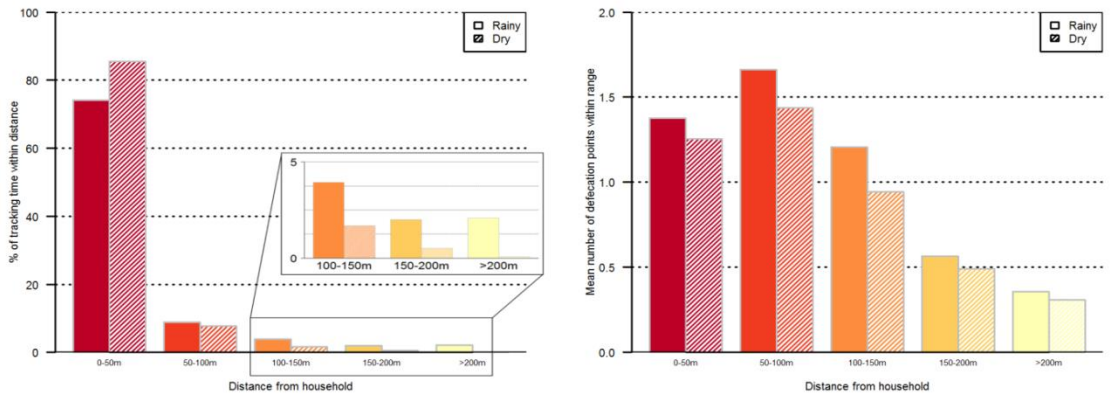


Fig 3.3. A. The median proportion of active time pigs spent at increasing distances from their households in rainy (n=53) and dry (n=55) seasons. **B.** The mean number of defecation points within the maximum LoCoH range of pigs at increasing distances from their households in rainy and dry seasons.

3.4.4 Roaming range areas

The areas of core, home, and maximum ranges are shown for all pigs in Figure 3.4. Range sizes were distributed exponentially, with the majority of pigs having maximum range areas of less than 30,000 m^2 and home range areas less than 5,000 m^2. However, a subset of pigs had substantially larger roaming areas, with maximum ranges up to 500,000 m^2 and home ranges up to 120,000 m^2.

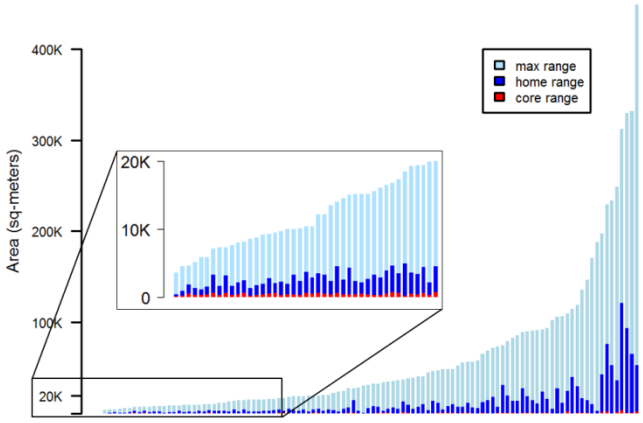


Fig 3.4. Areas of LoCoH core, home, and maximum ranges for all 108 pigs tracked.

In multivariable regression models, village of residence and season were the only variables significantly associated with log-transformed LoCoH areas. Age, household herd size, and household density all had significant bivariate associations, but became non-significant after adjustment for village and season, and pig sex was not significant in any model (Table 3.2). Across all villages, home ranges were 61% (95% CI: 47-72%) smaller in the dry season, compared to the rainy season, and there was significant variation in home range areas by village. Figure 3.5 shows representative maps of 3 pigs tracked in both seasons.

Table 3.2. Regression coefficients for home range area and defecation sites in home range. Bivariate and multivariate linear regression models for log-area of home range, and negative binomial models for the number of open defecation sites within home ranges.

	Home range area (e^{β} coefficients (95% CI))		**Defecation sites in home range (incidence rate ratio (95% CI))**	
	Bivariate	**Multivariate**	**Bivariate**	**Multivariate**
Village				
Village A	Ref.	Ref.	Ref.	Ref.
Village B	0.48 (0.30, 0.76)**	§0.47 (0.31, 0.70)**	7.06 (3.83, 13.01)**	§7.94 (4.28, 14.7)**
Village C	0.24 (0.15, 0.39)**	§0.23 (0.16, 0.35)**	1.25 (0.63, 2.49)	§1.25 (0.57, 2.70)
Season				
Rainy	Ref.	Ref.	Ref.	-
Dry	0.40 (0.27, 0.59)**	§0.39 (0.28,0.53)**	0.69 (0.39, 1.21)	-
Household density†				
≤25	0.95 (0.92, 0.97)**	-	1.03 (1.00, 1.07)	1.07 (1.04, 1.10)**
>25	1.05 (1.02, 1.09)**	-	1.03 (0.99, 1.08)	0.95 (0.93, 0.98)**
Herd size (per additional pig)	1.06 (1.02, 1.10)**	-	0.97 (0.91, 1.03)	-
Pig sex				
Female	Ref.	-	Ref.	Ref.
Male	0.78 (0.51, 1.19)	-	0.94 (0.54, 1.66)	§1.45 (1.01, 2.08)*
Pig age (per month)	1.04 (1.0, 1.08)*	-	0.98 (0.93, 1.03)	-
Log-area of home range	-	-	1.50 (1.13, 2.0)**	1.76 (1.43, 2.16)**

p-value: **<0.01, *<0.05

§Significant statistical interactions (by village) not shown (see Appendix A for full model associations)

†Number of households within 100m radius, linear spline at 25 households/100m

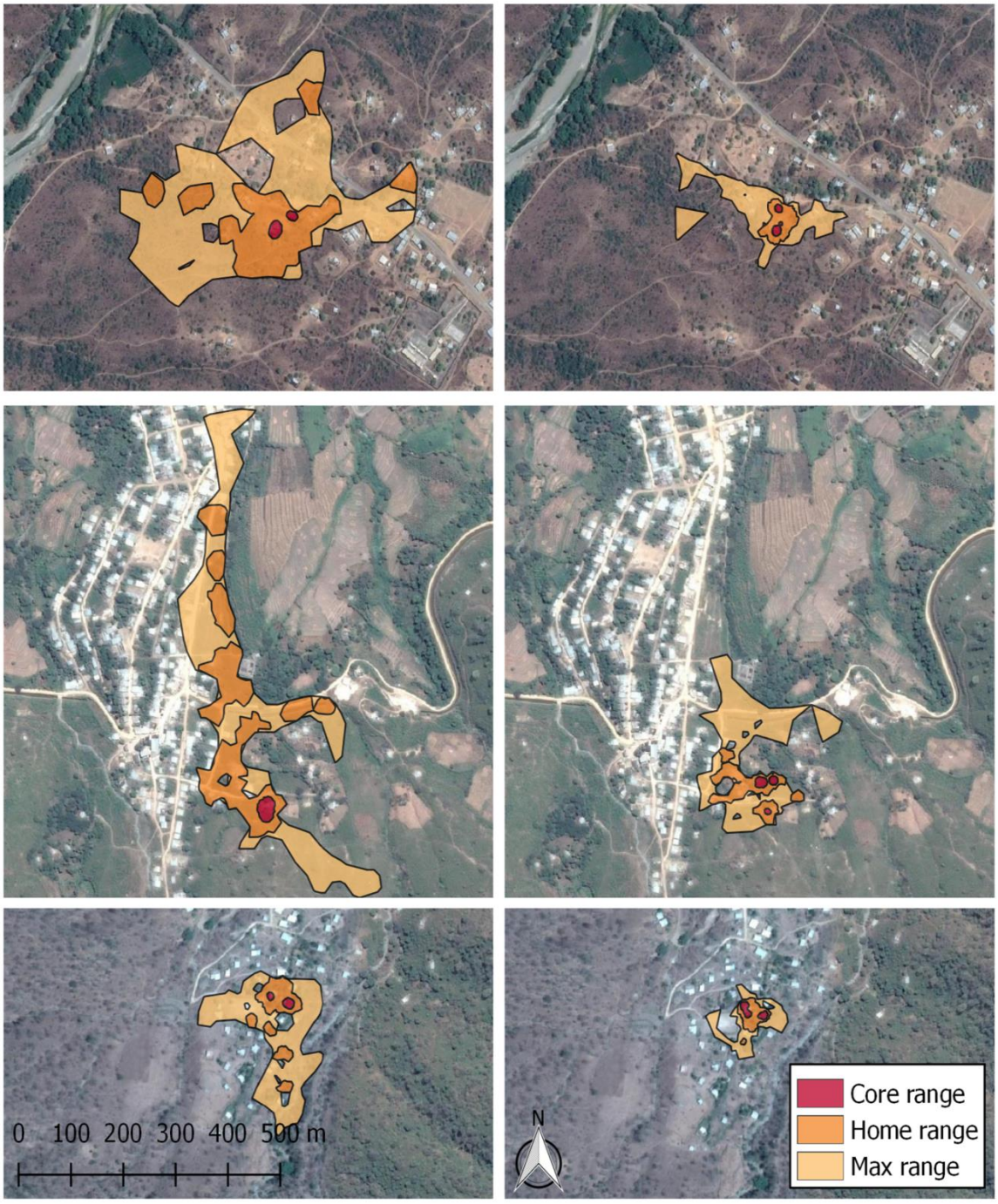


Fig 3.5. LoCoH home range maps of 6 representative pigs from 3 study villages (top to bottom: Village A, Village B, and Village C). Adjacent maps are from pigs of the same household in the rainy (left) and dry (right) seasons. LoCoH range levels represent densest 50% (Core), 90% (Home), and 100% (Max) of roaming area. *Satellite images from Google Satellite Hybrid extension for QGIS. Last update, April 05, 2017.*

Although not shown in Table 3.2, the degree of reduction observed between the rainy and dry seasons was significantly different in between villages ($p<0.001$ for village*season interaction). Villages A and B had significant reductions of 76% and 71%, respectively from the rainy to dry seasons, and Village C, the village with the smallest home ranges overall, had a non-significant 30% reduction in home range area. Home range areas by season and village are shown in Figure 3.6, and full tables of all regression outputs, including regression models for core and maximum ranges can be found in Appendix A.

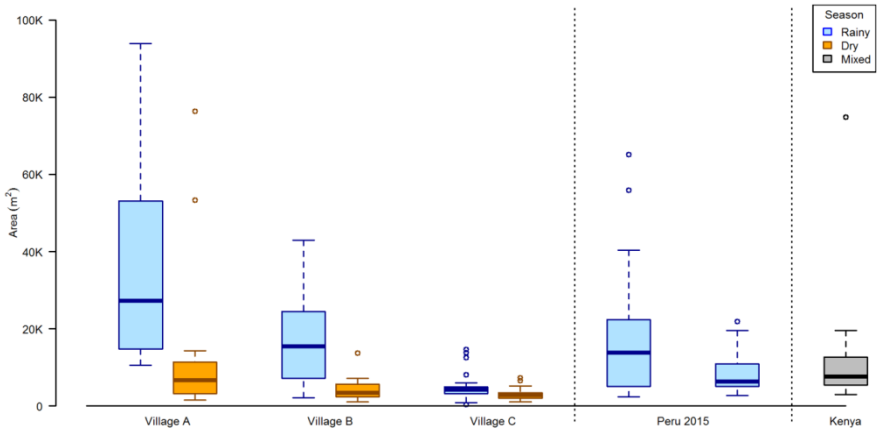


Fig 3.6. Box plot of home range areas by season and village show significant reduction in home ranges by season and between villages. Additional boxes show the home ranges extracted from pilot study in Peru [88], n=37 pigs in rainy season; and GPS tracking of 10 pigs in Kenya [89] from a mix of rainy and dry season tracking.

3.4.5 Contact with defecation sites

Overall, 56% of pigs had at least one defecation site in their home range and 85% had at least one defecation site in the maximum range. The rate of contact with defecation sites was not significantly different between the rainy and dry seasons (mean of 2.1 vs. 1.5 defecation sites in home ranges during the rainy vs. dry seasons, $p=0.18$), but did vary significantly between villages. Pigs from Village B had an average of 4.0 defecation areas in their home ranges, compared to averages of 0.6 and 0.7 in Villages A and C, respectively ($p < 0.01$, one-way ANOVA). Of the three study villages, Village B was the village with the smallest land-area, the highest density of households, and by far the most defecation sites found overall. In a negative

binomial model of contact with defecation sites (Table 3.2), residence in Village B, male sex, increased housing density up to 25 households/100-meter radius, and increased home-range area were significantly associated with the rate of contact with defecation sites. Tracking season, pig age, and herd size were not significantly associated with defecation contact (see Appendix A).

3.4.6 Pig roaming and land-type

We also analyzed the amount of active time pigs spent roaming in different land-types. Overall, pigs spent the majority of active roaming in the peri-domestic habitat, while proportionally less time was spent in vegetation and roads/paths, and very little time was spent in farmland. Season, village, household density, and home-range size were all significantly associated with roaming land-type (Table 3.3). Pigs spent significantly more time in peri-domestic areas during the dry season (64% vs. 55%, $p = 0.04$), and were also more likely to spend time in peri-domestic areas if they had smaller home ranges (74% vs. 55%, $p< 0.01$), or lived in higher-density areas of the village (66% vs. 54%, $p<0.01$). Contact with open defecation sites occur most frequently in peri-domestic and vegetated zones, less frequently along roads/paths, and was not observed in farmland (mean defecation sites in range = 2.0, 1.9, 0.9, and 0, respectively).

Table 3.3. Pig roaming land-type, by selected covariates. Mean percentage of *active* roaming time spent in given land-type. Farmland not shown due to infrequent roaming; other pig variables not shown (pig sex, age, and household herd size) were not significantly associated with any roaming land-type.

	Land-type		
	Peri-domestic	**Vegetation**	**Roads/paths**
Season[†]			
Rainy	54.8 (48.7, 60.9)*	26.7 (20.8, 32.7)	17.3 (12.8, 21.8)
Dry	64.2 (57.3, 71.1)*	20.1 (13.6, 26.6)	15.4 (11.3, 19.5)
Village[§]			
Village A	64.9 (57.2, 72.7)**	26.2 (18.2, 34.4)	8.5 (6.1, 10.9)**
Village B	46.1 (37.6, 54.5)**	26.1 (18.0, 34.3)	26.0 (20.0, 32.1)**
Village C	67.9 (61.5, 74.3)**	18.8 (11.6,25.9)	13.3 (8.9, 17.6)**
Home-range size[†]			
<3000 m^2	73.8 (67.7, 79.8)**	9.6 (0.6, 13.4)**	16.5 (10.1, 23.0)
>3000 m^2	54.6 (49.1, 60.2)**	28.2 (22.8, 33.6)**	16.2 (12.8, 19.7)
Household density[†,‡]			
≤ 10	53.8 (47.0, 60.5)**	29.1 (22.9, 35.2)**	16.7 (12.1, 21.3)
> 10	66.1 (60.2, 72.1)**	16.9 (11.0, 22.9)**	15.9 (12.0, 19.8)
# open defecation sites in range (mean, sd)	1.99 (2.3)	1.94 (2.5)	0.87 (1.2)

p-value: **<0.01, *<0.05

† Two-sample t-test used to derive p-values and 95% confidence intervals

§ One-way analysis of variance (ANOVA) used to derive p-value and 95% confidence intervals

‡Number of households within 100m radius

3.5 Discussion

The purpose of this study was to examine the roaming patterns of pigs in northern Peru, and to identify areas within their ranges where *T. solium* transmission could occur via contact with human feces. We found that pigs spent the majority of their active roaming time within 50 meters of their household. This home-centered range was concentrated in the peri-domestic habitat and predominated across both seasons and all villages (median: 79% of active time within 50m). Most of the areas of overlap between defecation sites and pig roaming ranges were found in this 50-meter zones or the wider 100-meter radius surrounding pig homes, suggesting that the majority of *T. solium* transmission risk is concentrated in these areas proximal to pigs' households.

These findings are generally consistent with our knowledge of limited pig roaming and focal *T. solium* transmission in this region. Prior spatial analyses of tapeworm carriers and

infected pigs have found that pigs living with 50 meters of a tapeworm are at significantly elevated risk of cyst infection [75,122], and our pilot GPS analysis of pig roaming in this region found that pigs spent 70% of their roaming time and 93% of their interactions with defecation sites within 50 meters of their homes [88]. Taken together, these studies provide consistent and convincing evidence that the *T. solium* transmission in this region occurs in close proximity to the home – areas where pig roaming and human defecation are concentrated – and that interventions targeting treatment resources to these hotspots of transmission are likely to be successful.

Although most pigs had limited roaming ranges and close contact with human feces near their home, many pigs spent at least some fraction of time foraging in more distant areas, and a subset of pigs had ample roaming ranges that revealed regular extended trips to distant areas. In these extreme cases, pigs ventured 1-3 kilometers from their homes, and spent nights away without returning home. These long-distance roamers are an important sub-group to consider in the context of control interventions, as they had higher rates of contact with open defecation areas, and, due to extended time away from home, may not be included in treatment, vaccination, or serological monitoring programs.

Another key finding in this study was the importance of season as a determinant of the area and distance pigs covered during roaming. Nearly all occurrences of extended roaming were observed during the rainy season, and rainy season home ranges were 61% larger than their dry season counterparts. Compared to the dry season, pigs in the rainy season also spent less time foraging in peri-domestic zones. This seasonal pattern is likely due to the increased availability of wilds fruits, vegetation, and natural streams during the rainy season. Pig owners frequently reported to us that their pigs roamed longer and further during the rainy summer months in search of wild fruits to eat and streams to bath in, and spent the dry winter months resting and grazing on domestic food sources. This seasonal pattern is consistent with a non-spatial study of pig behavior conducted in Mexico, which found that pigs spent more time feeding and walking during the rainy season, and more time resting and consuming feces during the dry season [33]. Despite our

finding of seasonality in roaming range areas, we did not detect any significant difference in contact with human feces between seasons, and therefore were not able to corroborate evidence of a seasonal pattern in *T. solium* transmission.

Apart from season, the most important determinant of the size of a pig's roaming area and its contact with defecation areas was its village of residence. Roaming areas in Village A were considerably larger than those observed in Villages B or C (median home ranges: 12570 m^2, 5697 m^2, and 3270 m^2, respectively), yet contact with defecation sites was more frequent in Village B (mean of 4.0 defecation sites in range vs. 0.6 and 0.7 in other villages). These differences highlight the importance of village-specific characteristics that may lead to heterogeneous transmission patterns between villages. For example, Village A is relatively flat with large and dispersed homesteads (6.9 households/100m) and a low rate of open defecation (97% of households had latrines), while Village B is a densely populated peri-urban settlement (26.1 households/100m) with a high rate of open defecation (only 66% of households owned latrines). Given that pig roaming patterns and contact with open defecation areas varied considerably between these villages, it is likely that spatial patterns of transmission and the degree of clustering in *T. solium* transmission differ as well. Control programs should consider the impact of these between-village heterogeneities when planning interventions. For example, the decision to select a mass or targeted intervention – or the selection of an appropriately sized treatment ring in a targeted approached – may differ depending on the degree of clustered transmission likely to be present. Knowledge of the local patterns in pig roaming, open defecation, and housing density may help to tailor intervention strategies local conditions.

This study had a few important strengths compared to prior research in this field. First, repeated tracking periods allowed us to investigate seasonal differences in roaming patterns. This aspect of pig roaming was not addressed in our prior analysis, and was not robustly evaluated in two other studies relating pig roaming to *T. solium* transmission risk – a GPS study in Kenya that tracked 5 pigs per season [89], and a non-spatial study of pig behavior in Mexico [33]. Our

current study tracked over 50 pigs per season across 3 villages, the most robust effort to date to study pig behavior as it relates to *T. solium* transmission. Second, our application of a 6-day tracking period (compared to 2 days in our prior study) and our selection of *active* roaming time were key improvements that reduced the impact of chance daily variations in roaming and the introduction of bias from periods of rest that would not contribute to transmission risk.

Despite these strengths, our study had a few important limitations. Due to the logistical challenges of mapping defecation sites in the communities, defecation mapping was only applied in the rainy season, and defecation sites were assumed to remain constant in the dry season. Although we are not aware of evidence from literature or local experts that open defecation practices vary by season, this remains a possibility, and could have affected estimates of contact with defecation in the dry season. Second, while we applied multiple measures to eliminate erroneous GPS points caused by signal disruption, some degree of imprecision in GPS points was inevitable. GPS imprecision likely introduced random error into our classification of pigs' land-usage, and reduced the accuracy of our algorithm to select periods of active roaming. Finally, roaming patterns and patterns of contact with human feces likely differ between endemic regions, and results obtained from these three villages may not be generalizable to other areas. That said, our findings are comparable to prior work on this topic from other regions [33,89] (see Fig 3.6), and spatial analyses from other regions that have detected clustered patterns of *T. solium* prevalence [74,75,79,90,116,122].

3.6 Conclusion

We found that the majority of pig roaming and contact with human defecation sites occurred in close proximity to pig homes – roaming was concentrated within 50 meters and contact with human defecation within 100 meters of pig households. When considered alongside prior GPS tracking studies and spatial analyses in this region, this study provides strong evidence

that *T. solium* transmission is most likely to occur in close proximity to the home, and supports control strategies that target treatment resources to these high-risk areas. When longer-range pig roaming occurred, it occurred more frequently in the rainy season, and depended on village characteristics such as density and topography. Therefore, while we did not find evidence that contact with feces or resulting transmission patterns vary by these factors, they could impact access to pigs, and we recommend that future control strategies consider seasonal patterns and local village heterogeneities when planning interventions such as treatment or vaccination.

The information provided here may also be useful for *T. solium* transmission models, which require precise estimates for behavioral factors that influence transmission patterns, such as pig roaming and open human defecation. Pig roaming and open human defecation are key features that cause clustered patterns of *T. solium* transmission, and modelers should account for this clustering, along with possible seasonal and village-specific differences in transmission patterns, when considering the structure and parameterization of future models. Ultimately, data from this study may fill an important gap in behavioral data needed for the development of accurate and validated *T. solium* transmission models. Advancements of *T. solium* modeling, including improved biological and behavioral data, is a need that has been highlighted by the World Health Organization as a priority for achieving control and elimination milestones [127].

Chapter 4: Exploring transmission of *Taenia solium* through sensitivity analysis of CystiAgent: a novel agent-based model

Ian W. Pray[1]*, Wayne Wakeland[2], William Pan[3], William E. Lambert[1], Hector H. Garcia[4], Armando E. Gonzalez[5] and Seth E. O'Neal[1], for the Cysticercosis Working Group in Peru.

[1]School of Public Health, Oregon Health & Science University and Portland State University, Portland, Oregon, United States

[2]Systems Science Program, Portland State University, Portland, Oregon, United States

[3]Global Health Institute, Duke University, Durham, North Carolina, United States

[4]School of Sciences, Department of Microbiology, Universidad Peruana Cayetano Heredia, Lima, Peru

[5]School of Veterinary Medicine, Universidad Nacional Mayor de San Marcos, Lima, Peru

*Corresponding author

4.1 Abstract

4.1.1 Background

The pork tapeworm (*Taenia solium*) is a serious public health problem in rural areas of developing countries, where the associated conditions of neurocysticercosis (NCC) and porcine cysticercosis cause substantial health and economic harms. The World Health Organization (WHO) has identified the development of an accurate and validated *T. solium* transmission model as a key priority for advancing towards control and elimination. Existing *T. solium* models, however, do not incorporate key spatial and behavioral features of *T. solium* transmission, and have not been tested against data from prospective studies.

4.1.2 Methodology and principal findings

We developed a spatially-explicit agent-based model (ABM) for *T. solium* ("CystiAgent") that was designed to capture the spatial and behavioral complexities of *T. solium* transmission in endemic regions. In this article, we introduce the structure and function of the model, and describe the results of sensitivity analyses conducted. Sensitivity analyses included Sobol's variance decomposition and partial-rank correlation coefficient (LHS-PRCC). The model parameters with the greatest impact on outputs were "tuning" parameters defining the probabilities of infection in humans and pigs given exposure to *T. solium*, and pig-raising practices in the simulation villages, the duration of human taeniasis, the roaming ranges of pigs, and use of latrines were also identified as key contributors to uncertainty in the model.

4.1.3 Conclusions and significance

CystiAgent represents an important new tool for investigating *T. solium* transmission, as it is the only available model to represent key spatial and behavioral aspects of transmission. Sensitivity analysis showed that application of the model to populations outside of Peru would benefit from local calibration of tuning parameters and additional knowledge of some key

behavioral parameters in order to be accurately applied to these areas. Next steps for this model include testing the model against results from prospective studies, and comparing and prioritizing available control strategies to meet WHO policy goals.

4.2 Author summary

The pork tapeworm (*Taenia solium*) is transmitted between humans and pigs in rural areas of developing countries, and represents a serious public health problem. If we are to achieve the ambitious goal of control and/or elimination of *T. solium* in affected regions, it will require innovation and collaboration from multiple disciplines, including the laboratory science, field epidemiology, and computer modeling. Here we describe our effort to develop and evaluate a novel model capable of simulating *T. solium* transmission in endemic settings. Our model, called "CystiAgent" is an agent-based model (ABM) that differs from prior *T. solium* models in that it represents the unique spatial and behavioral features of *T. solium* transmission. In our evaluation of the model, we found that a select set of parameters of the model were responsible for the majority of variation in outputs. Based on these results, we recommend that application of the model to specific endemic villages be paired with local data and statistical calibration of these key parameters. The results of this analysis will allow us to improve future versions of the model, and ultimately, produce a validated model that can be used to evaluate the effectiveness of proposed control or elimination strategies.

4.3 Introduction

Human cysticercosis, caused by the pork tapeworm (*Taenia solium*), remains a major public health concern in poor rural areas of the world. In endemic regions, up to one third of seizure disorders are attributed to neurocysticercosis (NCC), a severe neurological infection caused by the parasite [3,4]. Further, livestock losses exact a substantial economic burden [8]. *T.*

solium is transmitted between humans and pigs, and is common in agricultural communities where small-holder pig farming is common and access to sanitation is limited. Humans acquire the adult-stage intestinal tapeworm (a condition called taeniasis) by consuming raw or undercooked pork that is infected with intermediate-stage larval cysts, while pigs acquire this cyst infection (a condition called porcine cysticercosis) through contact with eggs present in the feces of infected humans.

Control and elimination of *T. solium* transmission in endemic areas is now known to be achievable [94,100] through strategic application of available drugs to treat human taeniasis [48,91] and porcine cysticercosis [34], and a vaccine to prevent infection in pigs [99,100]. Despite these effective tools, relatively few prospective studies have been conducted that compare available strategies, owing largely to the high cost and time required to carry out this work. To address this challenge, efforts to control and/or eliminate other neglected tropical disease (NTDs) have relied upon infectious disease models to compare and contrast available strategies prior to rolling out large-scale control programs [128,129]. It is at this critical juncture that *T. solium* control and elimination is now situated, as the World Health Organization (WHO) recently called upon modelers to identify a set of approaches that can be scaled and implemented in several countries by 2020 [127].

A variety of *T. solium* models have been developed in recent years that attempt to fill the above needs [12–16]. While each of these attempts have moved the *T. solium* modeling agenda forward, limitations in structure and data quality have prevented these models from providing the detailed insights needed to inform future control strategies. These existing models, like many traditional infectious disease models, rely on assumptions of spatial homogeneity, closed-populations, and parameters that are averaged across large populations. Transmission of *T. solium*, however, is uniquely difficult to model under traditional assumptions due to the complex social, biological, and environmental factors that perpetuate transmission in endemic areas. Local variations in pig-raising practices, sanitation, diet, and migration all interact to create locally

specific transmission patterns that differ from one endemic village to the next [30]. Even within villages, spatial heterogeneities caused by pig-roaming patterns and open defecation cause clustering that is important for a model to capture [75,88,122]. Importantly, models that fail to account for these heterogeneities are susceptible to overestimating the effect of control interventions [117] and yield unrealistic predictions for achieving control and elimination targets [121].

To avoid the pitfalls described above, complex ecological systems like *T. solium* transmission are well-suited for agent-based modeling (ABM). ABMs are increasingly used for modeling complex systems because they have the flexibility to simulate dynamic non-linear processes and can be applied in a spatially explicit environment [130,131]. In ABMs, the simulated population is made up of individuals ("agents") that each have a unique set of characteristics and behave according to the rules defined in the model's structure. This "bottom-up" structure allows for the modeler to easily manipulate the behaviors or the modeled environment and observe the emergent patterns that are produced by such manipulations. In the context of *T. solium* transmission, this structure facilitates application of the model to a variety of transmissions settings, and allows for testing a wide range of available control strategies, including spatially targeted strategies (e.g., "Ring Strategy" [50]), and other behavioral and structural interventions.

Of course, ABMs are not without limitations of their own. In order to account for the complex interactions between agents and the environment, ABMs have to be supplied with adequate data to define the structures and probabilities inherent to the modeled system. Until recently, this level of detailed information was not available for *T. solium.* Significant efforts to advance *T. solium* research in Peru over the past decade, however, have made such precision attainable [75,88,94]. A further limitation of ABMs is that their complexity can cause problems when attempting to understand and validate the dynamics of the system – for example, which behaviors are most important in driving transmission, and what are the key sources of uncertainty

in the model? To address these concerns, it is recommended that complex ABMs be subjected to rigorous sensitivity analyses prior to application for the purposes of prediction or evaluation of policy goals [132,133].

Our objective was to develop an ABM for *T. solium* transmission, and to subject the model to rigorous sensitivity analysis in order to evaluate its structures and parameters. In this article, we present the newly available model, called CystiAgent, with a detailed description of its structure and data sources. We also present the results of two sensitivity analyses applied to the model. The sensitivity analyses were conducted with three key objective in mind: 1) to investigate which parameters contribute more prominently to disease transmission, and determine the shape and direction of the relationship they have with outputs; 2) to identify key parameters that contribute to model uncertainty; and 3) to evaluate the robustness of the model to variability in parameter settings, a step that will help to assess the model's generalizability to other endemic settings. Finally, we should be clear that this article (Chapter 4) is intended as only an introduction and initial assessment of a new and hopefully useful tool. This article does not include validation of the model to a specific endemic setting, and does not include a prospective evaluation of available control or elimination strategies (see Chapter 5).

4.4 Methods

4.4.1 Model description

Model overview. CystiAgent is a spatially explicit ABM that is able to simulate endemic transmission of *T. solium* and test a variety of population-level interventions designed to control or eliminate *T. solium*. CystiAgent was developed in NetLogo 6.0.4 (Northwestern University, Evanston, IL), an open-access ABM software that was chosen for its ability to represent spatial data and display simulations through a graphical interface.

In CystiAgent, there are two agent classes – humans and pigs – that represent the primary and intermediate hosts of *T. solium*, respectively. Human and pig agents are individually assigned a set of features at baseline and behave in time and space according to the structural assumptions and parameters defined in the model. All humans and pigs are assigned to discrete household units that are distributed across the simulation village, and whose locations are given by set of input coordinates that can represent real or fictitious villages. Currently, CystiAgent is designed to simulate transmission in one village at a time (pop. up to ~2,000), but can be applied to any population with corresponding input coordinates. Each time-step of the model represents one week of cumulative activities and exposures.

Model outcomes. Humans may be infected with the adult-stage intestinal tapeworm (i.e., *T. solium* taeniasis), and pigs may be infected with larval-stage metacestodes (i.e., porcine cysticercosis). Pig infection is categorized as heavy (≥ 100 cysts) or light (< 100 cysts) cyst burden, while pig exposure includes the possibility of antibody response to allow comparison with serological assays used in field studies. Human cysticercosis, including NCC or NCC-related seizure disorders, is not included in this model as it does not contribute to transmission.

Model flow. Model processes can be roughly categorized into six steps that loop continuously in order to simulate natural endemic transmission (see Fig 4.1):

(1) Village setup. Households are assigned to a set of geographic coordinates and are populated with pigs and people. Input characteristics specific to each village are assigned, if known. These include the pig and human population size, and the proportion of households that own latrines, raise pigs, and own pig corrals. A proportion of humans and pigs are then assigned initial infection at baseline. Households are assigned the presence/absence of latrines and corrals, as well as the herd size for each pig-raising household

(2) Pig sale, import/export, and slaughter. Pigs due for slaughter may be butchered at home, sold within the village, or exported. Potentially infected pigs from external villages may

also be imported. When pigs are slaughtered in the village, their meat is either consumed exclusively at home, sold to other households, or shared between the two.

(3) Tapeworm infection, maturation and death. When consumed pork is infected with *T. solium* cysts, all members of the consuming households are exposed to potential tapeworm infection. If humans acquire a tapeworm infection, the intestinal tapeworm reaches maturity after 8 weeks [22,27], and begins expelling infectious eggs at that time. Tapeworm infections naturally clear after pre-determined infectious durations [22,27].

(4) Human travel. Humans that are designated as travelers leave the community at regular intervals, may contract tapeworm infections while traveling in other endemic areas, and return to the village after travel. Upon return, infected travelers resume contamination of their environment if applicable.

(5) Environment contamination and egg decay. Human tapeworm carriers that do not own or use a latrine release *T. solium* eggs and proglottid segments into the environment surrounding their household location. When tapeworm infections clear, humans stop releasing proglottid segments, but contamination of the environment with eggs persists until the eggs naturally degrade [134].

(6) Pig roaming and cyst infection. Pigs that are designated as free-roaming (i.e., not contained in corrals) are exposed to *T. solium* proglottids and eggs that are present in their home-range areas. Exposure to proglottid segments may lead to heavy cyst infection, while exposure to eggs may lead to light cyst infection. Either may result in antibody response and seropositivity. Free-roaming pigs are exposed to an additional risk of infection or seropositivity that is proportional to the number of tapeworm carriers in the village and naïve to the pig's location. This represents exposure to pigs that results from roaming and consumption of human feces from open defecation that occur outside of the home area.

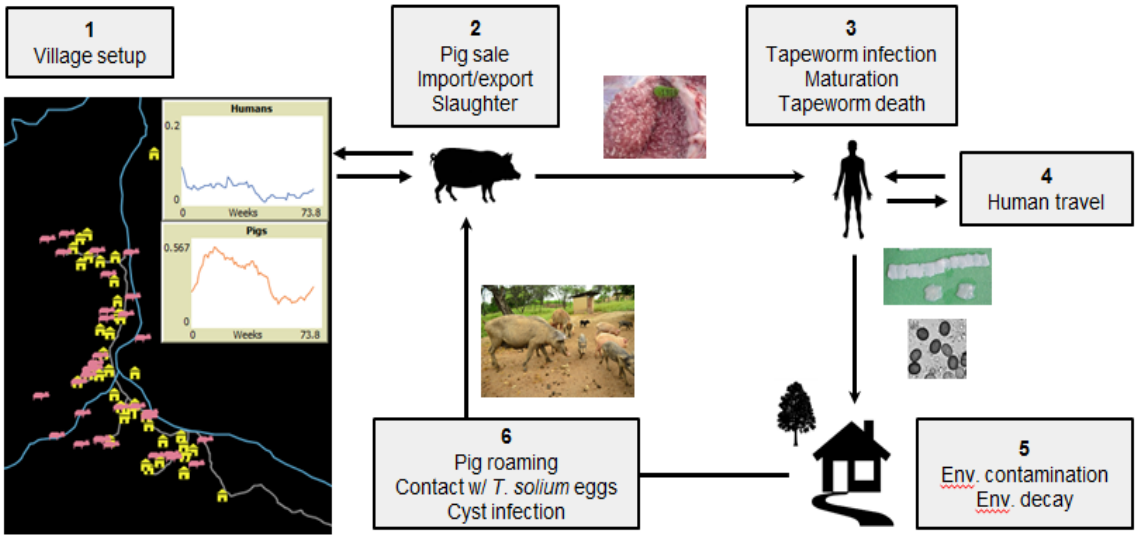


Fig 4.1. Diagram of CystiAgent model flow including visualization of the model in NetLogo depicting transmission in a simulated village (left). Line graphs represent the simulated prevalence of human taeniasis (upper) and porcine cysticercosis (lower).

Parameters. Each model process above is defined mathematically by a probability distribution and corresponding parameter(s) (Appendix B2). Depending on the model activity they represent, most parameters correspond to the central value (e.g., mean) and spread (e.g., variance) of a chosen distribution. During setup and running of the model, continuous features are assigned to participants based on random number generation from the designated probability distribution, while categorical features and randomly assigned from a binomial distribution.

A variety of sources, including primary data, literature review, and expert opinion, were utilized to determine the values and distributions for model parameters. For the majority of parameters, we used data collected in the Piura region of northern Peru. A full description of the methods and data sources used to estimate each parameter value can be found in Appendix B1. For the purposes of sensitivity analyses, we designated a "plausible range" of values for each parameter in addition to its estimated central value. This is a range of values across which the model was evaluated to determine their impact on model outputs. In some cases, the plausible range was represented by the range of mean values observed across a group of endemic villages,

and in other cases we manually widened the range to account for additional uncertainty and variability in the parameter.

Tuning parameters. In addition to the above suite of biological, behavioral, and environmental parameters, CystiAgent utilizes a set of tuning parameters to adjust the model to different local conditions and endemic prevalence levels. When the model is applied to specific observed prevalence levels for validation, this set of tuning parameters must be calibrated independently for each unique village using an approximated Bayesian computation (ABC) algorithm [135]. For the purposes of this sensitivity analysis, we intentionally set wide plausible ranges for tuning parameters in order to represent a broad range of possible transmission levels and measure their impact on the model.

There are six tuning parameters that represent different probabilities of exposure or infection in the model. Two tuning parameters define the probabilities of tapeworm infection after slaughter of heavily ("ph2h") and lightly ("pl2h") infected pigs; two other tuning parameters define the probability of heavy and light pig infection after exposure to proglottid segments ("heavy-inf"), and eggs ("light-inf") present in the environment; and the remaining two parameters determine the probability of exposure to proglottid segments ("heavy-all") or eggs ("light-all") during pig-roaming outside of a pig's home-range area.

Interventions. CystiAgent has the ability to simulate a variety of population-level interventions designed to control or eliminate *T. solium* transmission. A generic function is available to administer anti-helminthic treatment for human taeniasis, either presumptively or after stool screening. Other functions include the treatment of pigs to cure cystic larval infection, or vaccination to prevent infection. For each intervention type, user-controlled options allow for specification of participation levels, the sensitivity of screening tests, and the efficacy of drugs and vaccines used. These interventions can then be implemented through mass or targeted approaches, while varying the duration and frequency of intervention applications. Unique to this spatial model is the ability to simulate spatially targeted interventions. "Ring strategy" [50] can

be applied by targeting treatment resources to households residing within a given distance of heavily infected pigs. Finally, behavioral and developmental interventions such as improved access to corrals and latrines are available as stand-alone interventions or in combination with other approaches.

4.4.2 Sensitivity analysis of CystiAgent

We performed all sensitivity analyses in R version 3.5.1, using the "RNetLogo" package [136] to execute model simulations in NetLogo from R. Sensitivity analyses included the application of Sobol' variance decomposition and Latin hypercube sampling partial rank correlation coefficients (LHS-PRCC) in three unique villages, each representing a different context of population size and housing density. Household coordinates for the three test villages were based on real endemic villages in northern Peru that recently participated in a large prospective trial ("Ring Strategy Trial", in peer review) [137]. For evaluation of the CystiAgent model, sensitivity analyses were applied to two model versions: the full model that contained all model parameters (k = 33 parameters), and a reduced model for which village input characteristics and tuning parameters were fixed (k = 22 parameters), allowing for a more in-depth evaluation of key biological and behavioral parameters. For the reduced model, fixed values for village input characteristics (i.e., humans and pigs per household, pig ownership, corral and latrine access) were based on data from the census applied in each village, while tuning parameters were estimated using an ABC algorithm to fit the model to observed levels of transmission in each village (i.e., baseline prevalence of human taeniasis and porcine cysticercosis in the parent study). Each run of the model in sensitivity analyses consisted of 1000 weeks of stable endemic transmission with no interventions applied. The summary statistics collected at the end of each run were defined as the incidence-density of human taeniasis (number of new infections / 100 person-years), and the lifetime cumulative incidence of porcine cysticercosis (cumulative number of infected pigs / cumulative pig population).

In order to achieve the computational resources needed to run the model through many thousands of simulations for each of these analyses, we executed all model simulations on the Amazon Web Service EC2 cloud computing platform. Model simulations were distributed across a 72-core parallel processor using the "parallel" R-package [138] and executed on the EC2 cloud using the R-Studio Shiny server [139].

Sobol' variance decomposition. Readers seeking a detailed description of the Sobol' methodology should look to seminal works by Sobol' [140,141] and in-depth examples of their application to complex models [133,142,143]. For our application of the Sobol' method, we first determined plausible ranges for each model parameter as described above, and sampled values from each parameter distribution using a Sobol' sequence. Compared to other common sampling methods (e.g., simple random, Latin-hypercube, etc.), a Sobol' quasi-random sequence has been found to cover the parameter space more efficiently and allows for smaller sample sizes in sensitivity analyses [144,145]. As described in [146], the sample was divided into two input matrices, and then further arranged into $k + 2$ design matrices for evaluation in the model. The computational cost of this method depends on the number of input parameters (k) and the chosen number of samples drawn for each parameter (N), totaling $N*(k + 2)$. For this analysis, we selected per-parameter sample sizes (N) of 5000, 1500, and 1500 for the low, medium, and high-density villages, respectively, to account for the extra computing time required for larger populations. $N \geq 500$ is recommended for complex models [147]. Evaluating $k = 33$ model parameters in the full model and $k = 22$ parameters in the reduced model led to final computational costs between 36,000 and 175,000 per analysis, depending on the village.

The results of model simulations were analyzed using the "sobol2007" function available in the "sensitivity" package in R. The Sobol' method quantifies sensitivity of the model to each parameter with two measures: first-order sensitivity index (S_i), and a total effects sensitivity index (ST_i). S_i estimates the independent contribution of each parameter to variance in the model outcomes, while ST_i estimates the full contribution of each parameter after considering

interactions with other parameters [144,148]. Equations for S_i and ST_i are given below, with V_t representing the overall variance in the output, V_i representing the variance due to the uncertainty in parameter *i*, and $S_{(-i)}$ representing the sum of all S_i indices other than index *i*. First-order indices were considered significant if $S_i > 0.02$ in the full model analysis, or $S_i > 0.01$ in the reduced model analysis; 95% confidence intervals for S_i and ST_i were generated with 100 bootstrapped replications [149]. First-order and total-effect indices were calculated for human taeniasis and porcine cysticercosis in each of the three villages analyzed.

$$S_i = V_i / V_t$$

$$ST_i = 1 - S_{(-i)}$$

Latin hypercube sampling-partial rank correlation coefficient (LHS-PRCC). A detailed description of LHS-PRCC method can be found elsewhere [150]. Briefly, LHS-PRCC provides a non-parametric measure of the strength of monotonic association between each parameter and the model output. It begins with a Latin hypercube sample of each parameter for which the parameter ranges are divided into *n* equal segments, and a random value is drawn from each segment, as described [151]. For LHS-PRCC analyses on both the full (k = 33 parameters) and reduced (k = 22 parameters) models, we chose equivalent sample sizes (n) of 175,000, 50,000, and 50,000 for low, medium, and high-density villages, respectively. We then ran the model through all parameter permutations and analyzed the results to determine partial-rank correlation coefficients for each parameter using the "sensitivity" and "ppcor" R packages. For this, the PRCC formula calculates the linear correlation, ρ, between the residuals of the rank-transformed parameter input and rank-transformed model output, while accounting for correlations with all other parameter inputs [150]. Importantly, the final PRCC estimates provide measures of the strength, direction, and statistical significance of the association between parameter inputs and model outputs. P-values were obtained with a Student's *t* distribution and

were evaluated with a Bonferroni adjustment for 33 multiple comparisons ($p < 0.0015$ for statistical significance).

4.5 Results

4.5.1 Sobol' sensitivity analysis

Figure 4.2 contains graphs of CystiAgent model parameters that had significant first-order (S_i) and total-effect (ST_i) indices in Sobol' sensitivity analyses. Appendix B3 contain graphs and indices from all villages and analyses.

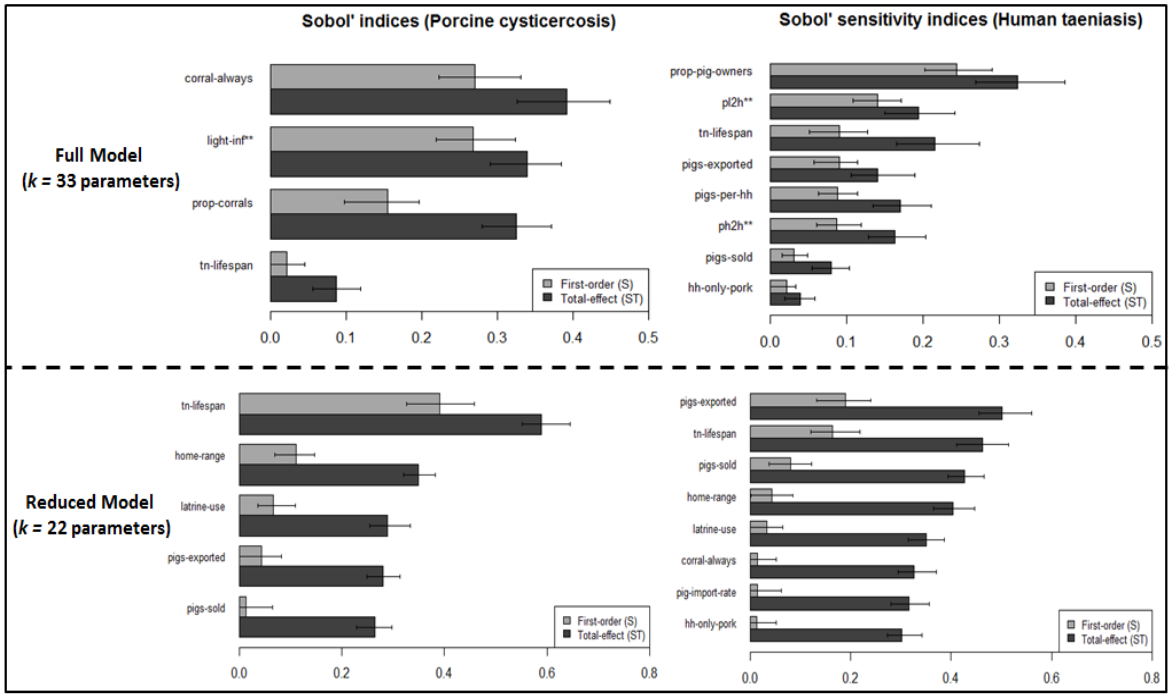


Fig 4.2. Sobol' first- and total-order indices for porcine cysticercosis (left) and human taeniasis (right), in the full model (top) and reduced model (bottom), medium-density village. Parameters with first-order indices $S_i > 0.02$ in the full model analysis and $S_i > 0.01$ in the reduced model analysis are shown. 95% confidence intervals produced with 100 bootstrap replications. See Appendix B2 for descriptions of parameter names and functions.

Full model analysis. Of the 33 parameters included in the analysis of the full CystiAgent model, parameters that were consistently identified as impactful on rates of porcine cysticercosis were the parameters defining the use of corrals to contain pigs, and pig-related tuning parameters.

Specifically, "always" using corrals for all owned pigs ("corral-always") had the most consistently high impact on output variance, with first-order indices of S = 0.10, 0.35, and 0.27 in low, medium, and high-density villages, meaning that 10%, 35%, and 27% of the variance in pig infection was attributed to the uncertainty range of this parameter. Similarly, the probability of light-infection after exposure to *T. solium* eggs ("light-inf") was highly impactful in each of the village analyses, with first-order indices of 0.25, 0.44, and 0.27 in the three villages.

The parameters that contributed most to variance in rates of human taeniasis in the full model analysis were those that determined the number of pigs in the population (and therefore more opportunities for infection) and the set of human-related tuning parameters. Specifically, the proportion of households raising pigs ("prop-pig-owners"), the mean number of pigs per household ("pigs-per-hh"), the proportion of pigs sold prior to slaughter ("pigs-sold"), and the proportion of sold pigs that were exported out of the village ("pigs-export") all had significant first-order indices in at least two of the three villages tested. For human-related tuning parameters, the probabilities of tapeworm infection after slaughter of a lightly ("pl2h") or heavily ("ph2h") infected pig were both highly impactful. The mean duration of tapeworm infections was also an important contributor to output variance in in two of the three villages ("tn-lifespan").

Reduced model analysis. When tuning parameters and village input characteristics were fixed for the reduced model analysis, the relationships between the remaining model parameters and model outputs changed considerably. Of the 22 parameters included in the reduced model analysis, the most consistently impactful parameter for both porcine cysticercosis and human taeniasis was the average duration of taeniasis infection ("tn-lifespan"), which accounted for 31%, 39%, and 29% of the total variation in pig infections, and 18%, 16%, and 17% of the total variation in human taeniasis rates across the three villages tested. After tapeworm lifespan, the second and third most impactful parameters in the reduced model analysis were the size of pig home-ranges ("home-range") and use of latrines ("latrine-use"), neither of which were identified as impactful in the full model analysis. These parameters accounted for an average of 11% and

8% of the variance in pig infection, and 5% and 4% of variance in human taeniasis, respectively, across the three villages evaluated. Finally, the proportion of pigs exported ("pigs-exported") and sold ("pigs-sold") were consistently identified as impactful parameters in the reduced model analysis.

Total effect indices. Total-effect indices (ST_i) in the full model analysis followed similar patterns as first-order effects (S_i), but were consistently larger to account for the extra variance due to interactions between parameters. Of the three test villages, the low-density village had the greatest disparity between first-order and total-effect indices. Similarly, total effect indices were greater in the reduced model analysis compared to the full model analysis, indicating that interaction effects between parameters contributed to a greater proportion of output variance in the reduced model.

4.5.2 LHS-PRCC

Figure 4.3 displays the parameters that were significant in the full and reduced versions of the LHS-PRCC sensitivity analysis for the medium-density village. A complete graphical comparison of LHS-PRCC results between both model versions and all three villages can be found in Table 4.1. Overall, the results of LHS-PRCC were similar to those produced by the Sobol' method. In both the full and reduced model analyses, LHS-PRCC identified the same set of highly influential parameters that had positive first-order Sobol' indices, but LHS-PRCC identified a larger set of lower impact parameters and was able to determine the direction of the associations between parameters and the model outputs.

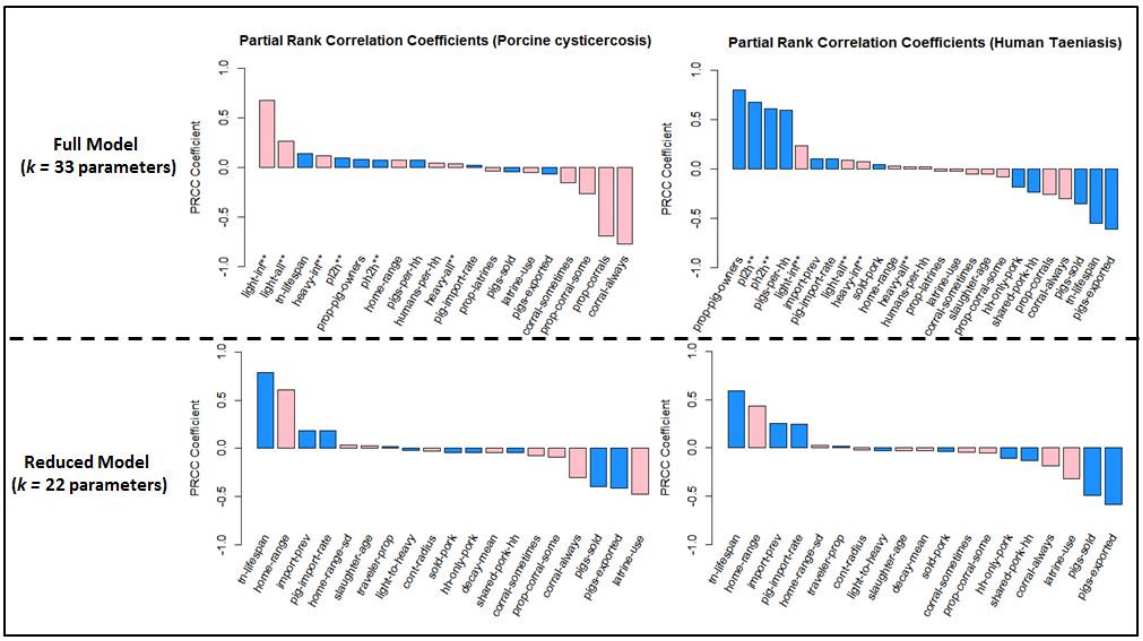


Fig 4.3. Partial rank correlation coefficients for porcine cysticercosis (left) and human taeniasis (right) in the full model (top) and reduced model (bottom), medium-density village. Bar colors represent the primary impact of each parameter (blue = human taeniasis, pink = porcine cysticercosis). Parameters with p-values < 0.0015 are shown. See Table 1 for a description of parameter names and functions.

Full model analysis. Similar to the results of the Sobol' method, the suite of parameters that define the use of corrals to contain pigs and the pig-related tuning parameters were the most consistently and strongly correlated with porcine cysticercosis. Most prominently, this included the proportion of pig-owners that "always" corralled their pigs ("corral-always"), which was highly protective for pigs with coefficients of ρ = -0.56, -0.78, and -0.82, in the low, medium, and high-density villages, respectively. Owning a corral ("prop-corrals") and use of the corrals on "some" pigs ("corral-sometimes" and "prop-corral-sometimes") were also correlated with decreased pig infection in the model. Among the pig-related tuning parameters assessed, the most impactful parameters were the probability of light cyst infection after exposure to environmental egg contamination ("light-inf") and the probability of exposure to environmental egg contamination outside of home-range ("light-all").

For human taeniasis as the model output, the four parameters most strongly correlated with increased incidence were the two human-related tuning parameters ("pl2h" and "ph2h") and

the two parameters responsible for determining the size of the local pig population: the proportion of households that raise pigs ("prop-pig-owners") and the mean number of pigs per household ("pigs-per-hh"). Parameters that were strongly associated with a decreased incidence of taeniasis in all three villages included the export of pigs out of the village ("pigs-exported"), the sale of pigs prior to slaughter ("pigs-sold"), and an increased duration of tapeworm infection ("tn-lifespan"). In addition to these strong correlations, the rate of pig import ("pig-import-rate") and the prevalence of cyst infection among imported pigs ("import-prev") were consistently correlated with small increases in taeniasis, while parameters that promoted consumption of pork at home ("hh-only-pork", "shared-pork-hh") were associated with small decreases in taeniasis.

Reduced model analysis. Similar to the results of the Sobol' analyses, when tuning parameters and village characteristics were fixed, the set of parameters that impacted transmission shifted considerably. The mean duration of taeniasis was the parameter most strongly correlated with increased rates of both porcine cysticercosis (ρ = 0.63, 0.79, and 0.71) and human taeniasis (ρ = 0.49, 0.59, and 0.57) in the low, medium, and high-density villages. The size of pig home-ranges ("home-range"), the rate of pig import ("pig-import-rate") and the prevalence of cyst infection among imported pigs ("import-prev") were all significantly correlated with increased incidences of porcine cysticercosis and human taeniasis in all three villages; while the use of latrines ("latrine-use"), proportion of pigs exported ("pigs-exported"), proportion of pigs sold ("pigs-sold"), and use of corrals to contain pigs ("corral-always") were all significantly correlated with reduced rates of both porcine cysticercosis and human taeniasis in all three villages.

Table 4.1. Illustrated comparison of LHS-PRCC results of full and reduced models across low, medium, and high-density villages. Parameters with significant LHS-PRCC coefficients ($p < 0.0015$) shown. See Appendix B1 for descriptions of parameter names and functions.

	Full Model (k = 33 parameters)			Reduced Model (k = 22 parameters)		
Parameters	High	Medium	Low	High	Medium	Low
Pig						
tuning-pig[§]	●●● ■■	●●● ■■	●●● ■■■			
corral-always	●●● ■■	●●● ■■	●●● ■■	●● ■■	●● ■	● +
prop-corrals	●●● ■	●●● ■■	●●● ■			
corral-sometimes	+	● +	● +	● +	+ +	+
prop-corral-some	+	●● +	● +	● +	+ +	+ +
home-range	+ +	+ +	● +	●●●■■■	●●● ■■	●● ■■
latrine-use	+	+ +	● +	●● ■■	●● ■■	●● ■■
humans-per-hh	+ +	+ +	● +			
prop-latrines	+	+ +	+ +			
slaughter-age	+	+	+	+	+ +	+
decay-mean			+	+ +	+ +	+ +
cont-radius				+	+ +	
home-range-sd				+ +	+ +	+ +
Human						
tuning-human[§]	■■■ +	■■■ ●	■■■ ●●			
prop-pig-owners	■■■ +	■■■ +	■■■ ●●			
tn-lifespan	■■■ +	■■■ ●	■■ ●●	■■■●●●	■■■●●●	■■ ●●●
pigs-per-hh	■■■ +	■■■ +	■■ ●			
pigs-exported	■■■ +	■■■ +	■■ ●	■■■ ●●	■■■ ●●	■■ ●●
pigs-sold	■■ +	■■ +	■■ ●	■■ ●●	■■ ●●	■■ ●●
hh-only-pork	■	■	■ +	+ +	■ +	+ +
shared-pork-hh	■	■	■ +	■ +	■ +	+ +
pig-import-rate	■	+ +	+ +	■■ ●	■ ●	■ ●
import-prev	■	■	+ +	■■ ●	■■ ●	■ ●
sold-pork	+	+	+	+ +	+ +	
travel-duration	+					
travel-incidence				+ +		+ +
light-to-heavy				+	+ +	+ +
traveler-prop					+ +	

■ = Human taeniasis, ● = Porcine cysticercosis

SOBOL first-order indices (S_i) : ■■■ / ●●● > 0.25; ■■ / ●● > 0.1; ■ / ● > 0.02

PRCC ($|\rho|$) : ■■■ / ●●● > 0.5; ■■ / ●● > 0.25; ■ / ● > 0.1; + / + < 0.1; (all $p < 0.0015$)

[§]tuning-pig and tuning-human refer collectively to the set of tuning parameters defining the probabilities of pig and human infection.

4.6 Discussion

Our primary objective of this research was to develop a functional ABM capable of simulating the complex behavioral, biological, and environmental factors that contribute to *T. solium* transmission in endemic areas. Our sensitivity analyses demonstrated that the CystiAgent model effectively replicated key aspects of the *T. solium* life-cycle, including structural and behavioral features of transmission that are not available in other existing transmission models. Features such as access to corrals and latrines, sale and export of pork, and roaming patterns of pigs were identified as highly impactful on transmission in the final calibrated model, and incorporating these features is a unique advantage of our spatial ABM.

Our long-term goal is to provide a validated *T. solium* model that can be used to prioritize candidate control and elimination strategies. The current analysis allowed us to move closer to this goal by both demonstrating the ability of the CystiAgent model to represent the complex dynamics of *T. solium* transmission, and identifying key model parameters that must be investigated in order to apply the model to specific endemic settings in the future.

In our full model analysis, we found that the parameters that had the strongest impact on model variability were the "tuning" parameters that defined probabilities of infection in the model. For porcine cysticercosis, these included the probabilities of heavy or light infection upon contact with *T. solium* eggs or proglottids in the environment, and for humans, these included the probabilities of tapeworm infection upon consumption of heavily or lightly infected pork. Due to their considerable impact on transmission in the model, and the wide range of values they can assume, statistical calibration of the values of these parameters is highly recommended for application of the model to any specific transmission setting. Bayesian approximation [135] or other available parameter estimation methods [152] can be employed for this purpose.

Apart from these tuning parameters, many of the highly impactful parameters identified in our full model analysis fell into the category of village characteristics. These were parameters that defined the number of households raising pigs, the number of pigs per household, and access to corrals to contain pigs. The impact of these parameters on transmission levels demonstrates the importance of local variation in population structure and pig-raising practices on *T. solium* transmission dynamics. In light their impact, determining local values for these village characteristics should be a priority when applying the model to specific endemic settings. Steps such as population census or consultation with local leaders to acquire information about the size and characteristics of the pig and human populations would allow for reduced uncertainty and improved model accuracy.

We conducted our reduced model analysis in order to see beyond the tuning parameters and village characteristics that were driving uncertainty in our first set of analyses (i.e., full model analyses). This reduced analysis allowed us to assess the impacts of a smaller set of biological and behavioral parameters in the context of transmission levels that were tuned to more realistic levels. In this reduced analysis, the average duration of tapeworm infections ("tn-lifespan") emerged as the most significant source of uncertainty in all villages and analyses. The size of pig home ranges ("home-range"), the proportion of households that regularly use latrines ("latrine-use"), and the sale ("pigs-sold") and export ("pigs-exported") of pigs were also consistently identified as impactful in this reduced model analysis.

The impacts attributed to parameters in this reduced model analysis reflect both the strength of the relationship they have with model outputs, and the amount of uncertainty defined in the parameter values themselves (i.e., the width of the defined "plausible range"), which exerts considerable leverage on a parameter's measured impact. Each of the key parameters identified above were varied across wide ranges due to our uncertainty in the true value of the parameter (e.g., mean tapeworm lifespan ranged from 6 months to 2 years, the percent pigs exported ranged from 34% to 100%, etc.; see Appendix B2).

For biological parameters like tapeworm lifespan, this high degree of uncertainty is due to limited knowledge from experimental studies [27,31], and data is unlikely to improve due to ethical constraints on experimental infection. For other parameters, wide uncertainty ranges are due to the variability that exists between endemic villages and regions. Each of these factors depends on cultural, behavioral, and economic practices that are context-specific. For example, estimates for the home ranges of free-roaming pigs were based on a GPS study recently completed in three villages of northern Peru (see Chapter 3), but even within this restricted locale, variations in topography, landscape, and pig management led to substantial differences between villages. Similar between-village variations were seen in the sale and export of pigs, which served as a primary economic activity in some rural villages evaluated, and a rare source of emergency income in others. Finally, the prevalence and use of latrines varied considerably between villages depending on whether state-sponsored latrine construction had included the village. Taken together, these local variations are important to take into account when applying the model to specific endemic settings. As with key village characteristics outlined above, investigation of these local behavioral features through surveys or expert consultation prior to application of the model would reduce parameter uncertainty and likely improve validity of the model for that setting.

The parameters identified in our sensitivity analyses are generally consistent with the only other published sensitivity analysis for a *T. solium* transmission model [15]. The EPICYST model is a deterministic mathematical model that includes human cysticercosis as a primary model outcome and was parameterized based on data from *T. solium* transmission in a sub-Saharan Africa. Consistent with our findings, an LHS-PRCC analysis of EPICYST revealed the most influential parameters to be "transmission coefficients" that define the rates of infection upon exposure, the expected duration of tapeworm infections, and the rate of pork consumption among humans. However, EPICYST is a population-level model and does not include individual

behaviors or a spatial framework. Therefore it is not able to provide a comparison to other important features of our model such as pig corralling, pig roaming, and latrine use.

There are a few important strengths and limitations of our approach to highlight. First, we chose to design CystiAgent within the framework of an ABM, which allowed us to account for the complex spatial and behavioral heterogeneities that affect *T. solium* transmission in endemic areas. Despite this strength, CystiAgent only begins to account for the complex heterogeneities that likely occur in real-world systems. Age-related differences in pig roaming patterns [88], seasonal and climate-related variations in transmission [33], acquired immunity [153], vector-borne transmission [86,87], and black-market distribution of infected pork [154] are only a few of the many additional factors that may impact transmission patterns and are not explicitly defined in CystiAgent.

Second, the parameter inputs used in CystiAgent were primarily sourced from a single region of northern Peru through extensive work conducted in the region over the past decade. The depth of data available in this region is a strength of our approach and made it possible to construct this detailed ABM. Nonetheless, parameter values that are accurate for this region of Peru may be vastly different from corresponding settings in other endemic regions. Therefore, application to new regions would require some degree of input data for key parameters and local calibration of tuning parameters. That said, the results of our sensitivity analyses showed that model outputs are robust to variations in all but most sensitive parameters.

Another notable limitation of the model is that much of the behavioral data used to define parameter values in CystiAgent is self-reported by participants, and, as such, prone to self-report biases. For example, usage rates for latrines and corrals are almost certainly over-reported by participants, as has been demonstrated with self-report of other health-related behaviors [155]. To account for this additional uncertainty, we widened the plausible ranges of these parameters for sensitivity analyses. Nonetheless, the impacts of these parameters on transmission rates should

be interpreted with caution, and with the knowledge that behaviors defined in the model represent optimal compliance.

Finally, an important strength of our sensitivity analyses was our use of two complementary methods (Sobol' and LHS-PRCC) and our application of the methods on three villages of differing population sizes and densities. The consistency of our results between methods and villages provides confidence that the key features of the model are robust to variation in population structure and methodology. Despite these promising findings, the model could be tested in additional endemic settings to provide further insight into parameter relationships. Perhaps most importantly, sensitivity analyses should be conducted in the context of control interventions, as key parameters that affect transmission at endemic equilibrium (e.g., human travel and migration, pig importation) may be different when control pressure is applied.

4.7 Conclusion

In this research, we developed a functional ABM that is able to represent the core features *T. solium* transmission observed in endemic settings. Our sensitivity analyses demonstrated that the CystiAgent model functioned as expected, with key biological, behavioral, and environmental parameters interacting to uniquely impact patterns of *T. solium* transmission. Despite significant uncertainty in some key model parameters, the robustness of our model to variations in all but the most sensitive parameters suggests that the model is likely to be transportable to other endemic settings outside of Peru, given local specification of these key parameters and calibration of tuning parameters to local levels of transmission. While the generalizability of the model to other populations outside of Peru will remain unknown until it is tested in these settings, we have conducted validation of CystiAgent model against data from prospective trials conducted in Peru, and present the results of this validation in Chapter 5. Ultimately, our goal is to provide this validated model as a tool for researchers and policy-makers

seeking to compare available control strategies for *T. solium* and prioritize promising strategies for evaluation in prospective trials.

Chapter 5: Validation of an agent-based model for *T. solium* transmission using two large prospective trials in Peru

Ian W. Pray[1]*, Claudio Muro[2], Ricardo Gamboa[2], Percy Vilchez[2], Wayne Wakeland[3], William Pan[4], William E. Lambert[1], Armando E. Gonzalez[5], Hector H. Garcia[2,6], and Seth E. O'Neal[1,2], for the Cysticercosis Working Group in Peru.

[1]School of Public Health, Oregon Health & Science University and Portland State University, Portland, Oregon, USA

[2]Center for Global Health Tumbes, Universidad Peruana Cayetano Heredia, Tumbes, Peru

[3]Systems Science Program, Portland State University, Portland, Oregon, USA

[4]Global Health Institute, Duke University, Durham, North Carolina, USA

[5]School of Veterinary Medicine, Universidad Nacional Mayor de San Marcos, Lima, Peru

[6]School of Sciences, Department of Microbiology, Universidad Peruana Cayetano Heredia, Lima, Peru

*Corresponding author

5.1 Abstract

5.1.1 Background

The pork tapeworm (*Taenia solium)* is a parasitic helminth that imposes a major health and economic burden on poor rural populations around the world. As recognized by the World Health Organization, a key barrier for achieving control of *T. solium* is the lack of an accurate and validated simulation model with which to study transmission and evaluate available control strategies.

5.1.2 Methods

We developed and validated an agent-based model for *T. solium* transmission. Our model, CystiAgent, is unique among *T. solium* models in its ability to represent spatial patterns and geographically targeted interventions, as well as key behavioral and environmental features of *T. solium* transmission. We tested the model against results from two large prospective trials conducted in Peru – the Cysticercosis Elimination Demonstration Project and the Ring Strategy Trial – which, together, comprised 40 villages and 10 unique intervention types.

5.1.3 Findings

CystiAgent was able to accurately replicate baseline levels of transmission observed in all 36 villages tested (4 of the original 40 villages were excluded) and adequately predicted declines in transmission when control strategies were applied. Model-predicted intervention effects were slightly overestimated overall, and accuracy varied by study and intervention type.

5.1.4 Interpretation

CystiAgent represents an important new tool to promote control and elimination of *T. solium.* Results from this validation will be used to improve future versions of the model with the long-term goals of validating the model on other populations, and ultimately, employing the model to make evidence-based recommendations for *T. solium* control.

5.2 Introduction

Cysticercosis is a neglected tropical disease (NTD) that exacts a substantial health and economic burden in low-income countries. The global burden of cysticercosis includes over 5 million people with epilepsy from neurocysticercosis (NCC) [1] and hundreds of millions of dollars in annual livestock losses from discarded pork [8]. Cysticercosis is caused by the pork tapeworm, *Taenia solium,* which is transmitted between humans and pigs, and is common in rural villages where sanitation is limited and small-holder pig farming is practiced. Humans may acquire the adult-stage intestinal tapeworm (i.e., taeniasis) by consuming raw or undercooked pork that is infected with intermediate-stage larval cysts, while pigs acquire this cyst infection (i.e., porcine cysticercosis) through contact with eggs present in the feces of infected humans.

Although global eradication of *T. solium* transmission is unlikely in the short-term, local control or elimination is now possible [94,100] due to the availability of new tools that can be deployed to interrupt transmission. These include effective treatment of taeniasis [48,91] and porcine cysticercosis [34], improved diagnostic tests [95,156], and a vaccine to prevent pig infection [99,100]. In 2012, shortly after the success of a large-scale elimination demonstration in Peru that effectively implemented many of these tools [94], the World Health Organization (WHO) declared ambitious targets for global control and elimination of *T. solium.* They called for validated control strategies to be available by 2015, and for these strategies to be scaled and implemented in several countries by 2020 [127]. Unfortunately, we have not yet identified a set of validated control strategies and the 2020 targets for large-scale implementation are unlikely to be met. Inability to meet these goals is due, in part, to the relative shortage of prospective trials that have been conducted to evaluate available strategies, owing largely to the high cost and time required to conduct these trials.

Infectious disease models are useful tools that can be deployed to identify and validate control strategies when prospective studies are infeasible [128,129]. To this end, in the

"framework for intensified control of *T. solium*" laid out by WHO , modeling was identified as a key sector that will need to be developed in order to reach agreed-upon targets for control [11]. Five models of *T. solium* transmission have been published to date that attempt to fill this gap [12–16]. Despite preliminary use of these models to compare and contrast available control strategies, there are significant concerns about the accuracy and utility of these models that must be addressed before reliable policy recommendations can be made.

First, all of the existing models have structural limitations that prevent them from capturing key heterogeneities in *T. solium* transmission. The most fundamental of these is a spatial structure, which is an important feature for replicating spatially clustered transmission patterns that have been observed in endemic villages [74,75,116,122] and would allow models to evaluate spatially targeted control strategies (e.g., "Ring strategy") that have been successful in field trials [50]. Another limitation of existing models is the lack of an open population structure that allows for human travel and migration, factors that are likely responsible for ongoing transmission and re-introduction of transmission in cleared areas [119,120]. Together, these deficits are likely to result in over-estimation of intervention effectiveness and unrealistic predictions for achieving control and elimination targets [117,121]. While the above structural limitations are significant barriers, perhaps the most serious deficit of existing *T. solium* models is that none have been validated against observations from prospective trials. Comparison of model predictions with observed outcomes is a critical step for ensuring the validity of a model, and should be a pre-requisite for any model that will be used to make evidence-based policy recommendations.

In order to address the above limitations we developed a novel agent-based model (ABM) called "CystiAgent." Notably, CystiAgent includes a spatial structure and open population, which allows it to represent key aspects of the *T. solium* life-cycle, and facilitates evaluation of spatially targeted control interventions. In the current study, we present results from our validation of CystiAgent using data from two large cluster-randomized trials conducted in

northern Peru. Our objectives were to evaluate the ability of CystiAgent to replicate baseline levels of transmission through model calibration and to evaluate the accuracy of the model for predicting observed reductions in transmission when control strategies were applied.

5.3 Methods

5.3.1 Model description

Model structure. CystiAgent is a spatially explicit ABM that simulates endemic transmission of *T. solium* and is able to test population-level control or elimination strategies. The model was developed in NetLogo 6.0.4 (Northwestern University, Evanston, IL), an open-access ABM software that was chosen for its ability to represent spatial data and display simulations through a graphical interface. A detailed description of the model structure and parameters, along with complete sensitivity analysis, can be found is a separate publication (See Chapter 3).

Agents. CystiAgent has two classes of agents – humans and pigs – that interact and transmit *T. solium* in a dynamic spatial environment. Humans may be infected with the adult-stage intestinal tapeworm (i.e., *T. solium* taeniasis) through consuming infected pork, while pigs may be infected with larval-stage metacestodes (i.e., porcine cysticercosis) through contact with *T. solium* eggs or proglottids in the environment. Exposure to eggs may cause light cyst infection (<100 cysts), while exposure to proglottid segments may lead to heavy cyst infection (≥ 100 cysts), and either may lead to seropositivity. Human cysticercosis, including NCC or NCC-related seizure disorders, is not included in this model as it does not contribute to transmission.

Environment. The model environment consists of households that are spatially distributed according to a set of input coordinates. The current model version must be applied separately in each village and functions for villages up to ~2,000 individuals if geographic coordinates for households are known. At baseline, humans and pigs are assigned to households, and households are given characteristics that include the presence of latrines and corrals to

contain pigs. In the current analysis, these characteristics were assigned based on census data that was available for each village included, but simplifications could be made if such data were not available.

Parameters. The activity-flow of the model consists of key biological and behavioral processes that propagate the *T. solium* life-cycle. Briefly, these processes include: the sale, export, and import of potentially infected pigs; pig slaughter, pork distribution, and pork consumption leading to possible tapeworm infection; tapeworm maturity, open human defecation, and environmental contamination with *T. solium* eggs; pig roaming, exposure to infectious eggs, and natural decay of eggs present in the environment; and human travel into and out of the village. Each of the above processes is defined mathematically by a probability distribution and corresponding parameter(s) (Table 5.1).

For the current model validation, a variety of sources, including primary data, expert opinion, and literature review were utilized to determine the values and distributions for model parameters. For the majority of model parameters, we used data collected in the Piura region of northern Peru. Apart from the model's tuning parameters, which were estimated separately for each village and are described below, we used the same version of the model and the same set of core parameter values for all village and all validation analyses described here. A full description of the methods and data sources used to estimate each parameter value can be found in Appendix B2.

Apart from the core biological and behavioral parameters included in the model, CystiAgent utilizes a set of tuning parameters to adjust the rates of transmission to match the observed prevalence of human and pig infection in each study village. To do this, the model includes eight tuning parameters that represent different probabilities of exposure or infection to pigs or humans in the system (Fig 5.1, Table 5.1). Since these parameters represent complex sequences of unknown probabilities and cannot be estimated through primary data or literature review, we developed an approximated Bayesian computation (ABC) algorithm to

mathematically derive their values for this analysis. For the current validation, a unique set of tuning parameters was independently estimated for each of the 36 villages evaluated across the two validation datasets.

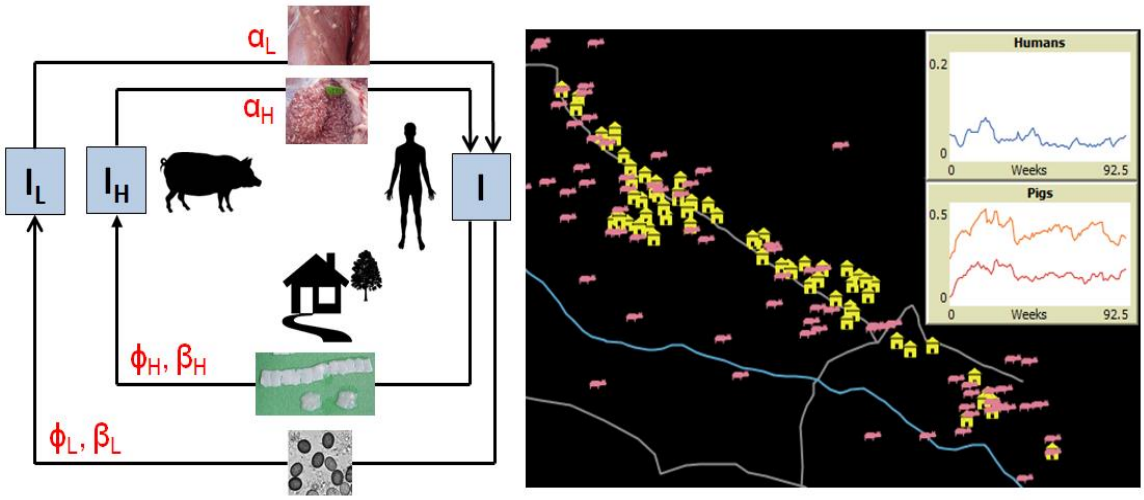


Fig 5.1. Left: diagram of transmission states and select parameters in CystiAgent (left). α_L and α_H are "tuning" parameters that represent the probabilities of human infection (I) after consuming lightly (I_L) or heavily (I_H) infected pork. β_L and β_H represent the probability of light (I_L) or heavy (I_H) cyst infection in pigs after consuming *T. solium* eggs or proglottids, respectively, in the environment. ϕ_L and ϕ_H represent the probability of exposure to *T. solium* eggs or proglottids and are scaled to the current number of tapeworm carriers (HT) according to $1 – (1-\phi)^{HT}$. Serological tuning parameters (ε_L and ε_H) are not shown here. See Table 5.1 for details. **Right: Snapshot of model layout in NetLogo 6.0.4 (right).** Plots represent the prevalence of human taeniasis (top) and porcine cysticercosis (bottom): light (orange) and heavy (red) cyst infection.

Table 5.1. CystiAgent model parameters.

Parameter	Distribution	Value	Source
Village input features			
Humans per household	Poisson	†	NA
Proportion of households raising pigs	Binomial		
Pigs per pig-raising household	Exponential		
Corral prevalence among pig-owner households	Binomial		
Latrine prevalence	Binomial		
Pig sale, slaughter, and tapeworm infection			
*Pig slaughter age (median)	Log-normal	9.8 months	[137]
Proportion of pigs sold prior to slaughter	Binomial	0.51	HH
Proportion of sold pigs exported	Binomial	0.73	HH
Rate of pigs imported from endemic areas (imports / pig / week)	Uniform	0.00105	HH
Prevalence of cyst infection among imports	Binomial	0.134	[75]

Proportion of infected imported pigs with light cyst burden	Binomial	0.76	[75]
Proportion of pork consumed by owner	Binomial	0.40	HH
Proportion of pork sold after slaughter	Binomial	0.12	HH
Proportion of shared pork eaten by owner	Binomial	0.8	HH
Tapeworm maturity and environmental contamination			
Incubation time to reach tapeworm maturity	Fixed	8 weeks	[22,27]
*Tapeworm lifespan (mean, sd = 1 year)	Normal	2 years	[22,27]
Latrine-use (prop. of households that "always" use latrine)	Binomial	0.25	GPS
*Radius of environmental contamination (median, meters from home)	Log-normal	26 meters	GPS
Rate of egg decay in environment (mean survival duration)	Exponential	8 weeks	[134]
Pig roaming and exposure to *T. solium* eggs			
Proportion of pig households with corrals that "always" corral pigs	Binomial	0.05	GPS
Proportion of pig households with corrals that "sometimes" corral pigs	Binomial	0.57	GPS
Proportion of pigs in "sometimes"-corral-households that are corralled	Binomial	0.32	GPS
*Radius of pig roaming "home-range" (median)	Log-normal	44 meters	GPS
Human travel			
Proportion of households with a frequent traveler	Binomial	0.42	HH
Frequency of travel to other endemic areas (every X weeks)	Uniform	8 weeks	HH
*Duration of travel	Exponential	1.75 weeks	HH
Incidence of *T. solium* taeniasis during travel (risk / person / week)	Uniform	0.00023	[75]
Tuning parameters			
α_L, α_H: Probability of human taeniasis upon consumption of lightly or heavily infected pork	Binomial	†	NA
β_L, β_H: Probability of light or heavy cyst infection upon exposure to *T. solium* eggs or proglottids in the environment (respectively)	Binomial		
ϕ_L, ϕ_H: Probability exposure to *T. solium* eggs or proglottids in the environment (respectively, per existing tapeworm carrier)	Binomial		
ε_L, ε_H: Probability of pig seropositivity upon exposure to *T. solium* eggs or proglottids in the environment (respectively)	Binomial		

*Unique values from distribution randomly assigned to individuals
†Values for 5 village input features and 8 tuning parameter were determined individually for each village
HH=Household survey; GPS=GPS pig tracking study; NA=not applicable; see Appendix B1 for details on primary data collection.

Interventions. CystiAgent has the ability to simulate a variety of population-level interventions designed to control or eliminate *T. solium* transmission. A generic function is available to administer anti-helminthic treatment for human taeniasis, either presumptively or after stool screening. Other functions include the treatment of pigs to cure cystic larval infection, or vaccination to prevent infection. For each intervention type, user-controlled options allow for specification of participation levels, the sensitivity of screening tests, and the efficacy of drugs and vaccines used. These interventions can then be implemented through mass or targeted approaches, while varying the duration and frequency of intervention applications. Unique to this spatial model is the ability to simulate spatially targeted interventions. "Ring strategy" [50] can be applied by targeting treatment resources to households residing within a given distance of heavily infected pigs. Finally, behavioral interventions such as improved access to corrals and latrines, along with corresponding rates of usage, are available as stand-alone interventions or in combination with other approaches.

5.3.2 CystiAgent model validation

We validated the CystiAgent model using data from two large prospective trials completed in Peru: (1) the Ring Strategy Trial (RST) and (2) the Cysticercosis Elimination Demonstration Project (CEDP). Validation attempts were performed individually for each village in the parent study. For each village included in the validation, a unique set of household coordinates and characteristics were applied to each village based on available census data from the studies, a unique set of model tuning parameters was estimated using our ABC calibration algorithm in order to match observed baseline prevalence levels in each village, and unique rates of participation in the interventions were applied. Other than these village-specific settings, the exact same model structure and set of core parameter values were used in all villages across the two datasets. The intervention sequences that were carried out in the field were simulated in the model villages (n=1000 simulations per village), and model outputs were compared to observed outcomes through quantitative and graphical analyses.

Ring Strategy Trial (RST). This prospective trial was carried out in 23 villages in the Piura region of northern Peru over a period of 24 months (2015-2017). A detailed description of this trial is currently under peer review [contact Seth O'Neal]. Study villages were divided into six arms each receiving unique intervention designs (Table 5.2). Four of the six arms received some variation of "Ring Strategy". This is an approach that targets antihelminthic treatment to humans and pigs that live within 100 meters of pigs found to be heavily infected through non-invasive tongue palpation [44]. Between the four ring-strategy arms, the approach varied based on the intervention applied within 100-meter rings. For humans, the two options included presumptive treatment of human taeniasis with two doses of oral niclosamide (NSM), or stool screening for taeniasis with the enzyme-linked immunosorbent assay for copro-antigen detection (CoAg-ELISA) and repeated follow-up treatment with oral NSM until cure. Treatment of pigs with a single oral dose of oxfendazole (OFZ) according to pigs' weights was added to the human interventions in a factorial design. All ring interventions were administered in study villages every 4 months throughout the 2-year study-period. The remaining two study arms received mass applied interventions in 6-month intervals. These two interventions included mass treatment of humans with a single oral dose of NSM, and combined human and porcine mass treatment. At the conclusion of the study, all human participants were offered NSM, and post-treatment stool samples were analyzed with CoAg-ELISA to evaluate the prevalence of human taeniasis. Additionally, serum samples from all pigs in the study communities were collected every four months and evaluated with the enzyme-linked immune-electro transfer blot (EITB) to determine the incidence of seroconversion among serial cohorts of pigs [45].

Table 5.2. Summary of intervention strategies used for model validation. Ring Strategy Trial (top, 21 villages) and Cysticercosis Elimination Demonstration Project (top, 15 villages).

Strategy	Interventions	Population	Data source
Ring treatment	Pig tongue screening, human treatment in rings (q4 months, 7x)	2 villages (~1200 humans, ~400 pigs)	Ring Strategy Trial
Ring treatment w/ pig treatment	Pig tongue screening, human and pig treatment in rings (q4 months, 7x)	4 villages (~1600 humans, 500 pigs)	
Ring screening	Pig tongue screening, human screen-and-treat in rings (q4 months, 7x)	4 villages (~1500 humans, 600 pigs)	
Ring screening w/ pig treatment	Pig tongue screening, human screen-and-treat, pig treat in rings (q4 months, 7x)	4 villages (~1500 humans, 400 pigs)	
Human mass treatment	Human MDA (q6 months, 5x)	4 villages (~1300 humans, ~400 pigs)	
Combined mass treatment	Human and pig MDA (q6 months, 5x)	3 villages (~1400 humans, ~500 pigs)	
Combined mass treatment w/ vaccine	Human MDA (3x) + Pig MDA (5x) + Pig vacc.	2 villages (~3500 humans, ~1100 pigs)	Cysticercosis Elimination Demonstration Project
Combined mass treatment	Human MDA (3x) + Pig MDA (5x)	6 villages (~2900 humans, ~1700 pigs)	
Combined mass screening w/ vaccine	Human screen-and-treat (2x) + Pig MDA (5x) + Pig vacc.	3 villages (~2200 humans, ~1300 pigs)	
Combined mass screening	Human screen-and-treat (2x) + Pig MDA (5x)	4 villages (~2400 humans, ~1300 pigs)	
TOTAL		**36 villages (~19000 humans, ~8000 pigs)†**	

*MDA = Mass drug administration, q=frequency; x=repetitions
†4 villages excluded (2 CEDP, 2 Ring Strategy) due to lack of observed transmission and/or data

Cysticercosis Elimination Demonstration Project (CEDP). This prospective trial was carried out in 17 villages in the Tumbes region of northern Peru over a 9-month period in 2006-07. Results of the trial along with a detail study protocol are published elsewhere [94]. Briefly, study villages were divided into four arms, all consisting of mass-applied interventions. All study arms included mass application of anti-helminthic treatment to pigs (OFZ) every 2 months (5 rounds total). This intervention was then combined with either presumptive treatment of human taeniasis (single-dose NSM), or stool screening (CoAg-ELISA) with follow-up NSM treatment. Human screen-and-treat approaches were administered in months 2 and 5 of the study, while presumptive human treatment was administered in months 2, 5, and 9. Each of these arms were

further divided into a vaccination and non-vaccination arm. The vaccination arm included administration of the TSOL18 vaccine [99] to pigs in two doses during month 5. For monitoring the intervention, humans in all study arms were screened for taeniasis at months 2 and 5, and pig serum was tested for antibody response with the EITB assay over eight time-points before, during and after the study. Necroscopic examination of selected seropositive pigs was conducted shortly after the conclusion of the study, and 12 months later to assess elimination status.

Parameter estimation. We used an approximated Bayesian computation (ABC) algorithm to estimate model tuning parameters separately for each village. Generally speaking, ABC methods are well-suited to complex non-linear models because they approximate likelihood functions empirically by running the model through Monte Carlo simulations [135,157]. The specific ABC method we employed followed a simple "rejection sampling" approach and was based on a variety of in-depth examples found in literature [152,158,159].

Our ABC method was used to estimate values for the six tuning parameters that define the probabilities of human and pig infection in the model (β_L, β_H, α_L, α_H, ϕ_L, and $\phi_{H;}$ see Fig 5.1 for details) and two serological parameters (ε_L and ε_H) defining the probability of antibody response after exposure for pigs. Of the 40 villages that participated in the trials, two villages in the RST were excluded due to few or no infected pigs being detected during the trial and two villages in the CEDP trial were excluded because household geographic coordinates were not available. This left a total of 36 villages for ABC estimation and subsequent validation analysis: 21 villages from the RST and 15 villages from the CEDP trial. For each village, we used a Sobol' quasi-random sequence [144,145] to sample 5,000 values from a uniform distribution for each of the tuning parameters. Because each of the tuning parameters represents a probability, we set initial limits of the distribution to between 0 and 1, but narrowed the starting ranges to improve efficiency after the first few villages. For each combination of parameter values, we ran the model though 1000 weeks of stable endemic transmission and recorded summary statistics at the end of each model run. For humans, we recorded the average prevalence taeniasis across the

simulation period, and for pigs, we recorded the average seroprevalence, and prevalence of light and heavy cyst infection. The "sensitivity" package in R was then used to calculate the Euclidean distance between these summary statistics and the baseline prevalence observed in the field data [135,160]. Following a rejection sampling scheme, we selected the top 1% of model runs that minimized the Euclidean distance and extracted posterior distributions from the selected parameter sets. We then repeated the algorithm with 10,000 parameter combinations sampled from these posterior distributions and selected our final parameter values based on the median values of the new posterior distributions produced for each parameter in this final step.

In order to achieve the computational resources needed to run the model through many thousands of simulations for each of village, we executed all model simulations on the Amazon Web Service EC2 cloud computing platform. Model simulations were distributed across multiple 72-core parallel processors using the "parallel" R-package [138] and executed on the EC2 cloud using the R-Studio Shiny server [139].

Simulation of intervention strategies. After calibration of model tuning parameters was complete for each village, we proceeded with validation of the model against observations from prospective trials. For simulation of the intervention strategies, we set up the model with unique sets of tuning parameters and input characteristics for each village, and applied the corresponding sequence of interventions that each village received in its field trial. Identical sets of core behavioral/biological parameters values (Table 5.1) were applied in all villages. The levels of participation that were applied to humans and pigs in the model reflected the participation rates that were observed in each village during the application of interventions (Table 5.3). Appendix C1 contains a detailed description of all intervention settings applied for model validation. To briefly summarize key settings, the sensitivity of the CoAg-ELISA for detecting *T. solium* taeniasis was set to 96.4% [156], and the efficacy of NSM for treatment of human taeniasis was set at 76.6% for one dose, 86.6% for two doses, and 93.3% for post-screening follow-up [93]. Treatment of pigs with OFZ was assumed to have an efficacy of 100% [96], while the efficacy of

the TSOL-18 vaccine was set to 99% for pigs receiving two doses [99,100] (no protective benefit was assumed for a single dose [161]).

Table 5.3. Village-specific participation rates applied to model simulations.

	Ring Strategy Trial (21 villages)	CEDP (15 villages)
Humans		
Rings		
Stool screening (CoAg-ELISA)	83.3% (74.5-89.5)	NA
Treatment (NSM)		NA
1 dose	16.1% (12.0-27.0)	NA
2 doses	70.7% (61.0-76.0)	NA
Post-screening treatment	91.8%†	
Mass application		
Stool screening (CoAg-ELISA)	NA	78.5% (64.5-85.5)
Treatment (NSM, 1 dose)	75.5% (70.3-81.5)	78.3% (72.1-85.0)
Post-screening treatment	NA	91.8%†
Pigs		
OXF treatment	68.9% (26.4-90.0)	90.6% (86.6-94.0)
Vaccine (TSOL18)	NA	
Round 1	NA	89.2% (78.4-95.9)
Round 2 (booster)	NA	85.5% (72.0-93.0)

†Applied uniformly in all villages

All simulations began with a 1000-week burn-in period, followed by the corresponding intervention sequence and 100-week post-intervention period. For each village, we repeated the same intervention sequence 1000 times. As above, we executed all model runs for the validation on a 72-core parallel processor using the Amazon Web Service EC2 cloud-computing platform.

Statistical analysis. For each village evaluated, we compared the model-predicted prevalence of human taeniasis, porcine cysticercosis, and pig seroprevalence (or seroincidence) with the prevalence observed in the corresponding field trial. For porcine cysticercosis, besides the final-round necroscopic examination performed in the CEDP trial, all prevalence estimates were generated by applying simple proportions to the number of seropositive pigs detected, which were based on prior necropsy studies conducted in the region [43,94,162]. For human taeniasis, the final-round prevalence in the RST and baseline prevalence in the CEDP trial were directly measured, but baseline prevalence in the RST was not measured. Therefore, we developed a

regression equation to estimate baseline prevalence in each village based on pig seroprevalence [75]. Appendix C1 contains a detailed description of the methods used to analyze observed field data and adjust outcomes for model comparison.

For each of the outcomes listed above, we calculated the mean absolute error (MAE) of model predictions against the observed summary statistics. MAE is a useful measure of model performance, as it captures both the accuracy and precision of model predictions by calculating the average distance between all model runs and the observed value (Equation 1). MAE was calculated separately for each outcome (human taeniasis, porcine cysticercosis, seroprevalence/incidence) for all time-points and villages. Results were then compared between villages, intervention types, and studies to assess the overall performance of the model.

$$Eq.\,1: \quad MAE = \frac{\sum_{i=1}^{n}|x_i - x|}{n}$$

5.4 Results

5.4.1 Baseline calibration

The average prevalence of human taeniasis (HT) at baseline was 1.8% (range: 0.3% to 6.6%) across the 36 villages included in the validation. Baseline taeniasis was slightly higher in the RST villages compared to the CEDP (2.1% vs. 1.4%; p = 0.13). For porcine cysticercosis (PC), the average prevalence at baseline across all 36 villages was 12.7% (range: 1.9% to 26.9%). There was considerable variation in this measure across all villages, and it was significantly higher among RST villages compared to CEDP (17.8% vs. 5.5%, p < 0.01).

The results of model calibration are shown in Fig 5.2. In most cases, the model was able to accurately duplicate levels of transmission observed in study villages at baseline. Villages that had higher MAE at baseline were more likely to have smaller populations (< 75 households). Between the two studies, calibration was more precise for CEDP villages (p < 0.01), likely due to the lower prevalence observed in these villages, which reduced variability in model outputs. No

differences in baseline calibration were observed between intervention types. Despite differences in precision, the median model outputs closely approximated the observed baseline prevalence for the majority of villages in both datasets, illustrating high overall accuracy of the model calibration (see Appendices C2 and C3 for village-specific results).

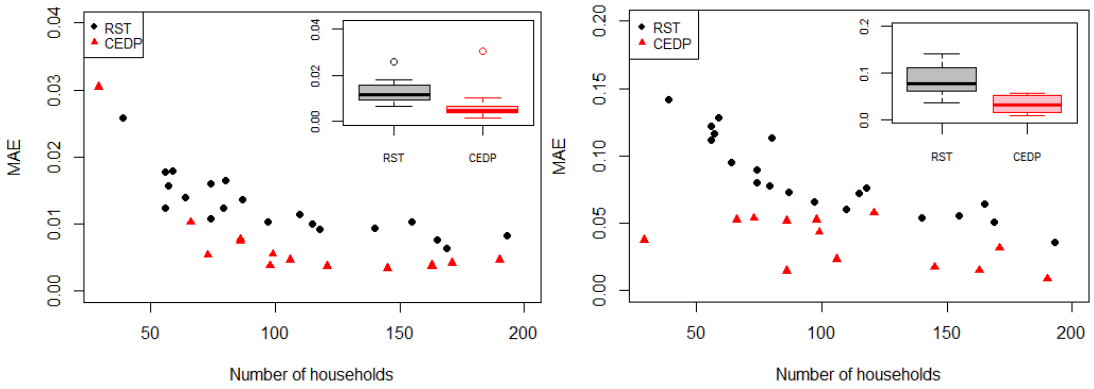


Fig 5.2. Mean absolute errors (MAE) for human taeniasis (left) and porcine cysticercosis (right) for baseline calibration. Scatter plots show the relationship between MAE and village population; inset boxplots show direct comparison of MAE by study: RST = Ring Strategy Trial (21 villages), CEDP = Cysticercosis Elimination Demonstration Project (15 villages).

5.4.2 Evaluation of CystiAgent with prospective trials

Ring Strategy Trial. In the Ring Strategy Trial (n=21 villages), the average prevalence of HT observed at the conclusion of the study was 0.8% (range: 0 to 4.3%), an average absolute reduction of 1.2% (range: 2.5 to -0.4%) from the estimated baseline prevalence. Three of 21 villages were found to be free of HT at study end. The estimated prevalence of PC at the conclusion of the study was 8.5% (range: 2.0 to 23.3%), an average absolute reduction of 9.3% (range: 18.1 to -2.7%) from the estimated baseline prevalence. None of the villages achieved completed elimination of HT and PC by the end of the trial.

A comparison of CystiAgent model predictions with observed results from the RST are shown in Fig 5.3 and all results are available by village in Appendix C2. In the majority of villages, CystiAgent predicted a stronger intervention effect than was observed in the actual trial. The median predicted prevalence of HT at the conclusion of the study averaged 0.1% (range: 0 to 0.4%) across the 21 study villages. These predictions were, on average, 0.8% lower (range: 4.2%

to -0.3%) than the corresponding prevalence of HT observed at study-end. For pigs, the median predicted prevalence of PC averaged 2.7% (range: 0 to 6.4%) at study-end, which measured 5.8% lower (range: 18.9% to -4.1%) than the estimated final-round prevalence of PC in the study villages. The model predicted *T. solium* elimination (i.e., median predicted prevalence = 0 for HT and PC) in two of the 21 villages by the end of the trial, and six villages by the end of the 100-week post-intervention period. Villages for which elimination was predicted had slightly lower baseline prevalence (1.9% vs. 2.2% for HT, p=0.62; 16.5% vs. 18.3% for PC, p=0.52), but did not differ by any other discernable factors, including participation rate or intervention type.

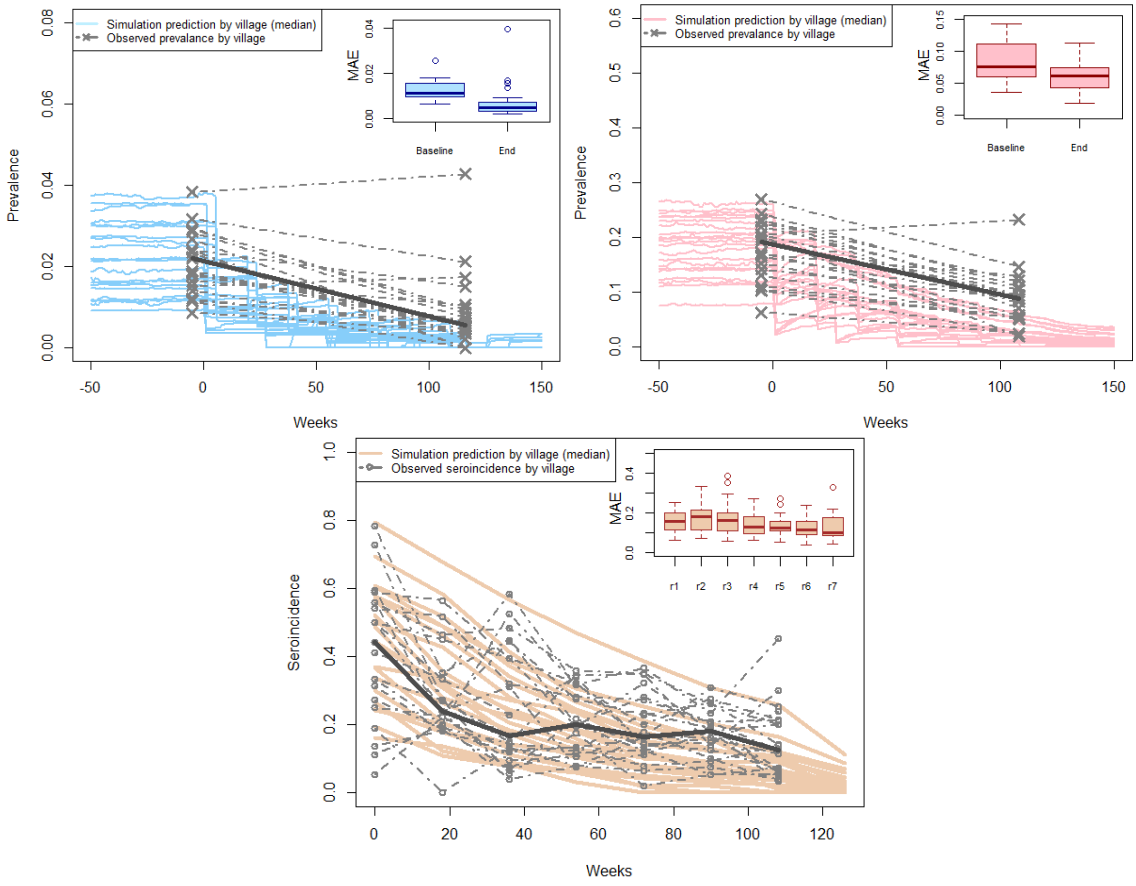


Fig 5.3. Simulation predictions versus observed rates of human taeniasis, porcine cysticercosis and porcine seroincidence in the Ring Strategy Trial. Graphs display the village-specific median predicted values (n=1000 simulations per village). Inset boxplot display the mean absolute error (MAE) of all prediction runs against the observed value at each time point (21 villages total).

Despite over-estimation of the intervention effect in most villages, MAE for both HT and PC were lower at study-end than at baseline (Fig 5.4). The most important factors associated with the accuracy of final-round predictions were the baseline prevalence of HT and PC in the study villages. Villages with lower baseline prevalence of HT and PC achieved lower final-round prevalence, which led to significantly improved final-round MAE values for both outcomes ($p<0.01$ for HT, $p=0.036$ for PC, respectively). Besides baseline prevalence, no other village-level features, including intervention type, population size, and participation rate, were associated with improved accuracy of model predictions.

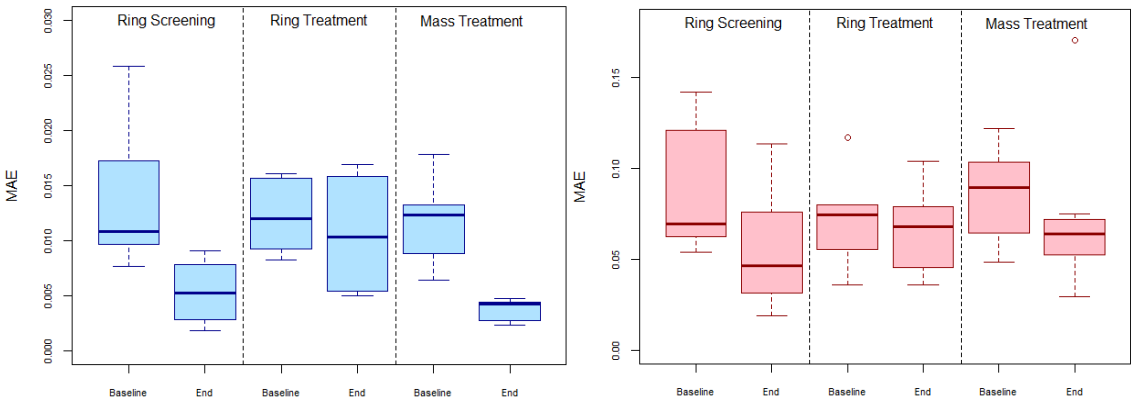


Fig 5.4. Mean absolute error (MAE) of model predictions for human taeniasis (left) and porcine cysticercosis (right) in the Ring Strategy Trial. Baseline and final-round accuracy is shown for each intervention arm (21 villages total).

Comparison of intervention types. Between the three primary interventions applied in the RST, the model predicted that each strategy would produce similar and substantial reductions in HT and PC by study-end (see Appendix C4 for graphical results by intervention type). This differed from observed results of the trial, which found that ring-screening and mass-treatment led to larger reductions in transmission compared ring-treatment. As a result of their larger effect sizes, the accuracy of model predictions for the final-round prevalence were slightly improved in ring-screening and mass-treatment arms compared to ring-treatment (Fig 5.4). With respect to the additional treatment of pigs with OFZ in half of the study arms, the model predicted a small

but non-significant additive effect of pig treatment on the final-round HT and PC prevalence, which is consistent with what was observed in the RST.

Seroincidence. In the observed RST data, the 4-month cumulative incidence of pig seropositivity was reduced from an average of 41% (range: 5.3 to 78.3%) at baseline to 14.7% (range: 3.3 to 45.2%) in the final sero-survey, which corresponds to an average absolute reduction of 26.7% (range: 1.6 to 71.3%). Similar to model predictions for HT and PC described above, final-round predictions for seroincidence were lower (7.8% lower; range: 36.2% to -18.7%) than their corresponding observed values (Fig 5.3). Prediction accuracy across the seven seroincidence measurements was poorest in rounds 2 and 3, time-points at which true seroincidence declined steeper than model predictions. In rounds 4 through 7, however, accuracy improved as model-predicted seroincidence continued to decline while the observed seroincidence plateaued. Between intervention arms, this sharp pattern of decline followed by a plateau in seroincidence was most apparent in the ring-screening arm, while ring-treatment and mass-treatment illustrated more gradual declines across the study period (see Appendix C4 for graphical results by intervention type). CystiAgent was not able to replicate the sharp initial decline in seroincidence observed in ring-screening, but accurately predicted the more gradual slope of seroincidence observed in most ring-treatment and mass treatment villages.

CEDP Trial. In the Cysticercosis Elimination Demonstration Project (n=15 villages), the average prevalence of HT decreased from 1.4% (range: 0.3 to 6.6%) at baseline to 0.8% (range: 0 to 2.7%) mid-way through the study. HT was not measured at study end. The average estimated prevalence of PC was reduced from 5.5% (range: 1.9 to 8.9%) at baseline to elimination of transmission at study-end, which was confirmed with necroscopic examination of pigs at both the conclusion of the study and one-year post-intervention.

Model predictions for the CEDP trial are shown in Fig 5.5, and all validation results by village are shown in Appendix C3. At the mid-study measurement of HT (month 5), the average predicted prevalence of HT was 0.4% (range: 0 to 2.1%) which is 0.4% (range: 1.9 to -0.4%) less

than the observed mid-study prevalence. Following this mid-period reduction, the model predicted elimination (i.e., median predicted prevalence = 0 for HT and PC) in 9 out of 15 villages by study-end, and 2 out of 15 villages at one-year post-intervention.

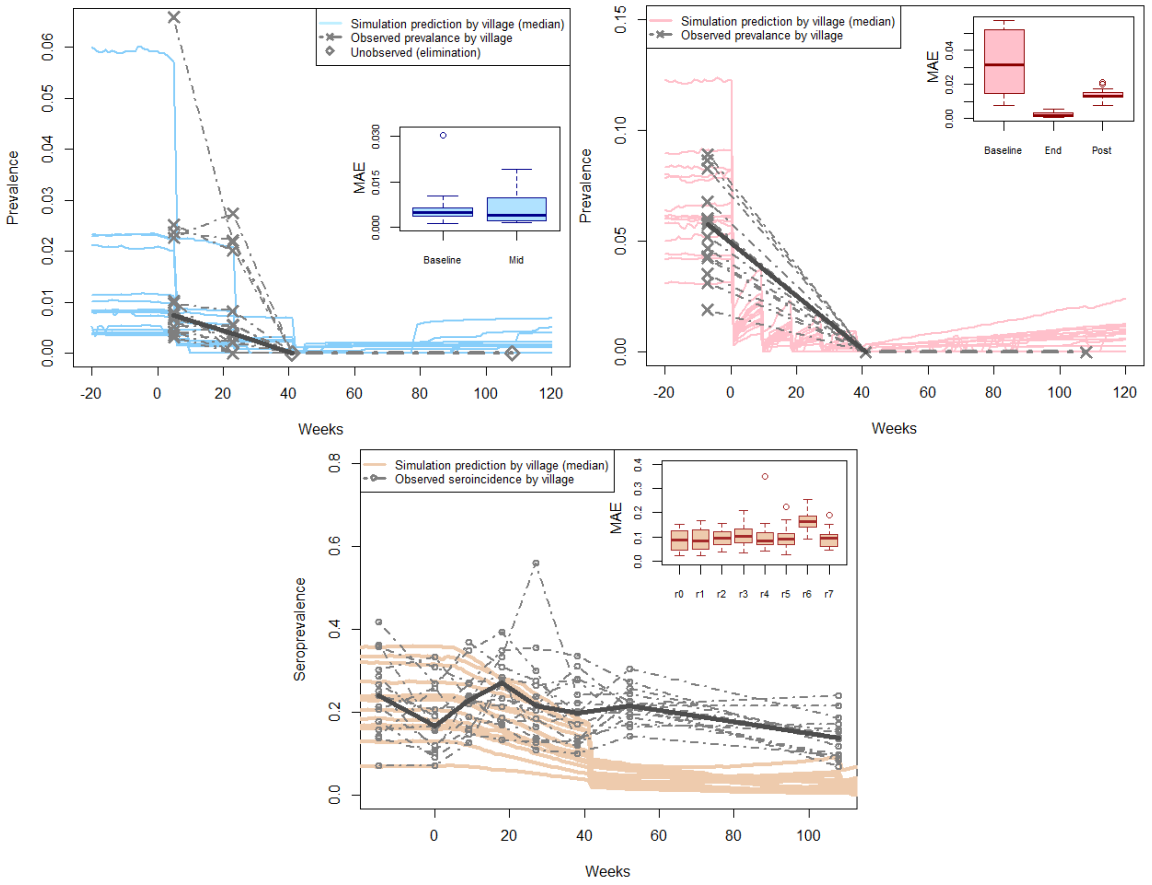


Fig 5.5. Simulation predictions versus observed rates of human taeniasis, porcine cysticercosis and porcine seroprevalence in the CEDP Trial. Graphs display the village-specific median predicted values (n=1000 simulations per village). Inset boxplot display the mean absolute error (MAE) of all prediction runs against the observed value at each time point (15 villages total).

The MAEs of model predictions varied considerably between outcomes assessed (Fig 5.6). For pig infection, MAEs were highest and most variable at baseline, but decreased sharply at study-end as elimination of transmission was correctly predicted in most villages. Errors for PC at the post-intervention time-point then rose as rebounds in transmission were predicted in most villages, despite continued elimination in the observed data. The only factor that significantly impacted model accuracy at study-end was the intervention type, as mass screening villages were

more likely to have more accurate predictions (lower MAE) compared to mass treatment villages, largely due to correct predictions of elimination in mass screening villages. At the post-intervention follow-up, villages with a higher baseline prevalence of PC were more likely to have stronger predicted post-elimination rebounds in transmission, and, as a result, increased MAE measures (p=0.046) compared to the observed elimination status in study villages.

For HT, errors were only calculated for the first two time-points available, and were not significantly different between the baseline and mid-period time-points (p=0.86). Prediction accuracy for the mid-period HT measurement was improved in mass screening villages (p=0.03) compared to mass-treatment, and was improved for villages with lower baseline and mid-period HT prevalence (p<0.01 for both). Population size and participation levels were not associated with improved prediction accuracy for either HT or PC.

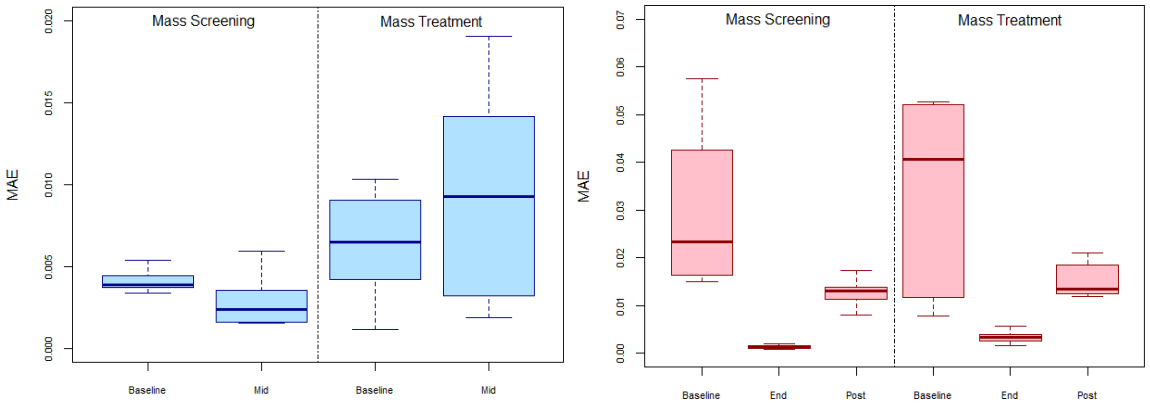


Fig 5.6. Mean absolute error (MAE) of model predictions for human taeniasis (left) and porcine cysticercosis (right) in the CEDP Trial. Baseline and final-round accuracy is shown for each intervention arm in the CEDP trial (15 villages total).

Comparison of intervention types. Consistent with observations in the CEDP trial, all seven villages participating in mass screening interventions were predicted to achieve elimination at study-end, with two of these villages predicting sustained elimination throughout the post-intervention phase (Appendix C5). Of the 8 villages participating in mass treatment interventions, only two were predicted to achieve elimination at study-end and one was predicted to have sustained elimination. Apart from intervention type, for which significant differences were

observed in elimination probability ($p<0.01$), no other village-level factors, including participation, population size, and baseline prevalence, were associated with predicted elimination at study-end. The addition of pig vaccination in 5 of the 15 study villages did not have a detectable impact on transmission or the likelihood of achieving elimination. The model predicted elimination by study end in four of these five villages. There was no difference in the magnitude of post-intervention rebound or prediction error in the post-intervention measurement based on the addition of pig vaccination.

Seroprevalence. In the observed CEDP data, the prevalence of seropositivity among pigs (2+ EITB bands) paradoxically increased between the baseline measurement (18.6%; range: 7.2 to 33.3%) and the measurement taken at study-end (19.8%; range: 9.9 to 33.5%); however, a small decrease was observed at the final post-intervention measurement (13.9%; range: 6.9 to 23.9%). The CystiAgent model predicted median baseline seroprevalence of (21.9%; range: 6.9 to 35.8%), an average of 2.0% (range: -6.2 to 9.1%) greater than the observed baseline seroprevalence. At study-end and post-intervention, the predicted seroprevalence averaged 8.0% (range: -2.8 to 23.8%) and 10.4% (range: 4.3 to 23.0%) lower than the observed seroprevalence at those time points, respectively. Prediction errors for seroprevalence were consistent throughout the core phase of the intervention, but increased sharply at time-point 6, when predicted seroprevalence declined without a corresponding decline in the observed seroprevalence. This drop was caused by the mass necropsy that took place at the conclusion of the intervention study and removed ~50% of seropositive pigs, but was not reflected by a corresponding drop in seroprevalence in the observed data.

5.5 Discussion

In this research, we developed a novel transmission model for *T. solium* (CystiAgent), and validated the model using data from two large cluster-randomized trials conducted in Peru.

Development and validation of a *T. solium* transmission model was identified as a key goal in WHO's 2014 framework for intensified control [11], yet despite the publication of three novel *T. solium* models in the intervening years, no model, to our knowledge, has been tested against observed data. This research, therefore, represents an important step towards delivering a validated a model that can be used to evaluate *T. solium* control strategies. Our results showed that CystiAgent consistently and accurately replicated baseline levels of transmission in the 36 study villages assessed, and, in most cases, accurately modeled declines observed in transmission when control strategies were applied; although accuracy varied considerably between villages, studies, and interventions applied.

5.5.1 Baseline calibration

With respect to reproducing baseline levels of transmission, CystiAgent was tuned specifically to transmission patterns in northern Peru by applying behavioral and environmental parameters that were collected in the Piura region of Peru, and then employing an ABC calibration algorithm to estimate values for model tuning parameters. While the majority of model parameters were held constant for all 36 villages modeled (see Table 5.1), we found it necessary to perform the ABC calibration to estimate unique values for model tuning parameters in each individual village due to variations in prevalence between villages. Future uses of this model to prospectively compare intervention strategies are not likely to require this level of rigorous local calibration. A single set of model tuning parameters could be calibrated to match the presumed baseline prevalence of a larger region, rather than tuning the model to individual village levels. In order to ensure that the optimal set of tuning parameters is selected in future model applications, the full range of possible tuning parameter values should be included in ABC calibration (uniform distributions between 0 and 1).

5.5.2 Validation of intervention strategies: CEDP Trial

When we applied control interventions to the model, we found that the model adequately predicted observed reductions in transmission, although predicted intervention effects tended to be slightly overestimated compared to observed results. The accuracy of model predictions depended on the study, with significantly lower prediction errors measured in the CEDP trial compared to the RST. In the CEDP trial, the model predicted elimination of transmission in 9 of 15 villages by study-end and 2 of 15 villages in the post-intervention period. Given that elimination was the end-point observed in the CEDP trial, model accuracy was very high in these villages. Of course, complete elimination of transmission is very difficult to confirm in the field, and it is likely that some infected pigs were missed in the final CEDP necropsy sample, or some tapeworm carriers persisted (as final-round taeniasis was not measured in the study). Therefore, model runs that predicted reductions of transmission to near elimination levels, and those that predicted a slight resurgence in transmission during the post-intervention phase should not be categorically considered as inaccurate. A final-round measurement of HT prevalence, and a more complete necropsy sample at study-end would have allowed for a more thorough assessment of model accuracy in this trial.

Given the uncertainty in final-round and post-intervention measures in the CEDP trial, the mid-study measurement of taeniasis prevalence is the most reliable measure with which to assess model accuracy. This measurement showed that our model slightly overestimated the initial effect of interventions on HT prevalence (average prevalence of 0.4% predicted vs. 0.8% observed). There are a few possible explanations for this error in the model's predictions. One is that the efficacy of NSM for curing taeniasis was overestimated in the model. While the applied single-dose efficacy of 76.6% has been reported in literature [93], variations in manufacturer formulation and shelf-life could impact the actual efficacy, and anecdotal reports from the field suggest that efficacy could be 5-10% lower than the published values. It is also possible that systematic non-participation of tapeworm carriers at baseline or migration of tapeworm carriers

into the study village during interventions could be responsible for prediction errors, although both participation rates and imported tapeworm infections were accounted for in model parameters. Finally, it is possible that immunity and acquired resistance to *T. solium* infections in pigs is reduced when control pressure is applied. Pig immunity is not explicitly included in CystiAgent, and model tuning parameters were fit to replicate natural levels of endemic transmission prior to the application of control interventions. Reductions in transmission caused by repeated treatment could reduce acquired immunity among pigs and lead to more rapid re-introduction of transmission after interventions are applied.

In contrast to the overall accuracy of HT and PC predictions in the CEDP study, predicted seroprevalence was relatively inaccurate. Serological measures of *T. solium* exposure are known to be highly volatile [46] and poorly specific [46,163] for representing true cyst infection in pigs, thus poor model accuracy for this measure is not surprising. Even so, the observed increases in seroprevalence over the course of the trial were unexpected and not able to be replicated by the model. This increasing seroprevalence does not have an obvious explanation based on our current knowledge of pig immunology [34,161]. In order to improve accuracy of CystiAgent for predicting serological outcomes, improved knowledge of the mechanisms of antibody response in pigs and the cause of false positivity in diagnostic test will be required. Despite the possible sources of error outlined above, overall model accuracy for the CEDP trial exceeded expectations, and indicated that CystiAgent was both well-calibrated to *T. solium* transmission in this region and able to appropriately model CEDP interventions.

5.5.3 Validation of intervention strategies: Ring Strategy Trial

Validation of CystiAgent against the RST produced slightly less accurate results compared to the CEDP trial. In 20 of the 21 villages tested, the model predicted intervention effects that were stronger than those observed in the actual study. This exaggerated intervention effect was particularly apparent for PC as an outcome, for which predicted final-round prevalence was an average of 5.8% lower (average prevalence of 2.7% predicted vs. 8.5% observed) than

their observed values. The final-round estimates for PC prevalence in this trial, however, should be interpreted cautiously since a necroscopic examination was not performed. Final-round prevalence of cyst infection in pigs was estimated by assuming that 44% of seropositive pigs at the conclusion of the study were truly infected, a figure based on two necropsy studies previously conducted in the region [43,162]. However, these necropsy studies were conducted in populations that had not received intervention programs. Based on our experience in the CEDP trial, rates of EITB false positivity increase after treatment is applied and transmission is controlled. Therefore, it is likely that the true final-round prevalence in the RST was lower than estimated and model predictions for PC were more accurate than the measured error-rates indicated.

While considerable uncertainty surrounds our measurement of PC at study-end, the final-round prevalence of HT was directly measured in all participants and serves as a reliable gauge of model accuracy in the RST. For HT, CystiAgent predicted a reduction in prevalence that was, on average, 0.8% lower (average prevalence of 0.1% predicted vs. 0.9% observed) than the reduction observed in the RST villages. Elimination of HT was achieved and correctly predicted by the model in three of the RST villages. Similar to the CEDP trial, the most likely explanations for the slight but consistently exaggerated effect sizes in the RST validation include overestimation of the efficacy of NSM for curing taeniasis, systematic non-participation of tapeworm carriers in treatment interventions, and the inability to account for dynamic changes in population immunity caused by the intervention.

In contrast to the poorly modeled seroprevalence trends in the CEDP trial, CystiAgent predictions of pig *seroincidence* were more consistent and accurate in the RST. Seroincidence is a more direct measure of current transmission levels and therefore was more reliably predicted due to less volatility in immune responses. Despite accurate predictions overall, CystiAgent was unable to replicate the unique shape of the seroincidence curve observed in the ring-screening intervention, which consisted of a sharp initial decline followed by a plateau. One possible explanation for the shape of decline in the observed data is highly clustered transmission patterns,

which would have led to rapid reductions in transmission when ring screening was applied, yet stagnation in later rounds as fewer heavily infected pigs were detected to initiate new rings. While the spatial structure of CystiAgent is designed to model clustered transmission patterns, the degree of clustering is determined by parameters dictating the home-range area of pigs, open defecation practices, and pork consumption patterns. Thus, misspecification of any of these parameters could have led to an incorrect degree of spatial clustering and inability to model the observed sharp decline in transmission.

5.5.4 Comparison of intervention strategies

While this validation study was not designed to evaluate the effectiveness of the intervention strategies themselves, the differences that were observed between studies and intervention arms merit brief consideration. In both the model and observed studies, the CEDP trial achieved greater reductions in transmission, and eliminated transmission in more villages, compared to the RST. This is due to the intensive approach employed by the CEDP trial, which consisted of repeated mass application of drugs and vaccine to humans and pigs with goal of rapidly interrupting the *T. solium* life-cycle and causing elimination. Our model predicted that mass screening of the human population was more likely to achieve elimination than presumptive mass treatment, an observation that is not surprising given the increased efficacy of NSM when targeted to cure. Thus, the addition of pig vaccination in 5 of the 15 study villages did not have a discernable impact on our estimates of intervention effectiveness; this is likely because the vaccine was not applied until mid-way through the intervention, and was only applied at two points in time. Our finding is consistent with projections from other *T. solium* models suggesting that vaccines must be applied with high coverage and repeatedly over many months in order to adequately protect the pig population and prevent transmission [16,164].

In contrast to the CEDP trial, the RST was a control strategy and was not designed to reach elimination levels. Between the three intervention arms in RST (ring screening, ring treatment, and mass treatment), CystiAgent predicted that all three would result in similar

reductions in transmission. However, in the actual trial, ring screening and mass treatment produced stronger control effects than ring treatment. As described above, improvements to model specifications will be required to accurately differentiate the effects of these three strategies. Nonetheless, the effectiveness of ring interventions overall indicates that they represent promising alternatives to mass-applied interventions, and merit consideration as viable control options.

5.5.5 Strengths and limitations

There are a few important strengths and limitations of our approach to highlight. First, and most importantly, this research represents the first attempt to validate a *T. solium* transmission model, which was accomplished by using two large prospective trials of *T. solium* control strategies. Using these datasets, the model was tested independently in 36 separate village and 10 unique intervention types, leading to a robust assessment of the model's accuracy in a variety of transmission settings. The two separate trials tested were from two different region of Peru (Piura and Tumbes). While the two regions have similarities in pig-raising practices, the consistency of the model's performance across regions in Peru provides initial support for the generalizability of the model to other transmission settings. Finally, the ABC parameter calibration tool we implemented allowed us to accurately replicate baseline levels of transmission in the 36 study villages, which permitted an unbiased assessment of the model's ability to reproduce intervention effects that were observed in the studies.

A potential limitation of our approach is that parameters used for this validation were primarily sourced from a single region of northern Peru (Piura region). While the depth of data available from this region made it possible to develop this detailed ABM and we successfully validated the model on two datasets from two different regions within Peru, some key parameters may differ in other endemic areas of the world, which could impact the model's accuracy when applied to these regions. Application of CystiAgent outside of northern Peru, therefore, would require some degree of local knowledge to set appropriate values for model parameters, and

would also benefit from re-calibration of model tuning parameters using ABC estimation methods in order to achieve desired levels of transmission.

Regarding strengths and limitations of the CystiAgent model itself, our ABM is the first *T. solium* transmission model to incorporate a spatially explicit structure and include detailed parameters to represent key behavioral and environmental components of *T. solium* transmission. This spatial structure allowed CystiAgent to evaluate Ring Strategy, a promising spatially targeted control intervention, and provides the flexibility to test other structural and behavioral interventions in the future.

Despite these advantages, the CystiAgent model will always be limited by uncertainty in the mechanisms and dynamics of *T. solium* transmission, which could impact its accuracy and validity. Some of the more important gaps in knowledge that may impact transmission yet are not incorporated into CystiAgent include age-related differences in pig roaming and environmental exposure to *T. solium* [33,88], distribution patterns of infected pork through black market channels [154], and the possibility of vector-borne transmission of *T. solium* eggs via dung beetles and flies [86,87]. Perhaps the most important aspect of *T. solium* transmission not included in the model, however, is immunity. Due to insufficient knowledge of the mechanisms of immunity, resistance, and susceptibility to *T. solium* infection in humans and pigs, we were not able to incorporate these features into the model. As described above, this may have contributed to the observed overestimation of intervention effects. Since probabilities of infection in the model (i.e., tuning parameters) were estimated using levels of transmission observed at baseline, the model is not able to account for changes in susceptibility or resistance that may occur in response to control pressure when an intervention is applied. This feature could be added to the model if appropriate data from experimental and/or field studies were available. Until such studies are carried out, future versions of CystiAgent may improve accuracy by not fixing tuning parameters to baseline levels of transmission, but allowing them to flex during interventions, or estimating multiple values for these parameters over time in order to approximate the effect of

dynamic population immunity. While these and many other factors are not yet features of CystiAgent, a key advantage of its ABM structure is that it has the flexibility to incorporate these novel features when data become available, and can serve as an accessible platform to develop and test hypotheses about *T. solium* transmission dynamics.

5.6 Conclusion

In this research, we developed and validated a novel transmission model for *T. solium* called CystiAgent. In this first large-scale validation of a *T. solium* transmission model, CystiAgent was able to accurately predict levels of transmission observed at baseline in the validation villages, and adequately replicate the effects of control interventions in the majority of villages. Of course, model predictions were not flawless, and overall the model overestimated intervention effects in many of the villages tested. These imperfect results, however, represent important data-points that can be used to adjust and improve future versions of the model. Moving forward we will continue to test and improve the CystiAgent model using data available from interventions in Peru, and evaluate the generalizability of the model through validation against data from other endemic regions. Ultimately, we aim to use a final validated version of CystiAgent to evaluated available strategies for *T. solium* control and deliver evidence-based recommendations for *T. solium* control that will meet the need for validated strategies emphasized by WHO.

Chapter 6: Synthesis of Research

6.1 Overview

The primary objective of this dissertation research was to develop and validate a novel *T. solium* transmission model. The three research aims that constitute this dissertation were designed as a comprehensive and integrated effort to achieve this objective. In Aim 1 (Chapter 3), I conducted a suite of field studies in northern Peru that provided the necessary data to develop and parameterize a new agent-based model for *T. solium* transmission. Notably, these studies included investigation of the shape and size of pig roaming ranges and the practice of open outdoor defecation in the human population – both key features that dictate the spatial pattern of *T. solium* transmission in endemic areas. In Aim 2 (Chapter 4), I used data generated from my prior field studies along with other sources to develop the CystiAgent transmission model. I then subjected the model to a series of sensitivity analyses in order to evaluate the function of the model and better understand key sources of uncertainty. Finally, in Aim 3 (Chapter 5), I validated the model by simulating two large cluster-randomized trials of *T. solium* control strategies in Peru, and comparing the model-predicted results to those observed in the actual trials.

6.2 Significance and contributions of this research

6.2.1 Why do we need a transmission model for *T. solium*?

Despite the availability of tools to treat and prevent taeniasis and cysticercosis in endemic areas, the global burden of *T. solium* remains high: across Asia, Latin America, and Africa, more than 50 million people are infected with the parasite [165], and many more will become infected every day unless we take strong and decisive steps to improve the situation. There are a variety of

formidable barriers standing between our current status of widespread endemicity and the ultimate goal of global eradication of *T. solium* transmission, many of which are beyond the grasp of this dissertation research. Overcoming economic, political, and cultural barriers to be able to implement widespread control programs will be a remarkable challenge that will require collaboration and creativity from many sectors over the ensuing decades. Before we can even face this barrier, however, we must answer a set of more basic questions about how *T. solium* is transmitted and how to best prevent or control its spread. More simply, if we aim to convince future policy-makers, researchers, and the public to implement widespread control of *T. solium*, we must be able to give them a formula for how to do it and how much it will cost.

This goal of developing a "formula" of possible control options is a challenge that is well-suited to transmission modeling, and was the long-term motivation for embarking on this dissertation research. Given that relatively few prospective trials have been conducted that evaluate potential control or elimination strategies, we simply don't yet know the answers to critical questions about the effectiveness and cost of many of the interventions that are available and currently under consideration by global policy organizations like WHO and the Pan-American Health Organization (PAHO). Some of these fundamental questions include: "Which combinations of pig- and human-directed interventions will be most effective?", "How frequently and for how long should these interventions be applied?", "What is the best way to target treatment resources – MDA, screen-and-treat, or some other targeted methods (e.g., Ring Strategy, school-aged children, etc.)?", and, "Which of these strategies is most cost-effective and feasible to implement?" The answers to these questions are not expected to be uniform across endemic areas. Local variations in transmission, resources, and acceptability of interventions, as well as the objectives of a given program (e.g., elimination vs. control, time-frame, etc.) will mean that the most "effective" or optimal strategy will likely be different in different regions.

Given the breadth of questions regarding *T. solium* control that remain unanswered, an accurate and flexible transmission model could be a hugely impactful tool to guide future

decision-making. Because of the speed and economy of modeling, millions of simulations can be executed rapidly to test different combinations of interventions on different populations, and estimate the range of outcomes to evaluate their likelihood of achieving given objectives. These results would identify a set of promising strategies for a given region that could then be tested in prospective trials in order to confirm the impacts predicted by models. This addition of modeling as a powerful tool to aid policy decisions has been successful in efforts to control or eliminate other NTDs [128,166,167], and would be a major advancement for *T. solium* if an adequate model were available.

As it turns out, we are not the first group to recognize the potential impact of a transmission model for *T. solium.* In their 2012 "Roadmap for Implementation" [127] and 2014 "Framework for Intensified Control of *T. solium,"* [11], WHO identified transmission modeling as one of the key areas that requires improvement in order to make progress against the disease. In response to this call, three *T. solium* models were published by different groups between 2016-19 [14–16], and a systematic review outlining recent progress on *T. solium* modeling was just released [168]. One of these models, EPICYST, has been presented to WHO and was called upon to help select strategies for control programs that are slated to begin in five countries in 2020 [11]. With the addition of our own CystiAgent model, there is now significant momentum and a demonstrated will to move the *T. solium* modeling agenda forward.

6.2.2 What is unique about CystiAgent?

Since a variety of *T. solium* models already exist, and WHO is already collaborating with a prominent modeling group to make policy decisions about *T. solium* control, it would be fair to ask why we felt the need to develop our model, CystiAgent, and what CystiAgent contributes that has not already been achieved by prior models.

In Chapters 4 and 5, I wrote in depth about the many novel features that we incorporated into CystiAgent that are not available in existing *T. solium* models. These features include a unique agent-based structure with a spatial configuration, and new behavioral parameters such as

pig movement, human sanitation, and human travel. Most directly, the spatial structure of CystiAgent allows the model to test spatially targeted interventions like "Ring Strategy," something no other *T. solium* model can do. Additionally, we hypothesized that incorporating spatial and behavioral features of transmission would improve the accuracy of our model by accounting for heterogeneities that are not captured in other available models.

While we believe that these are important features that will improve the accuracy and flexibility of CystiAgent, the most fundamental difference between CystiAgent and other *T. solium* models lies in our approach, which centers on validation and data-driven improvements to the model. Fig 6.1 presents a simplified diagram of the approach we adopted for developing and improving CystiAgent. The most important feature to note is that the entire process is cyclical and continuous. In an ongoing process, data from field studies are incorporated into newer versions of the model, and each successive version is evaluated with sensitivity analysis and validation to assess its accuracy and identify gaps that can be addressed through additional field studies. Therefore, this dissertation research represents just the first lap through this sequence of field studies (Chapter 3), sensitivity analysis (Chapter 4) and validation (Chapter 5), and the cycle will continue iteratively as long as there are questions to address and inaccuracies to explore.

The advantage of this iterative approach to model development is that the model is continually improving, and validation of the model against prospective data serves as a checkpoint to understand its current level of accuracy before proceeding to more impactful policy analyses. In this way, validation of the model serves as both a benchmark that can be used to improve future versions of the model and as a source of credibility and context that can be used to interpret current outputs.

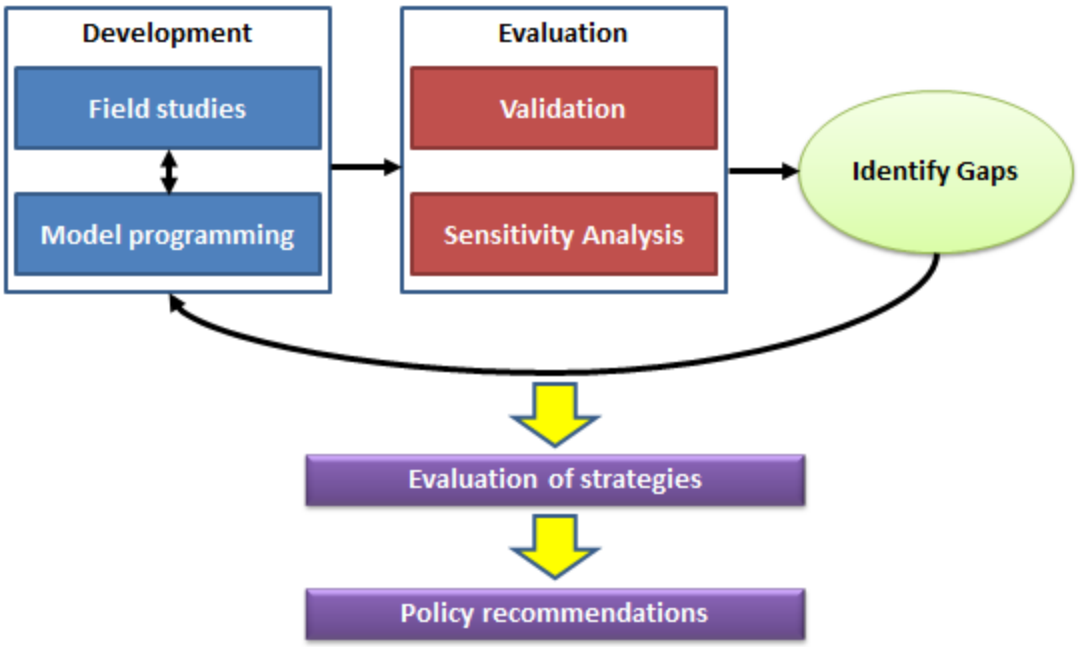


Fig 6.1. Conceptual flowchart of model development

Importantly, CystiAgent is the only existing *T. solium* model to have followed such a data-driven and iterative path. To our knowledge, no prior *T. solium* model has incorporated as much field data into its structure and parameterization, and no model has had its outputs validated with data from prospective intervention trials. Despite this lack of validation, each of the three recently released models have been used in published analyses to prospectively compare different combinations of interventions [14–16]. For CystiAgent, we plan to take a data-driven approach that ensures the validity of any future analyses. While our initial validation results were largely positive, there are important inaccuracies that remain, and we plan to move forward with analyses of control and elimination strategies only when we are confident that the model is accurately replicating *T. solium* transmission. This commitment highlights what we believe is an important contribution to prevention and control of *T. solium*: model validation against intervention trial data must be a pre-requisite for the use of any model in important policy decisions.

The final motivation for developing CystiAgent that I would like to highlight here involves promoting collaboration and data-sharing between modelers and field researchers. Given

the importance of model validation as described above, it is worth considering why prior *T. solium* models have not yet undergone this crucial process. The likely answer is that the steps of model development depicted in Fig 6.1 require an abundance of prospective data that few groups outside of Peru possess. By publishing results from the validation of CystiAgent, we hope to encourage increased data-sharing between the current (and future) research groups developing *T. solium* models. The availability of datasets published alongside our validation studies will allow for new and existing models to be validated against the same datasets. If this validation challenge is accepted by other groups, we would have an unprecedented opportunity to make unbiased comparisons between *T. solium* models and work synergistically to achieve the ultimate goal of supporting evidence-based decisions for *T. solium* control.

6.3 Future directions

In one sense, validation of the CystiAgent model represented a satisfying culmination of this dissertation research. We produced a good model that can be used to better understand *T. solium* transmission, and may ultimately be useful for conducting evaluations of control strategies if necessary improvements and adjustment can be made. However, as described above, the current version of CystiAgent represents only the first step of what we hope will be an enduring and cyclical process of validation and improvement to address the shortcomings of CystiAgent and other *T. solium* models.

Given the availability of other *T. solium* models, the first and most important follow-up to this research should be to replicate our rigorous validation of CystiAgent on these other models. While we designed CystiAgent to accurately represent components of *T. solium* transmission that we believed to be fundamental, there may be advantages of other model structures that we did not originally consider. In fact, it may be true that different model structures are better suited to different types of research questions or different endemic settings. For example, while the agent-

based structure of CystiAgent may allow it to test a wider variety of targeted strategies or structural/environmental interventions, a deterministic model may be adequate for testing mass-applied interventions while improving computational efficiency and the ability to explore other mathematical relationships (e.g., basic reproductive rate, incidence). Only through comparative validation of multiple *T. solium* models can we begin to understand the contribution and validity of each model, and to capitalize on the synergy that is possible through such collaborations.

Until such cross-model validations can be completed, however, we have only the results of our current work on CystiAgent to direct research priorities for future model improvements. For this purpose, the sensitivity analyses we conducted (Chapter 4) were essential in identifying parameters that had significant impacts on transmission in the model. Targeting these key components of transmission for additional research would not only improve the accuracy and precision CystiAgent, but would help uncover unknown mechanisms of *T. solium* transmission that could have broad impacts on control. In the following section, I briefly describe a few of the most important features of transmission that could be targeted with additional field work.

First, while our GPS tracking study provided important estimates for the size of pig roaming ranges and areas of contact with human feces, more research is needed to determine if pig behavior and human defecation patterns are similar in other endemic areas. The area of pig home-ranges and the prevalence of latrine-use were both identified as a highly impactful parameters and significant sources of uncertainty in our sensitivity analyses. Thus, replication of this study in other endemic regions would allow us to test the model in these regions, and may ultimately uncover information about differences in transmission patterns that could impact decisions about prevention and control.

Another important and largely unknown factor of *T. solium* transmission that was identified in our sensitivity analyses involves the distribution of infected pork throughout and between villages. Our surveys on this topic only crudely assessed the sale and distribution of pork, but did not specifically ask about other features of pork distribution including black market

sale of infected pork [154], donation to family members, traditional methods of curing or boiling meat to sanitize it [169], or report of infected meat to health officials [123]. What is particularly challenging about this topic is that such practices are difficult to assess given the stigma surrounding taeniasis and tainted pork, and may vary considerably between different cultures in endemic regions. Therefore, replication of these surveys in other endemic regions, and consideration of alternative research methods such as focus groups and personal interviews will be needed to fully explore this topic.

The average duration of tapeworm infections is another key aspect of transmission that was identified as highly influential on transmission in CystiAgent, but is subject to a high degree of uncertainty. Given the lack of animal models for *T. solium* infection and ethical barrier of monitoring active human infections, the natural duration of the tapeworm infections cannot be assessed directly. If the natural incidence of tapeworm infections in endemic population were known, duration could be estimated indirectly (Duration = Prevalence / Incidence). Unfortunately, the incidence of taeniasis has not yet been measured in observational epidemiologic studies, and would require substantial investment to reach necessary sample sizes to detect incident cases. Therefore, future work to reduce uncertainty in this key model parameter will likely require creative employment of modeling techniques. This could include model-based calibration of the parameter value to arrive at an estimate that reduces measured error in model outcomes.

Finally, the role of pig immunity in developing resistance to infection and causing antibody response without true cyst infection [46] is a field of research with considerable uncertainty that likely impacted the accuracy of CystiAgent in our validation study. Pig resistance and immunity were not factored into CystiAgent in any way due to limited knowledge on the topic. Experimental infection studies to untangle the complicated relationship between exposure to eggs, immune activation, and resistance would provide important knowledge to improve future versions of the model.

For each of the above components of transmission, a lack of quantitative data has potentially limited our correct specification of the feature in CystiAgent, and in each case, this uncertainty has potentially impacted variability of the model's outputs. With this in mind, our research group has proposed a suite of new epidemiologic and experimental studies aimed addressing these gaps. The proposed field work includes an enhanced GPS tracking study of pig movement, rigorous experimental infection studies to better understand resistance and immunity in pigs, and the validation of CystiAgent with prospective data collected in sub-Saharan Africa (Zambia) to evaluate its generalizability to this substantially different ecological setting.

When this research is completed, the resulting transmission model should have the ability to conduct the types of robust evaluations and comparisons of *T. solium* control strategies that are urgently needed for policy decisions. A final evaluation of control strategies conducted with this validated and generalizable model would include comparison of different combinations of interventions, including timing and frequency of treatment, vaccination, and behavioral interventions, to determine optimal sets of actions to maximize intervention effectiveness . These strategies could be tested in diverse settings and populations, and would be expected to include measurement of cost to assess feasibility. An evaluation of this magnitude would change the landscape of knowledge and tools available for *T. solium* control, and is now within reach because of the important foundation built through this dissertation research.

References

1. Coyle CM, Mahanty S, Zunt JR, Wallin MT, Cantey PT, White AC, et al. Neurocysticercosis: neglected but not forgotten. PLoS Negl Trop Dis. 2012 Jan;6(5):e1500. PMID: 22666505
2. Winkler AS. Neurocysticercosis in sub-Saharan Africa: a review of prevalence, clinical characteristics, diagnosis, and management. Pathog Glob Health. 2012;106(5):261–74. PMID: 23265550
3. Ndimubanzi PC, Carabin H, Budke CM, Nguyen H, Qian Y-J, Rainwater E, et al. A systematic review of the frequency of neurocyticercosis with a focus on people with epilepsy. PLoS Negl Trop Dis. 2010 Jan;4(11):e870. PMID: 21072231
4. Moyano LM, Saito M, Montano SM, Gonzalvez G, Olaya S, Ayvar V, et al. Neurocysticercosis as a cause of epilepsy and seizures in two community-based studies in a cysticercosis-endemic region in Peru. PLoS Negl Trop Dis. 2014;8(2). PMID: 24551255
5. Montano SM, Villaran M V., Ylquimiche L, Figueroa JJ, Rodriguez S, Bautista CT, et al. Neurocysticercosis: Association between seizures, serology, and brain CT in rural Peru. Neurology. 2005;65(2 SUPPL. 1):229–34. PMID: 16043791
6. Bruno E, Bartoloni A, Zammarchi L, Strohmeyer M, Bartalesi F, Bustos JA, et al. Epilepsy and neurocysticercosis in Latin America: A systematic review and meta-analysis. PLoS Negl Trop Dis. 2013;7(10). PMID: 24205415
7. Del Brutto OH, Santibáñez R, Idrovo L, Rodríguez S, Díaz-Calderón E, Navas C, et al. Epilepsy and neurocysticercosis in Atahualpa: A door-to-door survey in rural coastal Ecuador. Epilepsia. 2005;46(4):583–7. PMID: 15816956
8. Schantz PM, Cruz M, Sarti E, Pawlowski Z. Potential eradicability of taeniasis and cysticercosis. Bull Pan Am Health Organ. 1993;27(4):397–403. PMID: 8312963
9. Rajkotia Y, Lescano AG, Gilman RH, Cornejo C, Garcia HH. Economic burden of neurocysticercosis: results from Peru. Trans R Soc Trop Med Hyg. 2007;101(8):840–6. PMID: 17507067
10. Torgerson PR, Devleesschauwer B, Praet N, Speybroeck N, Willingham AL, Kasuga F, et al. World Health Organization Estimates of the global and regional disease burden of 11 foodborne parasitic diseases, 2010: A data synthesis. PLoS Med. 2015;12(12):1–22. PMID: 26633705
11. World Health Organization (WHO). Assembling a framework for intensified control of taeniasis and neurocysticercosis caused by Taenia solium: Report of an informal consultation. 2015;(July):17–8.
12. Gonzalez AE, Gilman RH, García HH, Lopez T. Use of a simulation model to evaluate control programmes against Taenia solium cysticercosis. In: Singh G, Prabhakar S, editors. Taenia solium Cysticercosis: From Basic to Clinical Science. CABI; 2002. p. 437–48.
13. Kyvsgaard NC, Johansen MV, Carabin H. Simulating transmission and control of Taenia solium infections using a Reed-Frost stochastic model. Int J Parasitol. 2007;37(5):547–58. PMID: 17250839
14. Braae UC, Devleesschauwer B, Gabriël S, Dorny P, Speybroeck N, Magnussen P, et al. CystiSim – an agent-based model for Taenia solium transmission and control. PLoS Negl Trop Dis. 2016;10(12):e0005184. PMID: 27984581

15. Winskill P, Harrison WE, French MD, Dixon MA, Abela-Ridder B, Basáñez M-G. Assessing the impact of intervention strategies against Taenia solium cysticercosis using the EPICYST transmission model. Parasit Vectors. 2017 Feb 9;10(1):73. PMID: 28183336

16. Sánchez-Torres NY, Bobadilla JR, Laclette JP, José M V. How to eliminate taeniasis/cysticercosis: porcine vaccination and human chemotherapy (Part 2). Theor Biol Med Model. 2019 Feb 26;16(1):4. PMID: 30803437

17. Nunn JF. Ancient Egyptian Medicine. Norman: University of Oklahoma Press; 2002. 71 p.

18. Cox FE. History of Human Parasitology. Clin Microbiol Rev. 2002 Oct 1;15(4):595–612. PMID: 12270904

19. Grove D. Tapeworms, Lice, and Prions: A compendium of unpleasant infections. Oxford: Oxford University Press; 2014. 23-36 p.

20. Küchenmeister F. Experimental evidence that Cysticercus cellulosae is transformed into Taenia solium in the human intestines. Weiner medizinische Wochenschrift. 1855;5:1–4.

21. Küchenmeister F. On animal and vegetable parasites of the human body. London: Then Syndenham Society; 1855. 452 p.

22. García HH, Gonzalez AE, Evans CAW, Gilman RH, Working C. Taenia solium cysticercosis. Lancet. 2003;362(9383):547–56. PMID: 12932389

23. Del Brutto OH, García HH. Taenia solium cysticercosis - The lessons of history. J Neurol Sci. 2015;359(1–2):392–5. PMID: 26320098

24. MacArthur WP. Cysticercosis as seen in the British Army, with special reference to the production of epilepsy. Trans R Soc Trop Med Hyg. 1934;27(4).

25. MacArthur WP. Cysticercosis of the brain. Br Med J. 1935;2:1229.

26. Pawlowsky Z. Taenia solium: basic biology and transmission. In: Singh G, Prabhakar S, editors. Taenia solium Cysticercosis: From Basic to Clinical Science. CABI; 2002. p. 1–14.

27. Yoshino K. On the subjective symptoms caused by parasitism of Taenia solium and its development in man (English summary). J Med Assoc Formosa. 1934;33:183–94.

28. Lightowlers MW. Control of Taenia solium taeniasis/cysticercosis: past practices and new possibilities. Parasitology. 2013;140(13):1566–77. PMID: 23947762

29. Pawlowski Z, Schultz MG. Taeniasis and cysticercosis (Taenia saginata). Adv Parasitol. 1972;10:269–343. PMID: 4559145

30. García HH, Gilman RH, Gonzalez AE, Verastegui M, Rodriguez S, Gavidia C, et al. Hyperendemic human and porcine Taenia solium infection in Perú. Am J Trop Med Hyg. 2003;68(3):268–75. PMID: 12685628

31. Allan JC, Velasquez-Tohom M, Garcia-Noval J, Torres-Alvarez R, Yurrita P, Fletes C, et al. Epidemiology of intestinal taeniasis in four, rural, Guatemalan communities. Ann Trop Med Parasitol. 1996 Apr;90(2):157–65. PMID: 8762405

32. Flisser A. Taeniasis and cysticercosis due to Taenia solium. Prog Clin Parasitol. 1994 Jan;4:77–116. PMID: 7948938

33. Copado F, De Aluja AS, Mayagoitia L, Galindo F. The behaviour of free ranging pigs in the Mexican tropics and its relationships with human faeces consumption. Appl Anim Behav Sci. 2004;88(3–4):243–52.

34. Sikasunge CS, Johansen MV, Willingham AL, Leifsson PS, Phiri IK. Taenia solium porcine cysticercosis: Viability of cysticerci and persistency of antibodies and cysticercal antigens after treatment with oxfendazole. Vet Parasitol. 2008;158(1–2):57–66. PMID: 18834668

35. Singh AK, Singh SK, Prasad KN, Singh A, Bajpai A, Rahman M, et al. Evaluation of ELISA, neck

muscle, tongue and eyelid examinations for the diagnosis of swine cysticercosis in a highly endemic area of north India. Exp Parasitol. 2013;134(3):313–7. PMID: 23578857

36. Lightowlers MWW, Assana E, Jayashi CM, Gauci CGG, Donadeu M. Sensitivity of partial carcass dissection for assessment of porcine cysticercosis at necropsy. Int J Parasitol. 2015;45(13):815–8.

37. Garcia HH, Gonzalez AE, Gilman RH, Cysticerosis Working Group in Peru. Diagnosis, treatment and control of Taenia solium cysticercosis. Curr Opin Infect Dis. 2003 Oct;16(5):411–9. PMID: 14501993

38. Gweba M, Faleke OO, Junaidu AU. Some risk factors for Taenia solium cysticercosis in semi-intensively raised pigs in Zuru , Nigeria. 2010;46(1):57–67.

39. Ngowi HA, Kassuku AA, Maeda GEM, Boa ME, Carabin H, Willingham AL. Risk factors for the prevalence of porcine cysticercosis in Mbulu District, Tanzania. Vet Parasitol. 2004;120(4):275–83. PMID: 15063938

40. Pouedet MSR, Zoli AP, Nguekam, Vondou L, Assana E, Speybroeck N, et al. Epidemiological survey of swine cysticercosis in two rural communities of West-Cameroon. Vet Parasitol. 2002;106(1):45–54. PMID: 11992710

41. De Aluja AS, Martinez JJ, Villalobos NM. Taenia solium cysticercosis in young pigs: Age at first infection and histological characteristics. Vet Parasitol. 1998;76(1–2):71–9. PMID: 9653992

42. Phiri IK, Dorny P, Gabriel S, Willingham AL, Sikasunge C, Siziya S, et al. Assessment of routine inspection methods for porcine cysticercosis in Zambian village pigs. J Helminthol. 2006;80(1):69–72. PMID: 16469176

43. Flecker RH, Pray IW, Santivañez SJ, Ayvar V, Gamboa R, Muro C, et al. Assessing ultrasonography as a diagnostic tool for porcine cysticercosis. PLoS Negl Trop Dis. 2017;11(1):e0005282.

44. Gonzalez AE, Cama V, Oilman RH, Tsang VCW, Pilcher JB, Chavera A, et al. Prevalence and comparison of serological assays, necroscopy, and tongue examination for the diagnosis of porcince cysticercosis in Peru. Am J Trop Med Hyg. 1990;43(2):194–9.

45. Tsang VC, Brand JA, Boyer AE. An enzyme-linked immunoelectrotransfer blot assay and glycoprotein antigens for diagnosing human cysticercosis (Taenia solium). J Infect Dis. 1989 Jan;159(1):50–9. PMID: 2909643

46. Jayashi CM, Gonzalez AE, Castillo Neyra R, Rodríguez S, García HH, Lightowlers MW. Validity of the Enzyme-linked Immunoelectrotransfer Blot (EITB) for naturally acquired porcine cysticercosis. Vet Parasitol. 2014;199(1–2):42–9. PMID: 24183647

47. Sikasunge CS, Phiri IK, Willingham AL, Johansen M V. Dynamics and longevity of maternally-acquired antibodies to Taenia solium in piglets born to naturally infected sows. Vet J. 2010;184(3):318–21. PMID: 19380243

48. Allan JC, Velasquez-Tohom M, Fletes C, Torres-Alvarez R, Lopez-Virula G, Yurrita P, et al. Mass chemotherapy for intestinal Taenia solium infection: Effect on prevalence in humans and pigs. Trans R Soc Trop Med Hyg. 1997;91(5):595–8. PMID: 9463679

49. Garcia HH, Gonzalez AE, Gilman RH, Moulton LH, Verastegui M, Rodriguez S, et al. Combined human and porcine mass chemotherapy for the control of T. solium. Am J Trop Med Hyg. 2006 May;74(5):850–5. PMID: 16687692

50. O'Neal SE, Moyano LM, Ayvar V, Rodriguez S, Gavidia C, Wilkins PP, et al. Ring-screening to control endemic transmission of Taenia solium. PLoS Negl Trop Dis. 2014 Sep;8(9):e3125. PMID: 25210748

51. Carabin H, Winkler AS, Dorny P. Taenia solium cysticercosis and taeniosis: Achievements from the past 10 years and the way forward. PLoS Negl Trop Dis. 2017 Apr;11(4):e0005478. PMID: 28426664

52. Del Brutto OH, Santibañez R, Noboa CA, Aguirre R, Díaz E, Alarcón TA. Epilepsy due to neurocysticercosis: analysis of 203 patients. Neurology. 1992 Feb;42(2):389–92. PMID: 1736171

53. Donadeu M, Fahrion AS, Olliaro PL, Abela-Ridder B. Target product profiles for the diagnosis of Taenia solium taeniasis, neurocysticercosis and porcine cysticercosis. PLoS Negl Trop Dis. 2017 Sep;11(9):e0005875. PMID: 28892472

54. Wilson M, Bryan RT, Fried JA, Ware DA, Schantz PM, Pilcher JB, et al. Clinical evaluation of the cysticercosis enzyme-linked immunoelectrotransfer blot in patients with neurocysticercosis. J Infect Dis. 1991 Nov;164(5):1007–9. PMID: 1940452

55. Garcia HH, Del Brutto OH. Neurocysticercosis: Updated concepts about an old disease. Lancet Neurol. 2005;4(10):653–61. PMID: 16168934

56. Moyano LM, O 'neal SE, Ayvar V, Gonzalvez G, Gamboa R, Vilchez P, et al. High prevalence of asymptomatic neurocysticercosis in an endemic rural community in Peru. 2016;142:1–11.

57. Medina MT, Durón RM, Martínez L, Osorio JR, Estrada AL, Zúniga C, et al. Prevalence, incidence, and etiology of epilepsies in rural Honduras: The Salamá study. Epilepsia. 2005;46(1):124–31. PMID: 15660778

58. Garcia-Noval J, Moreno E, de Mata F, Soto de Alfaro H, Fletes C, Craig PS, et al. An epidemiological study of epilepsy and epileptic seizures in two rural Guatemalan communities. Ann Trop Med Parasitol. 2001 Mar;95(2):167–75. PMID: 11299123

59. Sarti E, Schantz PM, Plancarte A, Wilson M, Gutierrez O I, Aguilera J, et al. Epidemiological investigation of taenia solium taeniasis and cysticercosis in a rural village of Michoacan state, Mexico. Trans R Soc Trop Med Hyg. 1994;88(1):49–52. PMID: 8154000

60. Sarti E, Schantz PM, Plancarte A, Wilson M, Gutierrez IO, Lopez AS, et al. Prevalence and risk factors for Taenia solium taeniasis and cysticercosis in humans and pigs in a village in Morelos, Mexico. Am J Trop Med Hyg. 1992;46(6):677–85. PMID: 1621892

61. Schantz PM, Sarti E, Plancarte A, Wilson M, Criales JL, Roberts J, et al. Community-based epidemiological investigations of cysticercosis due to Taenia solium: comparison of serological screening tests and clinical findings in two populations in Mexico. Clin Infect Dis. 1994 Jun;18(6):879–85. PMID: 8086547

62. Coral-Almeida M, Gabriël S, Abatih EN, Praet N, Benitez W, Dorny P. Taenia solium human cysticercosis: A systematic review of sero-epidemological data from endemic zones around the world. PLoS Negl Trop Dis. 2015;9(7):1–20. PMID: 26147942

63. World Health Organization (WHO). Endemicity of Taenia solium. 2015 (available from http://www.who.int/mediacentre/factsheets/fs376/en/).

64. O'Neal SE, Flecker RH. Hospitalization frequency and charges for neurocysticercosis, United States, 2003-2012. Emerg Infect Dis. 2015 Jun;21(6):969–76. PMID: 25988221

65. Ong S, Talan DA, Moran GJ, Mower W, Newdow M, Tsang VCW, et al. Neurocysticercosis in radiographically imaged seizure patients in U.S. emergency departments. Emerg Infect Dis. 2002;8(6):608–13. PMID: 12023918

66. Sánchez AL, Medina MT, Ljungström I. Prevalence of taeniasis and cysticercosis in a population of urban residence in Honduras. Acta Trop. 1998;69(2):141–9. PMID: 9588234

67. Secka A, Grimm F, Marcotty T, Geysen D, Niang AM, Ngale V, et al. Old focus of cysticercosis in a senegalese village revisited after half a century. Acta Trop. 2011;119(2–3):199–202. PMID: 21605539

68. Agudelo Florez P, Palacio LG. [Prevalence of Taenia solium antibodies in humans and in pigs in an endemic area of Colombia]. Rev Neurol. 36(8):706–9. PMID: 12717645

69. Mwanjali G, Kihamia C, Kakoko DVC, Lekule F, Ngowi H, Johansen MV, et al. Prevalence and risk factors associated with human Taenia solium infections in Mbozi District, Mbeya Region,

Tanzania. PLoS Negl Trop Dis. 2013;7(3). PMID: 23516650

70. Lescano AG, Garcia HH, Gilman RH, Gavidia CM, Tsang VCW, Rodriguez S, et al. Taenia solium cysticercosis hotspots surrounding tapeworm carriers: Clustering on human seroprevalence but not on seizures. PLoS Negl Trop Dis. 2009;3(1). PMID: 19172178

71. Camacho SPD, Ruiz AC, Peraza VS, Ramos MLZ, Medina MF, Lozano R, et al. Epidemiologic study and control of Taenia solium infections with praziquantel in a rural village of Mexico. Am J Trop Med Hyg. 1991;45(4):522–31. PMID: 1951862

72. Allan JC, Velasquez-Tohom M, Torres-Alvarez R, Yurrita P, Garcia-Noval J. Field trial of the coproantigen-based diagnosis of Taenia solium taeniasis by enzyme-linked immunosorbent assay. Am J Trop Med Hyg. 1996;54(4):352–6. PMID: 8615446

73. Rodriguez-Hidalgo R, Benitez-Ortiz W, Praet N, Saa LR, Vercruysse J, Brandt J, et al. Taeniasis-cysticercosis in Southern Ecuador: Assessment of infection status using multiple laboratory diagnostic tools. Mem Inst Oswaldo Cruz. 2006;101(7):779–82.

74. O'Neal SE, Moyano LM, Ayvar V, Gonzalvez G, Diaz A, Rodriguez S, et al. Geographic correlation between tapeworm carriers and heavily infected cysticercotic pigs. PLoS Negl Trop Dis. 2012 Jan;6(12):e1953. PMID: 23285305

75. Pray IW, Ayvar V, Gamboa R, Muro C, Moyano LMLM, Benavides V, et al. Spatial relationship between Taenia solium tapeworm carriers and necropsy cyst burden in pigs. PLoS Negl Trop Dis. 2017 Apr;11(4):e0005536. PMID: 28406898

76. Mwape KE, Phiri IK, Praet N, Muma JB, Zulu G, van den Bossche P, et al. Taenia solium infections in a rural area of Eastern Zambia-A community based study. PLoS Negl Trop Dis. 2012;6(3):1–9. PMID: 22479664

77. Prasad KN, Prasad A, Gupta RK, Pandey CM, Singh U. Prevalence and associated risk factors of Taenia solium taeniasis in a rural pig farming community of north India. Trans R Soc Trop Med Hyg. 2007;101(12):1241–7. PMID: 17603090

78. De N V, Le TH, Lien PTH, Eom KS. Current status of taeniasis and cysticercosis in Vietnam. Korean J Parasitol. 2014;52(2):125–9. PMID: 24850954

79. Raghava MV, Prabhakaran V, Jayaraman T, Muliyil J, Oommen A, Dorny P, et al. Detecting spatial clusters of Taenia solium infections in a rural block in South India. Trans R Soc Trop Med Hyg. 2010;104(9):601–12. PMID: 20638091

80. Bardosh K, Inthavong P, Xayaheuang S, Okello AL. Controlling parasites, understanding practices: The biosocial complexity of a One Health intervention for neglected zoonotic helminths in northern Lao PDR. Soc Sci Med. 2014;120C:215–23. PMID: 25261615

81. Willingham AL, Wu HW, Conlan J, Satrija F. Combating Taenia solium cysticercosis in Southeast Asia: an opportunity for improving human health and livestock production. Adv Parasitol. 2010;72(C):235–66. PMID: 20624534

82. Garvey B. Increasing porcine cysticercosis case reporting via cash transfer and microinsurance. In: Oregon Health and Science University: School of Publich Health Grand Rounds. Portland, Ore; 2017.

83. Gonzalez AE, Gilman R, Garcia HH, McDonald J, Kacena K, Tsang VCW, et al. Use of sentinel pigs to monitor environmental Taenia solium contamination. Am J Trop Med Hyg. 1994;51(6):847–50. PMID: 7810821

84. Sikasunge CS, Phiri IK, Phiri AM, Dorny P, Siziya S, Willingham AL. Risk factors associated with porcine cysticercosis in selected districts of Eastern and Southern provinces of Zambia. Vet Parasitol. 2007;143(1):59–66. PMID: 16956727

85. Pondja A, Neves L, Mlangwa J, Afonso S, Fafetine J, Willingham AL, et al. Prevalence and risk factors of porcine cysticercosis in Angónia District, Mozambique. PLoS Negl Trop Dis. 2010;4(2). PMID: 20126403

86. Lawson JR, Gemmell MA. Transmission of taeniid tapeworm eggs via blowflies to intermediate hosts. Parasitology. 1990;100 Pt 1:143–6. PMID: 2314928

87. Gomez-Puerta LA, Lopez-Urbina MT, Garcia HH, Gonzalez AE. Longevity and viability of Taenia solium eggs in the digestive system of the beetle Ammophorus rubripes. Rev Bras Parasitol veterinária. 2014 Mar;23(1):94–7. PMID: 24728368

88. Pray IW, Swanson DJ, Ayvar V, Muro C, Moyano LM, Gonzalez AE, et al. GPS tracking of free-ranging pigs to evaluate ring strategies for the control of cysticercosis/taeniasis in Peru. PLoS Negl Trop Dis. 2016 Apr;10(4):e0004591. PMID: 27035825

89. Thomas LF, de Glanville WA, Cook EA, Fèvre EM. The spatial ecology of free-ranging domestic pigs (Sus scrofa) in western Kenya. BMC Vet Res. 2013 Jan;9:46. PMID: 23497587

90. Lescano AG, García HH, Gilman RH, Guezala MC, Tsang VCW, Gavidia CM, et al. Swine cysticercosis hotspots surrounding Taenia solium tapeworm carriers. Am J Trop Med Hyg. 2007;76(2):376–83. PMID: 17297051

91. Sarti E, Schantz PM, Avila G, Ambrosio J, Medina-Santillán R, Flisser A. Mass treatment against human taeniasis for the control of cysticercosis: a population-based intervention study. Trans R Soc Trop Med Hyg. 2000;94(1):85–9. PMID: 10748908

92. Cruz M, Davis A, Dixon H, Pawlowski ZS, Proano J. Operational studies on the control of Taenia solium taeniasis/cysticercosis in Ecuador. Bull World Health Organ. 1989;67(4):401–7. PMID: 2805217

93. Bustos JA, Rodriguez S, Jimenez JA, Moyano LM, Castillo Y, Ayvar V, et al. Detection of Taenia solium taeniasis coproantigen is an early indicator of treatment failure for taeniasis. Clin Vaccine Immunol. 2012 Apr;19(4):570–3. PMID: 22336287

94. Garcia HH, Gonzalez AE, Tsang VCW, O'Neal SE, Llanos-Zavalaga F, Gonzalvez G, et al. Elimination of Taenia solium transmission in northern Peru. N Engl J Med. 2016 Jun 16;374(24):2335–44. PMID: 27305193

95. Allan JC, Wilkins PP, Tsang VCW, Craig PS. Immunodiagnostic tools for taeniasis. Acta Trop. 2003;87(1):87–93. PMID: 12781382

96. Gonzalez AE, Falcon N, Gavidia C, Garcia HH, Tsang VCW, Bernal T, et al. Time-response curve of oxfendazole in the treatment of swine cysticercosis. Am J Trop Med Hyg. 1998;59(5):832–6. PMID: 9840607

97. Gonzalez AE, Gavidia C, Falcon N, Bernal T, Verastegui M, Garcia HH, et al. Protection of pigs with cysticercosis from further infections after treatment with oxfendazole. Am J Trop Med Hyg. 2001;65(1):15–8. PMID: 11504400

98. Pondja A, Neves L, Mlangwa J, Afonso S, Fafetine J, Willingham AL, et al. Use of oxfendazole to control porcine cysticercosis in a high-endemic area of Mozambique. PLoS Negl Trop Dis. 2012;6(5):1–5. PMID: 22666509

99. Jayashi CM, Kyngdon CT, Gauci CG, Gonzalez AE, Lightowlers MW. Successful immunization of naturally reared pigs against porcine cysticercosis with a recombinant oncosphere antigen vaccine. Vet Parasitol. 2012;188(3–4):261–7. PMID: 22541797

100. Assana E, Kyngdon CT, Gauci CG, Geerts S, Dorny P, De Deken R, et al. Elimination of Taenia solium transmission to pigs in a field trial of the TSOL18 vaccine in Cameroon. Int J Parasitol. 2010;40(5):515–9. PMID: 20138046

101. Okello AL, Thomas L, Inthavong P, Ash A, Khamlome B, Keokamphet C, et al. Assessing the impact of a joint human-porcine intervention package for Taenia solium control: Results of a pilot study from northern Lao PDR. Acta Trop. 2016;159:185–91. PMID: 26992295

102. Braae UC, Magnussen P, Ndawi B, Harrison W, Lekule F, Johansen MV. Effect of repeated mass drug administration with praziquantel and track and treat of taeniosis cases on the prevalence of taeniosis in Taenia solium endemic rural communities of Tanzania. Acta Trop. 2017;165:246–51.

PMID: 26597324

103. Wohlgemut J, Dewey C, Levy M, Mutua F. Evaluating the efficacy of teaching methods regarding prevention of human epilepsy caused by Taenia solium neurocysticercosis in Western Kenya. Am J Trop Med Hyg. 2010;82(4):634–42. PMID: 20348512

104. Mwidunda SA, Carabin H, Matuja WBM, Winkler AS, Ngowi HA. A school based cluster randomised health education intervention trial for improving knowledge and attitudes related to Taenia solium cysticercosis and taeniasis in Mbulu district, northern Tanzania. PLoS One. 2015;10(2):e0118541. PMID: 25719902

105. Ngowi HA, Carabin H, Kassuku AA, Mlozi MRS, Mlangwa JED, Willingham AL. A health-education intervention trial to reduce porcine cysticercosis in Mbulu District, Tanzania. Prev Vet Med. 2008;85(1–2):52–67. PMID: 18243375

106. Sarti E, Flisser A, Schantz PM, Gleizer M, Loya M, Plancarte A, et al. Development and evaluation of a health education intervention against Taenia solium in a rural community in Mexico. Am J Trop Med Hyg. 1997;56(2):127–32. PMID: 9080868

107. Bulaya C, Mwape KE, Michelo C, Sikasunge CS, Makungu C, Gabriel S, et al. Preliminary evaluation of Community-Led Total Sanitation for the control of Taenia solium cysticercosis in Katete District of Zambia. Vet Parasitol. 2015 Jan 30;207(3–4):241–8. PMID: 25591408

108. Thys S, Mwape KE, Lefèvre P, Dorny P, Marcotty T, Phiri AM, et al. Why latrines are not used: communities' perceptions and practices regarding latrines in a Taenia solium endemic rural area in Eastern Zambia. PLoS Negl Trop Dis. 2015 Mar;9(3):e0003570. PMID: 25739017

109. Widdowson M-A, Cook AJC, Williams JJ, Argaes F, Rodriguez I, Dominguez JL, et al. Investigation of risk factors for porcine Taenia solium cysticercosis: a multiple regression analysis of a cross-sectional study in the Yucatan Peninsula, Mexico. Trans R Soc Trop Med Hyg. 2000;94(6):620–4. PMID: 11198643

110. Morales J, Velasco T, Tovar V, Fragoso G, Fleury A, Beltrán C, et al. Castration and pregnancy of rural pigs significantly increase the prevalence of naturally acquired Taenia solium cysticercosis. Vet Parasitol. 2002;108(1):41–8. PMID: 12191898

111. Hollingsworth TD, Adams ER, Anderson RM, Atkins K, Bartsch S, Basáñez M-G, et al. Quantitative analyses and modelling to support achievement of the 2020 goals for nine neglected tropical diseases. Parasit Vectors. 2015;8(1):630. PMID: 26652272

112. Heesterbeek H, Anderson RM, Andreasen V, Bansal S, De Angelis D, Dye C, et al. Modeling infectious disease dynamics in the complex landscape of global health. Science. 2015 Mar 13;347(6227). PMID: 25766240

113. Johansen MV, Trevisan C, Gabriël S, Magnussen P, Braae UC. Are we ready for Taenia solium cysticercosis elimination in sub-Saharan Africa? Parasitology. 2016;1–6. PMID: 27094170

114. Basáñez M-G, McCarthy JS, French MD, Yang G-J, Walker M, Gambhir M, et al. A research agenda for helminth diseases of humans: modelling for control and elimination. PLoS Negl Trop Dis. 2012;6(4):e1548. PMID: 22545162

115. Vynnycky E, White RG. An Introduction to Infectious Disease Modeling. New York: Oxford University Press; 2010.

116. Ngowi HA, Kassuku AA, Carabin H, Mlangwa JED, Mlozi MRS, Mbilinyi BP, et al. Spatial clustering of porcine cysticercosis in Mbulu district, northern Tanzania. PLoS Negl Trop Dis. 2010;4(4). PMID: 20386601

117. Burr TL, Chowell G. Signatures of non-homogeneous mixing in disease outbreaks. Math Comput Model. 2008;48(1–2):122–40.

118. Roberts MG, Heesterbeek JAP. A new method for estimating the effort required to control an infectious disease. Proc R Soc London Ser B-Biological Sci. 2003;270(1522):1359–64. PMID: 12965026

119. Gilman RH, Gonzalez AE, Llanos-Zavalaga F, Tsang VCW, Garcia HH. Prevention and control of Taenia solium taeniasis/cysticercosis in Peru. Pathog Glob Health. 2012 Sep;106(5):312–8. PMID: 23265557

120. Gonzales I, Miranda JJ, Rodriguez S, Vargas V, Cjuno A, Smeeth L, et al. Seizures, cysticercosis and rural-to-urban migration: the PERU MIGRANT study. Trop Med Int Health. 2015;20(4):546–52. PMID: 25581851

121. Klepac P, Metcalf CJE, McLean AR, Hampson K. Towards the endgame and beyond: complexities and challenges for the elimination of infectious diseases. Philos Trans R Soc Lond B Biol Sci. 2013;368(1623):20120137. PMID: 23798686

122. Lescano AG, Pray IW, Gonzalez AE, Gilman RH, Tsang VCW, Gamboa R, et al. Clustering of necropsy-confirmed porcine cysticercosis surrounding Taenia solium tapeworm carriers in Peru. 2019;100(2):314–22.

123. Beam M, Spencer A, Fernandez L, Atto R, Muro C, Vilchez P, et al. Barriers to participation in a community-based program to control transmission of Taenia solium in Peru. Am J Trop Med Hyg. 2018 Jun;98(6):1748–54. PMID: 29663901

124. Getz WM, Wilmers CC. A local nearest-neighbor convex-hull construction of home ranges and utilization distributions. Ecography (Cop). 2004;27(4):489–505.

125. Getz WM, Fortmann-Roe S, Cross PC, Lyons AJ, Ryan SJ, Wilmers CC. LoCoH: Nonparameteric Kernel methods for constructing home ranges and utilization distributions. PLoS One. 2007;2(2). PMID: 17299587

126. Lyons A. T-LoCoH for R: Tutorial and User Guide. http://tlocoh.r-forge.r-project.org/tlocoh_tutorial_2014-08-17.pdf. 2014 (available from http://tlocoh.r-forge.r-project.org/tlocoh_tutorial_2014-08-17.pdf).

127. Savioli L, Daumerie D. Accelerating Work to Overcome the Global Impact of Neglected Tropical Diseases: A Roadmap for Implementation. World Heal Organ. 2012;1–42.

128. Winnen M, Plaisier AP, Alley ES, Nagelkerke NJD, Van Oortmarssen G, Boatin BA, et al. Can ivermectin mass treatments eliminate onchocerciasis in Africa? Bull World Health Organ. 2002;80(5):384–90. PMID: 12077614

129. Hladish TJ, Pearson CAB, Chao DL, Rojas DP, Recchia GL, Gómez-Dantés H, et al. Projected impact of dengue vaccination in Yucatán, Mexico. PLoS Negl Trop Dis. 2016;10(5):1–19. PMID: 27227883

130. Marshall BDL, Galea S. Formalizing the role of agent-based modeling in causal inference and epidemiology. Am J Epidemiol. 2015 Jan 15;181(2):92–9. PMID: 25480821

131. Heckbert S, Baynes T, Reeson A. Agent-based modeling in ecological economics. Ann N Y Acad Sci. 2010;1185:39–53. PMID: 20146761

132. Schindler J. About the uncertainties in model design and their effects: an illustration with a land-use model. J Artif Soc Soc Simul. 2013;16(4):6.

133. Ligmann-Zielinska A, Kramer DB, Cheruvelil KS, Soranno PA. Using uncertainty and sensitivity analyses in socioecological agent-based models to improve their analytical performance and policy relevance. PLoS One. 2014;9(10). PMID: 25340764

134. Feacham RG, Bradley DJ, Garelick H, Mara DD. Taenia, Taeniasis, and Cysticercosis. In: Sanitation and Disease: Health Aspects of Excreta and Waste Management. p. 463–72.

135. Lintusaari J, Gutmann MU, Dutta R, Kaski S, Corander J. Fundamentals and recent developments in approximate Bayesian computation. Syst Biol. 2017;66(1):e66–82. PMID: 27798401

136. Thiele JC, Kurth W, Grimm V. Facilitating parameter estimation and sensitivity analysis of agent-based models : A cookbook using NetLogo and R. J Artif Soc Soc Simul. 2014;17(3):11.

137. O'Neal SE, Pray IW, Vilchez P, Gamboa R, Muro C, Moyano LM, et al. Community cluster-randomized trial of geographically-targeted interventions versus mass drug administration for control of Taenia solium cysticercosis. Lancet Glob Heal. 2019;in review.

138. R-core. Parallel Package for R. 2018 (available from https://stat.ethz.ch/R-manual/R-devel/library/parallel/doc/parallel.pdf).

139. Learn Shiny (Web Tutorials). Shiny from RStudio. 2017 (available from https://shiny.rstudio.com/tutorial/).

140. Sobol M. Sensitivity estimates for nonlinear mathematical models. Math Model Comput Exp. 1993;1(4):407–14.

141. Sobol IM. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. Math Comput Simul. 2001;55:271–80.

142. Zhang XY, Trame MN, Lesko LJ, Schmidt S. Sobol sensitivity analysis: A tool to guide the development and evaluation of systems pharmacology models. CPT Pharmacometrics Syst Pharmacol. 2015;4(2):69–79. PMID: 27548289

143. Fonoberova M, Fonoberov VA, Mezić I. Global sensitivity/uncertainty analysis for agent-based models. Reliab Eng Syst Saf. 2013;118:8–17.

144. Homma T, Saltelli A. Use of Sobol's quasirandom sequence generator for integration of modified uncertainty importance measure. J Nucl Sci Technol. 1995;32(11):1164–73.

145. Kucherenko S, Albrecht D, Saltelli A. Exploring multi-dimensional spaces: a comparison of Latin Hypercube and Quasi Monte Carlo sampling techniques. Cornell University Library. 2015 (available from https://arxiv.org/abs/1505.02350). p. 1–30.

146. Saltelli A, Annoni P, Azzini I, Campolongo F, Ratto M, Tarantola S. Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. Comput Phys Commun. 2010;181(2):259–70. PMID: 15246654

147. Convertino M, Muñoz-Carpena R, Chu-Agor ML, Kiker GA, Linkov I. Untangling drivers of species distributions: Global sensitivity and uncertainty analyses of MaxEnt. Environ Model Softw. 2014;51:296–309.

148. Chan K, Saltelli A, Tarantola S. Sensitivity analysis of model output: variance-based methods make the difference. In: Proceedings of the 1997 Winter Simulation Conference. 1997. p. 261–8.

149. Iooss B, Janon A, Pujol G. Package "sensitivity": Global sensitivity analysis of model outputs. 2018 (available from https://cran.r-project.org/web/packages/sensitivity/sensitivity.pdf).

150. Wu J, Dhingra R, Gambhir M, Remais J V. Sensitivity analysis of infectious disease models: methods, advances and their application. J R Soc Interface. 2013 Sep 6;10(86):20121018. PMID: 23864497

151. McKay MD, Beckman RJ, Conover WJ, Mckay MD, Beckman RJ. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics. 1979;21(2):239–45.

152. Rasmussen R, Hamilton G. An approximate bayesian computation approach for estimating parameters of complex environmental processes in a cellular automata. Environ Model Softw. 2012;29(1):1–10. PMID: 67626402

153. de Aluja AS, Villalobos AN, Plancarte A, Rodarte LF, Hernandez M, Zamora C, et al. Taenia solium cysticercosis: immunity in pigs induced by primary infection. Vet Parasitol. 1999 Feb 25;81(2):129–35. PMID: 10030755

154. Gonzalez AE, Castro M, Gilman RH, Vargas G, Sterling CR, Garcia HH, et al. The marketing of cysticercotic pigs in the Sierra of Peru. Bull World Health Organ. 1993;71(2):223–8. PMID: 8490986

155. Delea MG, Nagel CL, Thomas EA, Halder AK, Amin N, Shoab AK, et al. Comparison of respondent-reported and sensor-recorded latrine utilization measures in rural Bangladesh: A cross-sectional study. Trans R Soc Trop Med Hyg. 2017;111(7):308–15. PMID: 29126213

156. Guezala MC, Rodriguez S, Zamora H, Garcia HH, Gonzalez AE, Tembo A, et al. Development of a species-specific coproantigen ELISA for human Taenia solium taeniasis. Am J Trop Med Hyg. 2009;81(3):433–7. PMID: 19706909

157. Marin JM, Pudlo P, Robert CP, Ryder RJ. Approximate Bayesian computational methods. Stat Comput. 2012;22(6):1167–80.

158. van der Vaart E, Beaumont MA, Johnston ASA, Sibly RM. Calibration and evaluation of individual-based models using Approximate Bayesian Computation. Ecol Modell. 2015;312:182–90.

159. Turner BM, Van Zandt T. A tutorial on approximate Bayesian computation. J Math Psychol. 2012;56(2):69–85. PMID: 10702483

160. Csilléry K, Lemaire L, François O, Blum MGB. Approximate Bayesian Computation (ABC) in R: A vignette. 2015 (available from ftp://202.162.217.53/CRAN/web/packages/abc/vignettes/abcvignette.pdf).

161. Flisser A, Gauci CG, Martinez-ocan J, Garza-rodriguez A, Dominguez-alpizar JL, Maravilla P, et al. Induction of protection against porcine cysticercosis by vaccination with recombinant oncosphere antigens. 2004;72(9):5292–7.

162. Gavidia CM, Verastegui MR, Garcia HH, Lopez-Urbina T, Tsang VCW, Pan W, et al. Relationship between serum antibodies and Taenia solium larvae burden in pigs raised in field conditions. PLoS Negl Trop Dis. 2013;7(5):1–8. PMID: 23658848

163. Sciutto E, Martínez JJ, Villalobos NM, Hernández M, José M V., Beltrán C, et al. Limitations of current diagnostic procedures for the diagnosis of Taenia solium cysticercosis in rural pigs. Vet Parasitol. 1998;79(4):299–313. PMID: 9831953

164. Lightowlers MW, Donadeu M. Designing a Minimal Intervention Strategy to Control Taenia solium. Trends Parasitol. 2017;33(6):426–34. PMID: 28236521

165. International Task Force for Disease Eradication. Disease considered as candidates for global eradication. Vol. 42. 2008.

166. Coffeng LE, Stolk WA, Hoerauf A, Habbema D, Bakker R, Hopkins AD, et al. Elimination of African onchocerciasis: Modeling the impact of increasing the frequency of ivermectin mass treatment. PLoS One. 2014;9(12):1–25. PMID: 25545677

167. Adewole M, Onifade A. A mathematical model of dracunculiasis epidemic and eradication. IOSR J Math. 2013;8(6):48–56.

168. Dixon MA, Braae UC, Winskill P, Walker M, Devleesschauwer B, Gabriël S, et al. Strategies for tackling Taenia solium taeniosis/cysticercosis: A systematic review and comparison of transmission models, including an assessment of the wider Taeniidae family transmission models. PLoS Negl Trop Dis. 2019;13(4):e0007301.

169. Rodriguez-Canul R, Argaez-Rodriguez F, De la GDP, Villegas-Perez S, Fraser A, Craig PS, et al. Taenia solium metacestode viability in infected pork after preparation with salt pickling or cooking methods common in Yucatán, Mexico. J Food Prot. 2002 Apr;65(4):666–9. PMID: 11952216

170. Gonzales AE, Garcia HH, Gilman RH, Gavidia CM, Tsang VCW, Bernal T, et al. Effective, single-dose treatment of porcine cysticercosis with oxfendazole. Am J Trop Med Hyg. 1996;54(4):391–4. PMID: 8615453

171. Galán-Puchades MT, Fuentes M V. Taenia asiatica: The most neglected human taenia and the possibility of cysticercosis. Korean J Parasitol. 2013;51(1):51–4. PMID: 23467406

Appendices

Appendix A: Supplemental tables for Chapter 3

Table A1. Characteristics of tracked pigs.

	Village A		Village B		Village C	
Characteristics	**Rainy**	**Dry**	**Rainy**	**Dry**	**Rainy**	**Dry**
Total pigs tracked	14	16	18	19	21	20
Sex (male, %)	6 (43%)	3 (19%)	4 (22%)	9 (47%)	13 (62%)	15 (75%)
Age, in months (mean, sd)	13 (6.0)	12 (5.9)	9 (6.9)	9 (5.6)	7 (4.4)	8 (4.2)
Herd size (mean, sd)	11 (7.3)	10 (6.4)	5 (3.3)	7 (4.6)	6 (3.8)	6 (4.2)

Table A2-1. Regression coefficients for the log-area of maximum, home, and core roaming ranges. Ordinary least squares regression models (Bivariate associations).

	Bivariate (e^{β} coefficients (95% CI))		
	Max. range	**Home range**	**Core range**
Village			
Village A	Ref.	Ref.	Ref.
Village B	0.50** (0.32, 0.77)	0.48** (0.30, 0.76)	0.53** (0.41, 0.69)
Village C	0.19** (0.12, 0.29)	0.24** (0.15, 0.39)	0.64** (0.49, 0.84)
Season			
Rainy	Ref.	Ref.	Ref.
Dry	0.40** (0.27, 0.59)	0.40** (0.27, 0.59)	0.70** (0.57, 0.88)
Household density†			
≤25	0.94** (0.92, 0.96)	0.95** (0.92, 0.97)	0.98* (0.97, 0.99)
>25	1.06** (1.03, 1.09)	1.05** (1.02, 1.09)	1.00 (0.99, 1.02)
Herd size (per pig)	1.07** (1.03, 1.11)	1.06** (1.02, 1.10)	1.01 (0.99, 1.04)
Pig sex			
Female	Ref.	Ref.	Ref.
Male	0.70 (0.45, 1.07)	0.78 (0.51, 1.19)	0.88 (0.70, 1.10)
Pig age (per month)	1.06** (1.02, 1.10)	1.04* (1.0, 1.08)	1.02* (1.0, 1.04)

p-value: **<0.01, *<0.05

§Significant statistical interactions (village*season)

†Number of households within 100m radius, linear spline at 25 households/100m

Table A2-2. Regression coefficients for the log-area of maximum, home, and core roaming ranges. Ordinary least squares regression models (Multivariate associations).

	Multivariate (e^{β} coefficients (95% CI))		
	Max. range	**Home range[§]**	**Core range**
Village			
Village A	Ref.	Ref.	Ref.
Village B	0.49** (0.33, 0.74)	***Rainy*:** 0.43** (0.24, 0.75) ***Dry*:** 0.51** (0.07, 0.22)	0.64** (0.50, 0.82)
Village C	0.19** (0.12, 0.28)	***Rainy*:** 0.14** (0.08, 0.24) ***Dry*:** 0.10** (0.05, 0.17)	0.52** (0.41, 0.67)
Season			
Rainy	Ref.	Ref.	Ref.
Dry	0.53** (0.38, 0.73)	***Village A*:** 0.24** (0.14, 0.44) ***Village B*:** 0.29** (0.17, 0.49) ***Village C:*** 0.70 (0.43, 1.16)	0.69** (0.57, 0.84)
Household density[†]			
≤25	-	-	-
>25	-	-	-
Herd size (per pig)	-	-	-
Pig sex			
Female	-	-	-
Male	-	-	-
Pig age (per month)	-	-	-

p-value: **<0.01, *<0.05

[§]Significant statistical interactions (village*season)

†Number of households within 100m radius, linear spline at 25 households/100m

Table A3-1. Regression coefficients for the *number of open defecation sites* within pig range areas (maximum, home, and core ranges). Negative binomial models for the number of defecation sites within maximum and home ranges (rate ratios [RR] with 95% confidence intervals [CI] displayed), and logistic regression model for the presence of ≥ 1 defecation site within the core range (odds ratios [OR] with 95% CI displayed). Bivariate associations.

	Bivariate		
	Max. range (RR, 95% CI)	**Home range (RR, 95% CI)**	**Core range (OR, 95% CI)**
Village			
Village A	Ref.	Ref.	Ref.
Village B	5.18** (3.53,7.60)	7.06** (3.83, 13.01)	6.72* (1.37, 33.0)
Village C	1.02 (0.67, 1.56)	1.25 (0.63, 2.49)	1.11 (0.17, 0.30)
Home range size (log-area, m^2)	1.47** (1.17, 1.84)	1.50** (1.13, 2.0)	2.45* (1.03, 5.82)
Household density†			
≤25	1.02 (0.99, 1.05)	1.03 (0.99, 1.07)	1.04 (0.97, 1.12)
>25	1.04* (1.0, 1.08)	1.03 (0.99, 1.08)	1.01 (0.94, 1.08)
Pig sex			
Female	Ref.	Ref.	Ref.
Male	0.80 (0.52, 1.23)	0.94 (0.54, 1.66)	1.03 (0.37, 2.93)
Season			
Rainy	Ref.	Ref.	Ref.
Dry	0.86* (0.72, 1.02)	0.69 (0.39, 1.21)	1.10 (0.39, 3.11)
Herd size (per pig)	0.97 (0.93, 1.01)	0.97 (0.91, 1.03)	1.01 (0.91, 1.11)
Pig age (per month)	0.99 (0.95, 1.02)	0.98 (0.93, 1.03)	1.01 (0.91, 1.12)
Roaming land-type (per 1% increase by type)‡			
Peri-domestic	0.18** (0.07, 0.43)	0.17** (0.05, 0.56)	0.29 (0.04, 2.24)
Vegetation	1.11 (0.41, 3.02)	2.13** (0.58, 7.81)	1.62 (0.19, 13.8)
Roads/paths	16.9** (4.52, 63.1)	5.74 (0.83, 39.7)	2.34 0.10, 53.7)

p-value: **<0.01, *<0.05

§Significant statistical interactions (village*season)

†Number of households within 100m radius, linear spline at 25 households/100m

‡Farming land-type not included due to insufficient roaming

Table A3-2. Regression coefficients for the *number of open defecation sites* within pig range areas (maximum, home, and core ranges). Negative binomial models for the number of defecation sites within maximum and home ranges (rate ratios [RR] with 95% confidence intervals [CI] displayed), and logistic regression model for the presence of ≥ 1 defecation site within the core range (odds ratios [OR] with 95% CI displayed). Multivariate associations.

	Multivariate		
	Max. range (RR, 95% CI)	**[§]Home range (RR, 95% CI)**	**Core range (OR, 95% CI)**
Village			
Village A	Ref.	Ref.	Ref.
Village B	5.3** (3.61,7.78)	***Female*:** 16.3** (6.25, 42.3) ***Male*:** 3.84** (1.77, 8.39)	21.8** (2.74, 173)
Village C	1.48 (0.91, 2.41)	***Female*:** 2.45 (0.76, 7.9) ***Male*:** 0.63 (0.25, 1.57)	4.59 (0.43,48.5)
Home range size (log-area, m^2)	1.87** (1.59, 2.21)	1.69** (1.39, 2.07)	5.38* (1.50, 19.2)
Household density[†]			
≤25	1.06** (1.05, 1.08)	1.08** (1.05, 1.11)	-
>25	1.28** (1.11, 1.48)	0.95** (0.93, 0.98)	-
Pig sex			
Female	-	Ref.	
Male	-	***Village A:*** 5.10** (1.72, 15.1) ***Village B:*** 1.21 (0.80, 1.81) ***Village C:*** 1.28 (0.57, 2.92)	
Season			
Rainy	-	-	-
Dry	-	-	-
Herd size (per pig)	-	-	-
Pig age (per month)	-	-	-
Roaming land-type (per 1% increase by type)[‡]			
Peri-domestic	-	-	-
Vegetation	-	-	-
Roads/paths	3.41** (1.79, 6.49)	-	-

p-value: **<0.01, *<0.05

[§]Significant statistical interactions (village*season)

†Number of households within 100m radius, linear spline at 25 households/100m

‡Farming land-type not included due to insufficient roaming

Appendix B: Supplemental material, tables, and figures for Chapter 4

Supplement B1. Data sources and statistical methods for CystiAgent parameters.

SOURCE #1: Ring Strategy Trial (RST)

Model parameters:

humans-per-hh	mean number of people assigned to each household
prop-pig-owners	proportion of households that raise pigs
pigs-per-hh	mean number of pigs assigned to each pig-raising household
prop-corrals	proportion of pig-raising households that own corrals for their pigs
prop-latrines	proportion of households that have access to a latrine
slaughter-age	mean age at which pigs are slaughtered

Description of study:
This was a large cluster-randomized trial comparing the effectiveness of "Ring-strategy" versus mass-treatment as control strategies to reduced *Taenia solium* transmission. The trial was carried out in 23 rural villages of the northern Peruvian region of Piura between 2015-2017. Data from this study that were used in the CystiAgent model included the baseline census, which collected demographic variables for all households in the study, and serial follow-up of cohorts of pigs, which were captured every 4 months to collect serum samples. The RST study was funded by NIH grant number NIH R01-NS080645 with Seth O'Neal as principle investigator.

Methods/results:
Village input parameters. A household-level census was conducted at baseline in all 23 study villages, and attempted to gather information on all residents of the study villages. Variables recorded for the census included demographics of all human inhabitants, the condition of the house, access to water and sanitation, and livestock including pigs. For each variable of interest, results were summarized by village. The plausible ranges for parameters were determined by the maximum and minimum means observed across all villages. Probability distributions for discrete parameters (*humans-per-hh, pigs-per-hh)* were determined by fitting raw census data to a range of distributions (e.g., normal, log-normal, Poisson, exponential) and selecting the distribution with the best fit based on AIC values (Akaike Information Criterion). The Poisson and exponential distributions chosen to represent *humans-per-hh* and *pigs-per-hh*, respectively are defined by a single mean value, and were truncated at 1 to prevent household with 0 inhabitants.

Parameter	Distribution	Mean	Lower	Upper
humans-per-hh	Poisson	3.89	3.32	4.94
prop-pig-owners	Binomial	0.49	0.25	0.75
pigs-per-hh	Exponential	2.44	1.74	4.21
prop-corrals	Binomial	0.5	0.23	0.92
prop-latrines	Binomial	0.64	0.19	0.97

slaughter-age: The age of pigs at slaughter was estimated for two cohorts of study pigs (n=1,284 pigs) that entered the study at baseline and month 4 (M4), when they were between 2-4 months

old (this is when age estimation is most accurate). Pigs were captured for serum-sampling every 4 months throughout the study. Therefore, when a pig from one of these cohorts was censored from the study (not captured in the following sample), its age at censorship (i.e., slaughter) was assumed to be the age it would have reached at the mid-point of the previous 4-month interval. Pigs still alive at the end of the study (n=47 pigs) were conservatively given slaughter-ages of their age in the last sampling round plus two additional months. Final slaughter ages were analyzed to determine the probability distribution and summary statistics to best describe them. A log-normal distribution was chosen as the optimal fit, and the log-mean and log-standard-deviation were estimated from the data. The "plausible range" for log-mean was derived from a 95% confidence interval produced through bootstrapped resampling (n=1000).

Slaughter-age to follow a LOG-NORMAL distribution			
	Value	**Lower**	**Upper**
Log-mean	2.279	2.249	2.305
Log-SD	0.515	-	-

SOURCE #2: Household survey (HH)

Model parameters:

pigs-sold	Proportion of pigs sold prior to slaughter
pigs-exported	Proportion of sold pigs that are exported to other villages
pig-import-rate	Rate of pigs imported from other endemic villages (import / pig / week)
hh-only-pork	Proportion of slaughtered pigs that are consumed exclusively by the owner's household
sold-pork	Proportion of slaughtered pigs that are exclusively sold to another household in the same village
shared-pork-hh	Among slaughtered pigs shared between the owner and other households, the proportion pork-meat eaten by the owner's household
traveler-prop	Proportion of households that have a member who regularly travels
travel-freq	Interval of time (in weeks) between trips to other endemic villages
travel-duration	Average duration of trips to other endemic villages

Description of study:

The "household survey" was a door-to-door survey applied in 7 rural villages of northern Peru (Piura region) between 2017-18. The survey was applied to all heads-of-household that resided in the study villages, and was carried out over two time-points four months apart (n1=420 and n2=410 households). The survey was applied as part of a community-based study that aimed to identify methods for improving reporting of infected pigs to the health post. Survey questions used for the CystiAgent model were embedded in a larger survey that assessed knowledge, behaviors, and attitudes towards cysticercosis. The parent study was funded by NIH grant number NIH R01-NS080645 with Seth O'Neal as principle investigator.

Methods/results:

pigs-sold: The survey question asked (translated from Spanish): "How many pigs from your household have you sold prior to slaughter in the past 4 months?" The total number of pigs sold in each village was divided by the total number of pigs slaughter *or* sold in the same interval to determine the proportion of pigs sold among those due for sale or slaughter. The final parameter value was determined by averaging the results of the two survey time-points. The "plausible range" was determined by selecting the minimum and maximum values from among the seven villages surveyed.

	Estimate	LL	UL
Total pigs sacrificed or sold	760		
Pigs SOLD	51.4% (391/760)	32.6%	75.4%

pigs-exported: The survey question asked (translated from Spanish): "How many pigs from your households have you sold prior to slaughter *outside of the village* in the past 4 months?" The total number of pigs sold *outside of the village* (i.e., exported) was divided by the total number of pigs sold to determine the proportion of sold pigs that were exported. The final parameter value was determined by averaging the results of the two survey time-points. The "plausible range" was determined by selecting the minimum and maximum values from among the seven villages surveyed.

	Estimate	LL	UL
Total pigs sold	391		
Pigs EXPORTED	73.1% (286/391)	34.2%	100%

pig-import-rate: The survey question asked (translated from Spanish): "How many live pigs have your purchased from *outside the village* in the past 4 months?" The total number of pigs imported was divided by the total number of pigs in the village and the 4-month period (17 weeks) to determine the number of pigs imported per pig in the population per week. The final parameter value was determined by averaging the results of the two survey time-points. The "plausible range" was determined by selecting the minimum and maximum values from among the seven villages surveyed.

Parameter	Estimate	LL	UL
Total pigs in the village	1956		
Total pigs purchased externally	35		
IMPORTED pigs (per pig per week)	0.00105	0	0.00384

hh-only-pork, sold-pork: The survey questions asked (translated from Spanish): "How many pigs from your households have you slaughtered in the past 4 months?" and, among those, "How many were consumed exclusively by members of your household?", "How many were sold or gifted after slaughter to other households?" and "How many were shared between members of your household and other households?" The proportion of pigs eaten at home, sold, and shared were calculated by dividing each total by the total number of pigs slaughtered. The proportion shared is not explicitly defined as a model parameter because it is represented by the proportion of pigs that remain after the first two parameters are applied. The final parameter value was determined by averaging the results of the two survey time-points. The "plausible range" was determined by selecting the minimum and maximum values from among the seven villages surveyed.

Parameter	Estimate	LL	UL
Total pigs slaughtered by household	366		
Pigs consumed exclusively by HOUSEHOLD	39.6% (145 / 366)	21.7%	71.4%
Pigs exclusively SOLD to other households in village	11.5% (42 / 366)	0%	50.0%
Pigs SHARED between household and other households in village	48.9% (179 / 366)	12.5%	75.0%

shared-pork-hh: The survey question asked (translated from Spanish): "How many kilos of pork meat have you purchased from within your village in the past 4 months?" This total was divided by the total number of kilos of pork meat shared or sold within the village, which was determined taking the total number of pigs sold or shared in the village and applying an average of 50 kg per pig times 0.3 to represent the edible portion of the pig. This allowed for estimation of the proportion of shared pork that was eaten at home versus shared for each village.

Parameter	Estimate	LL	UL
Total sacrificed pigs whose meat was SHARED/SOLD	221		
Number of kg of pork (pigs*50kg*0.3)	7735kg		
Number of kg of pork purchased from WITHIN	1545.5kg		
Proportion of pork eaten by HOUSEHOLD	80% (6189.5 / 7735kg)	0%	83.9%

Travel-related parameters: A section of the survey asked respondents to list all trips taken in the past 4 months by any member of the household for which the traveler spent at least 1 night outside of the village. The person, location, and duration of each trip were recorded. In analysis, destinations were evaluated to determine if they were endemic for *T. solium* transmission. Given that travel was likely to be underreported and most destinations were endemic areas, we chose not to exclude non-endemic areas. Final parameter values were determined by averaging the results of the two survey time-points. The "plausible range" was determined by selecting the minimum and maximum values from among the seven villages surveyed.

traveler-prop: The proportion of households that have a traveler was determined by dividing the number households that reported at least one trip by the total number of households.

Parameter	Estimate	LL	UL
Number of households	828		
At least 1 TRIP ("TRAVELER HH")	42.3% (350/828)	24.6%	65.4%

travel-freq: The frequency of travel among travelers was determined by counting the total number of trips completed by each household in the past 4-months, and using this to determine the mean number of trips per week and the mean interval between trips.

Parameter	Estimate	LL	UL
Number of travelers	350		
MEAN # of trips in 4-month period	2.16	1.08	3.5
MEAN # of trips / week	0.12	0.06	0.20
MEAN frequency of trips (weeks between trips)	8.00	16.02	4.94

travel-duration: The mean duration of travel was determined by averaging the duration of every trip reported by a traveling household. Final travel-durations were analyzed to determine the probability distribution and summary statistics to best describe them. An exponential distribution was chosen as the optimal fit by comparing AIC values for evaluated distributions.

Parameter	Estimate	LL	UL
Number of travelers	350		
MEAN trip duration (days)	12.28	5.9	23.5
MEAN trip duration in weeks:	1.75	0.84	3.36

SOURCE #3: GPS pig tracking study (GPS)

Model parameters:

latrine-use	Among households with access to latrines, proportion that are in "good" condition and are "always" used by all household inhabitants
cont-radius	Distance from household at which open defecation occurs among household not using latrines
corral-always	Among households that raise pigs and own pig-corrals, the proportion of corrals that are in "good" condition and are "always" used to contain all pigs
corral-sometimes	Among households that raise pigs and own pig-corrals, the proportion of corrals that are "sometimes" used to contain pigs
prop-corral-some	Among pigs raised in households that "sometimes" contain pigs in corrals, the proportion of pigs contained at any given time
home-range	Radius of the area pigs cover when roaming (i.e., not contained in corrals), and within which exposure to *T. solium* is assumed to occur

Description of study:
A detailed description of this study is can be found in Chapter 3. The study consisted of two separate activities, both carried out in 2018 in three rural villages of northern Peru (Piura region). First, we conducted a door-to-door survey of all households in the villages. The survey asked adult heads-of-household about the presence and use of latrines and pig-raising practices. Second, we conducted GPS tracking of a sample of free-roaming pigs. Overall, we tracked 108 pigs for 6 days each, and included GPS tracking in both the rainy and dry season. Roaming patterns were then analyzed to determine the size of each pig's "home-range."

Methods/results:
latrine-use: Heads-of-households were asked if their household had access to a latrine/indoor bathroom, or if they used outdoor areas to defecate. If a latrine was present, we then asked how often members of the household used the latrine, giving the options "always," "sometimes," or "never." The option of "always" was only recorded if *all* members of the household, including children, were reported to *always* use the latrine/bathroom. We also inspected the condition of the latrine, and recorded it as "good," "normal," or "bad," and inspected the areas around the house for evidence of feces or soiled paper. In the model, "latrine-use" was defined as the proportion of households with latrines that reported "always" using latrines, and for which latrine condition was "good" and no evidence of feces was observed in the household area. The final parameter value was determined by averaging the results of the three villages. The "plausible range" was determined by selecting the minimum and maximum values from among the villages surveyed.

Among households with latrines, the proportion of HH's that are in GOOD condition and are used ALWAYS		
Estimate	**LL**	**UL**
73% (142 / 194)	57%	86%

cont-radius: For households that reported the practice of open defecation, respondents were asked to indicate the location of the defecation area, and a GPS point was recorded at the location where feces or soiled paper was visualized. We then calculated the distance between the open defecation point and the household (GPS point taken at the front door), and labeled this distance as the "contamination radius" for each house. Final contamination radii were analyzed to determine the probability distribution and summary statistics to best describe them. A log-normal distribution was chosen as the optimal fit by comparing AIC values for evaluated distributions. The "plausible

range" was derived from a 95% confidence interval of the log-normal distribution. When applied in the model, tapeworm carriers will shed infectious *T. solium* eggs and proglottids onto a location that is determined by generated a random value from the *cont-radius* distribution and applying a random angle from the household location.

	Estimate (m)	**LL**	**UL**
Log-mean	3.27	3.14	3.40
exp(log-mean)	**26.3**	**23.0**	**30.1**
Log-SD	**0.546**		

corral-always, corral-sometimes: For households that reported raising pigs, pig-owners were asked if they allowed their pigs to roam freely or kept their pigs contained in corrals. A response of "always" indicated that *all* pigs (including piglets) were *always* kept enclosed (or tied). Reports were confirmed or adjusted based on observation of pig-owners current practices, and were further validated by cross-referencing reports with pig-level data collected in the same month in a separate pig sero-survey (each pig reported as "free" of "enclosed"). The "plausible range" was determined by selecting the minimum and maximum values from among the villages surveyed. For application in the model, pig-raising households that possessed corrals were assigned as either "always," "sometimes," or "never" users of corrals. Pigs in households that "sometimes" used corrals were either corralled or not corralled depending on *prop-corral-some*.

	Estimate	**LL**	**UL**
Always	35% (36 / 102)	33%	39%
Sometimes	57 % (58 / 102)	53%	62%
Never	8% (8 / 102)	4%	13%

prop-corral-some: The proportion of pigs that were corralled among households that "sometimes" corral their pigs was determined by cross-referencing survey responses (household-level) with pig-level data from a serological survey that took place in the same month. Households that reported "sometimes" using corrals were found in the serological data, and the numbers of pigs that were "contained" vs. "free-roaming" at the time of the sero-survey were recorded. The "plausible range" was determined by selecting the minimum and maximum values from among the villages surveyed.

	Estimate	**LL**	**UL**
Total pigs in "sometimes" households	261		
Contained	32% (83 / 261 pigs)	15%	44%
Free-roaming	68% (178 / 261 pigs)	56%	85%

home-range: The area of each pig's "home-range" was calculated using the Localized Convex Hulls (LoCoH) home range algorithm. Home-range areas represent the area of active foraging that represents the densest 90% of a pigs range. The total area (m^2) was calculated for each pig, and was converted to a radius that assumed each range followed a circular shape. Final home-range radii were analyzed to determine the probability distribution and summary statistics to best describe them. A log-normal distribution was chosen as the optimal fit by comparing AIC values for evaluated distributions. In the sub-analysis, a significant difference between villages and seasons was detected. Therefore, upper and lower limits were determined by extracting the mean-log from the largest and smallest village/season combinations.

	Estimate	LL	UL
mean-log	**3.79**	**3.39**	**4.56**
exp(mean-log)	44.2	29.7	95.5
sd-log	**0.552**		

SOURCE #4: Literature review and expert opinion

Model parameters:

import-prev	Proportion of imported pigs that have cysticercosis
light-to-heavy	Proportion of infected imported pigs that have light vs. heavy cyst infection
travel-incidence	Rate of human taeniasis infection during travel to external endemic areas
tn-incubation	Latency period after initial infection (taeniasis) before beginning to expel eggs
tn-lifespan	Mean duration of taeniasis infections
decay-mean	Mean time eggs remain viable in the environment before decaying

import-prev: The proportion of imported pigs that had cyst infection was assumed to be consistent with the prevalence of cyst infection expected in a standard endemic village of northern Peru. We therefore estimated this prevalence using a necropsy study conducted in this region of Peru in 2017 [75]. Since only a sample of seropositive pigs were necropsied in this study, but all pigs were tested with the EITB assay, we estimated the total number of infected pigs by applying the proportion infected at each EITB band level (1-7 positive bands) to the total number of pigs at each band level. Given the limited data for this parameter, the "plausible range" of 0 to 30% was set manually based on plausible limits determined by our expertise.

WB # bands	Prevalence of cyst infection in necropsy sample	Pig population at band level (n=828 pigs)	Expected # with cyst infection
0	0/0 = 0%	395/828 = 47%	395*0 = 0
1	2/22 = 9.1%	87/828 = 10.5%	87*0.091= 7
2	4/42 = 9.5%	127/828 = 15.3%	127*0.095 = 12
3	17/52 = 32.7%	142/828 = 17.2%	142*.327 = 46
4	6/18 = 33.3%	37/828 = 4.5%	37*0.333 = 12
5	6/10 = 60%	16/828 = 1.9%	16*0.60 = 10
6	5/5 = 100%	9/828 = 1.1%	9*1.0 = 9
7	9/9 = 100%	15/828 = 1.8%	15*1.0 = 15
TOTAL	**49 / 158 = 31%**	**433 / 828 pigs = 52%**	**111 / 828 = 13.4%**

light-to-heavy: The proportion of infected imported pigs that had light vs. heavy cyst infection was based on the same necropsy study referenced above [75]. Among infected pigs that were found on necropsy, those with < 100 viable cysts were considered lightly infection and those with ≥ 100 viable cysts were heavily infected. Given the limited data for this parameter, the "plausible range" of 50% to 100% of infected imported pigs with *light* infection was set manually based on what our experts believed could be the plausible limits.

Parameter	Estimate
Total pigs necropsied	158
Necropsy-positive (1+ cyst)	49
Light infection (1-99 cysts)	**75.5% (37 / 49)**
Heavy infection (>100 cysts)	**24.5% (12 / 49)**

travel-incidence: The incidence of human tapeworm infections (taeniasis) among travelers traveling to other endemic areas was assumed to be the same as the incidence observed in a standard endemic village. However, the incidence of taeniasis in endemic villages is unknown and has not been published. We therefore estimated taeniasis incidence based on the known prevalence in this region and an estimate of the duration of infection using the equation: Incidence = Prevalence / Duration. The prevalence was based on a cross-sectional study of taeniasis prevalence conducted in this region [75]. See below for explanation of taeniasis duration used for this calculation (2 years). The "plausible range" was determined by calculating incidence with the minimum and maximum prevalence observed among the 7 villages included in the original study. Taeniasis incidence was then applied in the model as the probability of infection per person per week of travel.

Parameter	**Estimate**	**LL**	**UL**
Taeniasis prevalence	2.4% (34/1420)	0.9% (2/218)	5.3% (8/151)
Taeniasis duration	104 weeks	26 weeks	208 weeks
INCIDENCE (probability of infection per week)	**0.024 / 104 = 0.000231**	**0.009 / 208= 0.000043**	**0.053 / 26 = 0.0020**

tn-incubation: The period of time after initial ingestion of a viable *T. solium* cyst and growth of a mature tapeworm that expels infection eggs and proglottids is widely reported in literature to be approximately 2 months (8 weeks). See [22,26,27]. No "plausible range" or probability distribution was utilized for this fixed parameter value.

tn-lifespan: Humans that are infected with tapeworm will have a natural duration of infection that is determined by this parameter. Taeniasis duration is unknown and has not been studied due to ethical barriers. A range of 2-4 years is typically reported, and is based on age-specific prevalence data, biological plausibility, and auto-infection of a scientist in 1935. See [Garcia, Yoshino]. A zero-truncated normal distribution with standard deviation of 50 weeks was chosen for simplicity, but is not supported by data. In the model, the duration of each individual tapeworm infection is randomly drawn from this distribution.

Estimate	**LL**	**UL**
Mean = 104 weeks (2 years)	26 weeks (0.5 years); Truncated at 0	208 weeks (4 years)
SD = 50	SD = 50 95% LL = 0 95% UL = 132 weeks (2.5 years)	SD = 50 95% LL = 125 weeks (2.4 years) 95% UL = 290 weeks (5.6 years)

decay-mean: If a tapeworm carrier is practicing open defecation, they will contamination their environment with *T. solium* eggs and proglottids. After the infection ends, eggs (but not proglottids) will persist in the environment until they naturally decay. The longevity of eggs in the environment is determined by this parameter. Estimates for the longevity of eggs in the environment are based on experimentation done on eggs of related species in the 1970s [134]. These experiments indicate that longevity is impacted by temperature and moisture. The mean and plausible range was derived from observations from a variety of experiments conducted in different conditions. In the model, a fixed probability of decay is applied each week, which creates an exponential survival function with the mean inversely related to the probability of decay.

Eggs decay in the environment according to the EXPONENTIAL SURVIVAL FUNCTION, where at each time (t) there is a constant probability of decay (λ), the survival function is therefore given by f(t) = exp(-λt), with mean decay time given by E(t) = 1/λ		
Estimate:	**LL**	**UL**
λ = 0.125	**1.0**	**0.038**
E(t) = 8 weeks	1 week	26 weeks

Table B2-1. CystiAgent model parameters and plausible ranges used in sensitivity analyses (part 1 of 2)

Parameter	Code	Distribution	Value	Plausible range		Source
				Lower	Upper	
Village input features						
Humans per household	humans-per-hh	Poisson	3.89	3.32	4.94	RST
Proportion of households raising pigs	prop-pig-owners	Binomial	0.49	0.25	0.75	
Pigs per pig-raising household	pigs-per-hh	Exponential	2.44	1.74	4.21	
Corral prevalence among pig-owner households	prop-corrals	Binomial	0.5	0.23	0.92	
Latrine prevalence	prop-latrines	Binomial	0.64	0.19	0.97	
Pig sale, import, export, and slaughter						
Pig slaughter age (median)	slaughter-age	Log-normal	9.8 months	9.5	10.0	RST
Proportion of pigs sold prior to slaughter	pigs-sold	Binomial	0.51	0.33	0.75	HH
Proportion of sold pigs exported	pigs-exported	Binomial	0.73	0.34	1	HH
Rate of pigs imported from endemic areas (imports / pig / week)	pig-import-rate	Uniform	0.00105	0	0.00384	HH
Prevalence of cyst infection among imports	import-prev	Binomial	0.134	0	0.3	[75]
Proportion of infected imported pigs with light cyst burden	light-to-heavy	Binomial	0.76	0.5	1	[75]
Proportion of pork consumed by owner	hh-only-pork	Binomial	0.40	0.22	0.71	HH
Proportion of pork sold after slaughter	sold-pork	Binomial	0.12	0	0.5	HH
Proportion of shared pork eaten by owner	shared-pork-hh	Binomial	0.8	0	0.84	HH
Tapeworm maturation and death						
Incubation time to reach tapeworm maturity	tn-incubation	Fixed	8 weeks	-	-	[22,27]
Tapeworm lifespan (mean, sd = 1 year)	tn-lifespan	Normal	2 years	0.5	4	
Human travel						
Proportion of households with a frequent traveler	traveler-prop	Binomial	0.42	0.24	0.65	HH
Frequency of travel to other endemic areas (every X weeks)	travel-freq	Uniform	8 weeks	5	16	HH
Duration of travel	travel-duration	Exponential	1.75 weeks	0.84	3.36	HH
Incidence of *T. solium* taeniasis during travel (risk / person / week)	travel-incidence	Uniform	0.00023	0.00004	0.002	[75]
Environmental contamination and decay						
Latrine-use (prop. of households that “always” use latrine)	latrine-use	Binomial	0.25	0	0.86	GPS
Radius of environmental contamination (median, meters from home)	cont-radius	Log-normal	26 meters	23	30	GPS
Rate of egg decay in environment (mean survival duration)	decay-mean	Exponential	8 weeks	1	26	[134]
Pig roaming and exposure to *T. solium* eggs						
Proportion of pig households with corrals that “always” corral pigs	corral-always	Binomial	0.05	0	0.39	GPS
Proportion of pig households with corrals that “sometimes” corral pigs	corral-sometimes	Binomial	0.57	0.25	0.61	
Proportion of pigs in “sometimes”-corral-households that are corralled	prop-corral-some	Binomial	0.32	0.15	0.44	
Radius of pig roaming “home-range” (median)	home-range	Log-normal	44 meters	30	96	

†Exposure probabilities (“light-all” and “heavy-all”, x) scaled to the current number of tapeworm carriers (HT) according to $1 – (1-x)^{HT}$
HH=Household survey; GPS=GPS pig tracking study, RST=Ring Strategy Trial

Table B2-2. CystiAgent model parameters and plausible ranges used in sensitivity analyses (part 2 of 2)

Parameter	Code	Distribution	Value	Plausible range		Source
				Lower	Upper	
Tuning parameters						
Probability of human taeniasis upon slaughter of lightly infected pig	pl2h	Binomial	NA	0.03	0.4	NA
Probability of human taeniasis upon slaughter of heavily infected pig	ph2h	Binomial	NA	0.003	0.04	
Probability of light cyst infection upon contact with *T. solium* eggs	light-inf	Binomial	NA	0.003	0.02	
Probability of heavy cyst infection upon contact with *T. solium* proglottids	heavy-inf	Binomial	NA	0.003	0.02	
Probability of exposure to *T. solium* eggs per human with taeniasis†	light-all	Binomial	NA	0	0.05	
Probability of exposure to *T. solium* proglottids per human with taeniasis†	heavy-all	Binomial	NA	0	0.05	

†Exposure probabilities ("light-all" and "heavy-all", x) scaled to the current number of tapeworm carriers (HT) according to $1 - (1-x)^{HT}$
HH=Household survey; GPS=GPS pig tracking study, RST=Ring Strategy Trial

Supplement B3: Graphical results of CystiAgent sensitivity analyses

Sobol' Sensitivity Analysis. All villages (low, medium, high-density); full and reduced models. Parameters with $S_i > 0.02$ in the full model and $S_i > 0.01$ in the reduced model shown.

	Full Model (k = 33 parameters)			Reduced model (k = 22 parameters)		
Parameters	High	Medium	Low	High	Medium	Low
Pig						
tuning-pig[§]	●●	●●●	●●●			
corral-always	●●●	●●●	●●	● ■	■	
prop-corrals	●●	●●	●			
corral-sometimes	●					
prop-corral-some				●		
home-range				●● ■■	●● ■	●●
latrine-use				● ■	● ■	●● ■
humans-per-hh			●			
cont-radius				● ■		
home-range-sd				■		
Human						
tuning-human[§]	■■	■■■	■■ ●			
prop-pig-owners	■■	■■■	■■■			
tn-lifespan	■■	■ ●	●●	■■ ●●●	■■ ●●●	■■ ●●●
pigs-per-hh	■■	■	■ ●			
pigs-exported	■	■	■	■■ ●	■■ ●	■■ ●
pigs-sold	■	■		■ ●	■ ●	■ ●
hh-only-pork		■			■	
shared-pork-hh				■ ●		
pig-import-rate				■ ●	■	■ ●
import-prev						■
sold-pork				●		
travel-duration				■		
travel-incidence				■ ●		

■ = Human taeniasis, ● = Porcine cysticercosis

SOBOL first-order indices (S_i): ■■■ / ●●● > 0.25; ■■ / ●● > 0.1; ■ / ● > 0.02

[§]tuning-pig and tuning-human refer collectively to the set of tuning parameters defining the probabilities of pig and human infection.

LHS-PRCC Analysis. All villages (low, medium, high-density); full and reduced models. Only parameters with significant correlation coefficients (p< 0.0015) are displayed.

	Full Model (k = 33 parameters)			Reduced Model (k = 22 parameters)		
Parameters	High	Medium	Low	High	Medium	Low
Pig						
tuning-pig[§]	●●● ■■	●●● ■■	●●● ■■■			
corral-always	●●● ■■	●●● ■■	●●● ■■	●● ■■	●● ■	● +
prop-corrals	●●● ■	●●● ■■	●●● ■			
corral-sometimes	+	● +	● +	● +	+ +	+
prop-corral-some	+	●● +	● +	● +	+ +	+ +
home-range	+ +	+ +	● +	●●●■■ ■	●●● ■■	●● ■■
latrine-use	+	+ +	● +	●● ■■	●● ■■	●● ■■
humans-per-hh	+ +	+ +	● +			
prop-latrines	+	+ +	+ +			
slaughter-age	+	+	+	+	+ +	+
decay-mean			+	+ +	+ +	+ +
cont-radius				+	+ +	
home-range-sd				+ +	+ +	+ +
Human						
tuning-human[§]	■■■ +	■■■ ●	■■■ ●●			
prop-pig-owners	■■■ +	■■■ +	■■■ ●●			
tn-lifespan	■■■ +	■■■ ●	■■ ●●	■■■●● ●	■■■●●●	■■ ●●●
pigs-per-hh	■■■ +	■■■ +	■■ ●			
pigs-exported	■■■ +	■■■ +	■■ ●	■■■ ●●	■■■ ●●	■■ ●●
pigs-sold	■■ +	■■ +	■■ ●	■■ ●●	■■ ●●	■■ ●●
hh-only-pork	■	■	■ +	+ +	■ +	+ +
shared-pork-hh	■	■	■ +	■ +	■ +	+ +
pig-import-rate	■	+ +	+ +	■■ ●	■ ●	■ ●
import-prev	■	■	+ +	■■ ●	■■ ●	■ ●
sold-pork	+	+	+	+ +	+ +	
travel-duration	+					
travel-incidence				+ +		+ +
light-to-heavy				+	+ +	+ +
traveler-prop					+ +	

■ = Human taeniasis, ● = Porcine cysticercosis

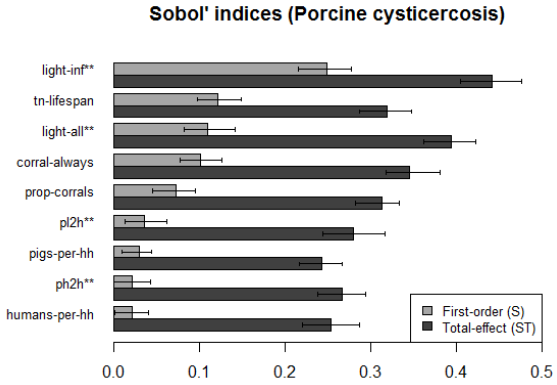
PRCC (absolute-value of ρ): ■■■ / ●●● > 0.5; ■■ / ●● > 0.25; ■ / ● > 0.1; + / + < 0.1; (all p < 0.0015)

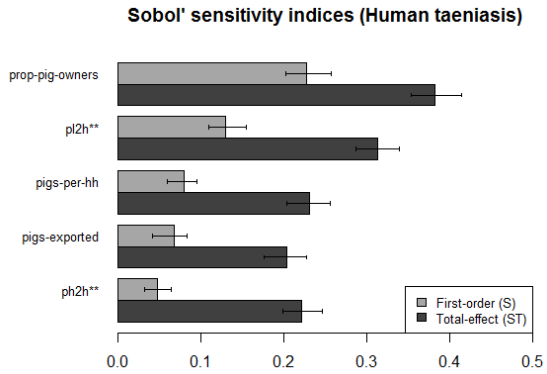
[§]tuning-pig and tuning-human refer collectively to the set of tuning parameters defining the probabilities of pig and human infection.

FULL ANALYSIS (k = 33 parameters)

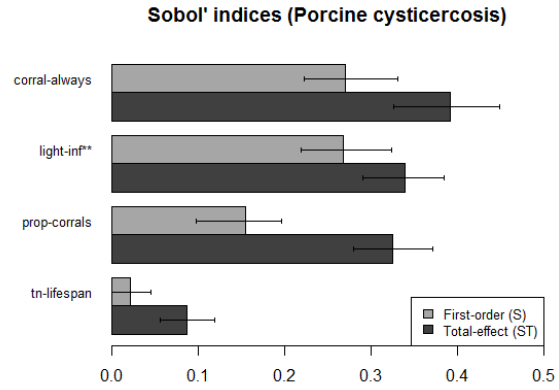
Sobol' sensitivity analyses indices

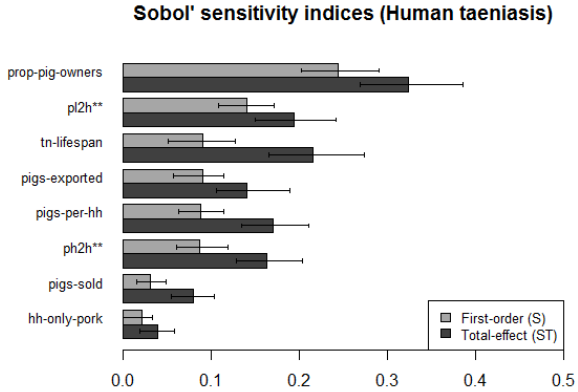
- Low-density village



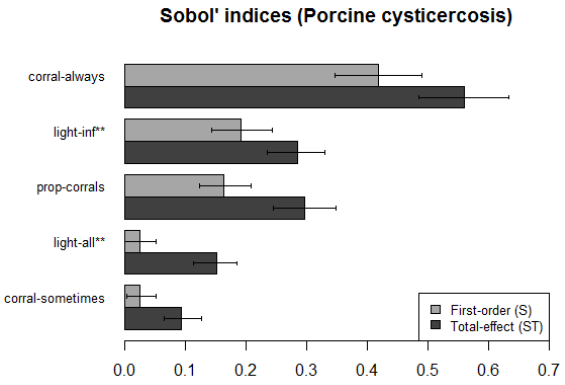


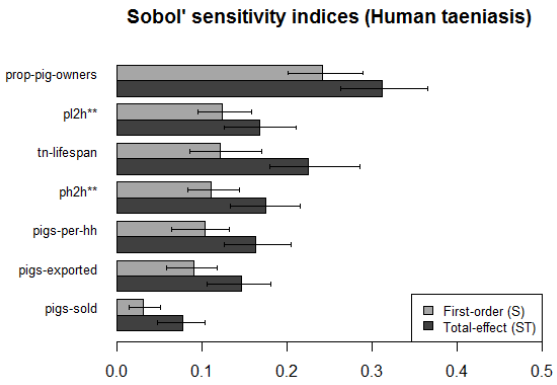
- Medium-density village





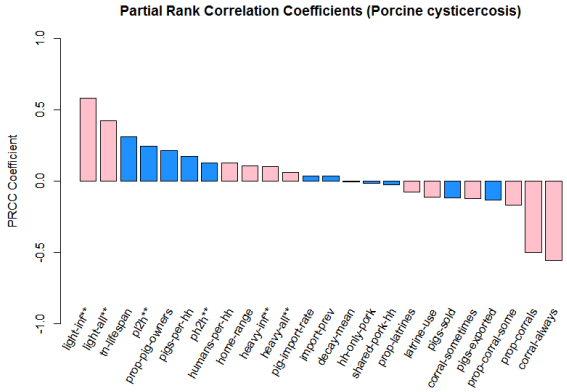
- High-density village

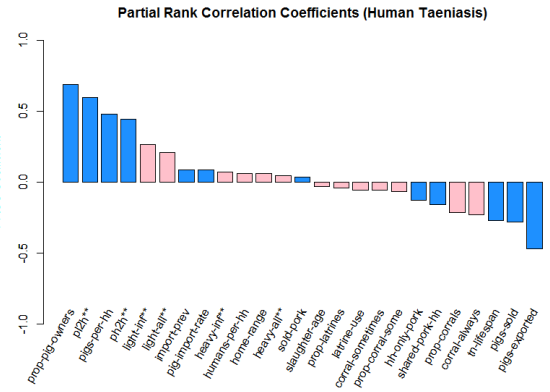




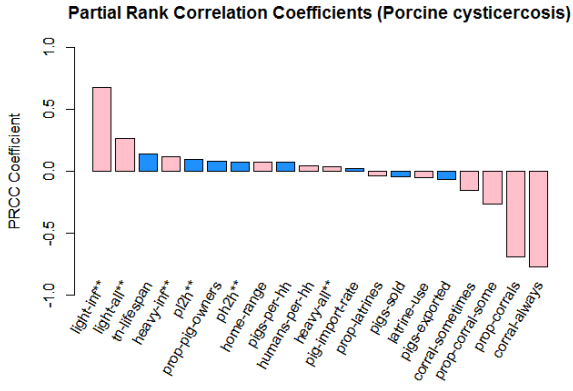
Latin hypercube sampling partial rank correlation coefficients (LHS-PRCC)

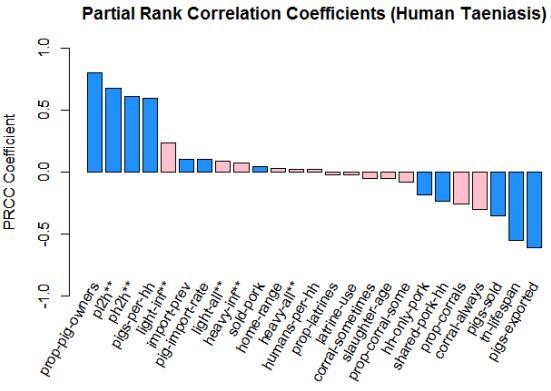
- Low-density village



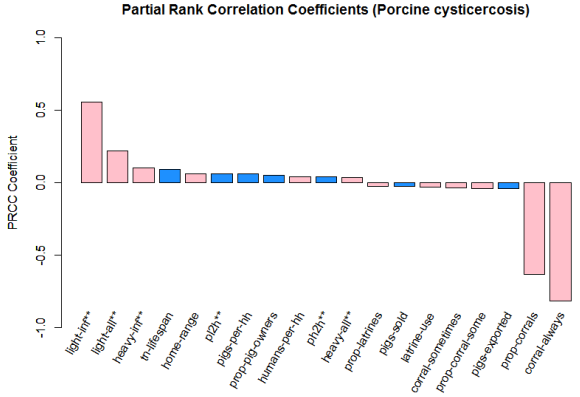


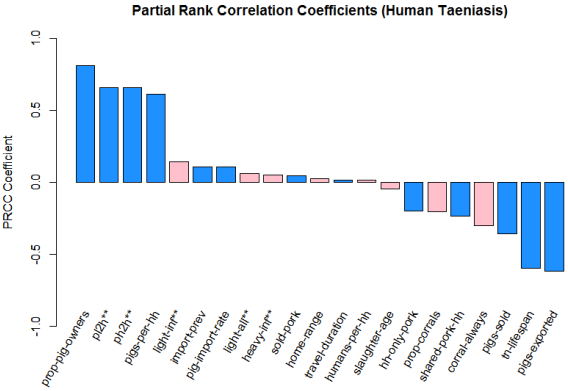
- Medium-density village





- High-density village

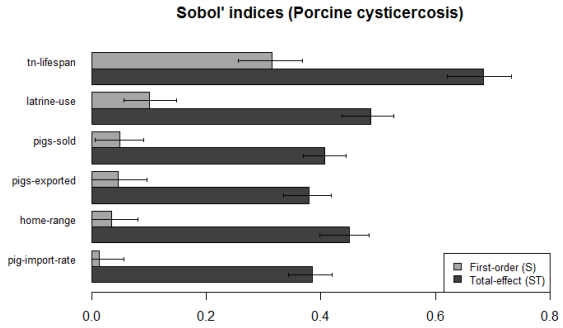


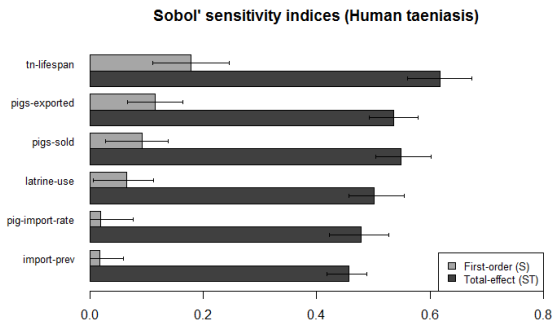


REDUCED ANALYSIS (k = 22 parameters)

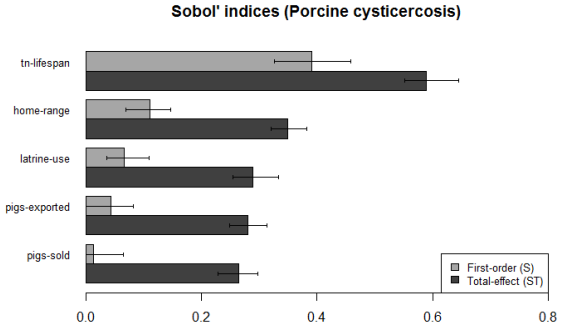
Sobol' sensitivity analyses indices

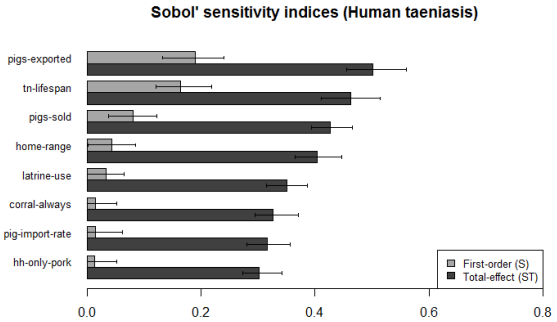
- Low-density village



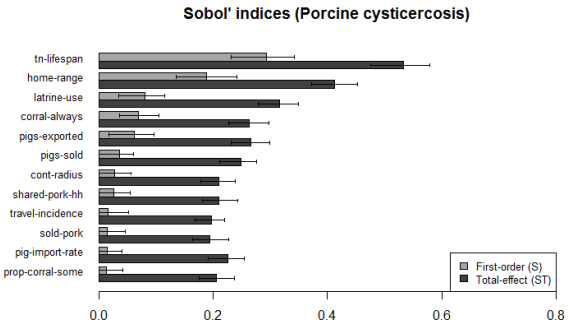


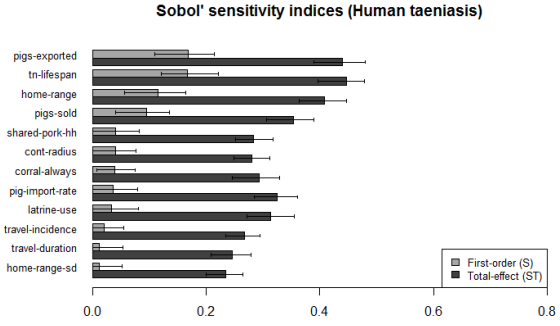
- Medium-density village





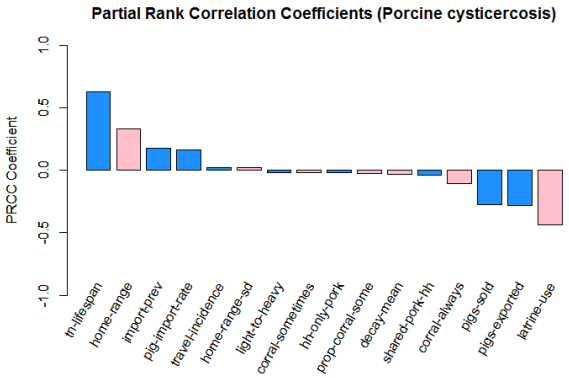
- High-density village

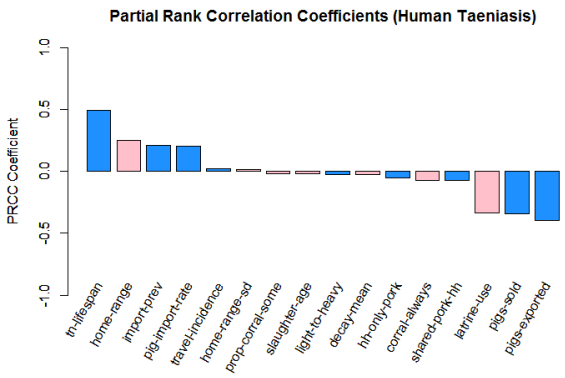




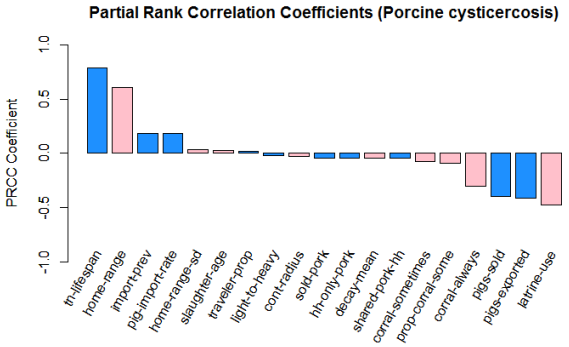
Latin hypercube sampling partial rank correlation coefficients (LHS-PRCC)

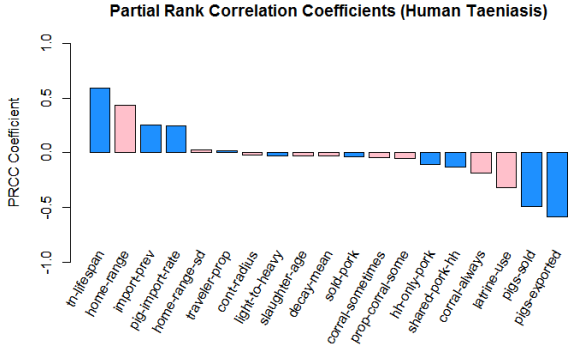
- Low-density village



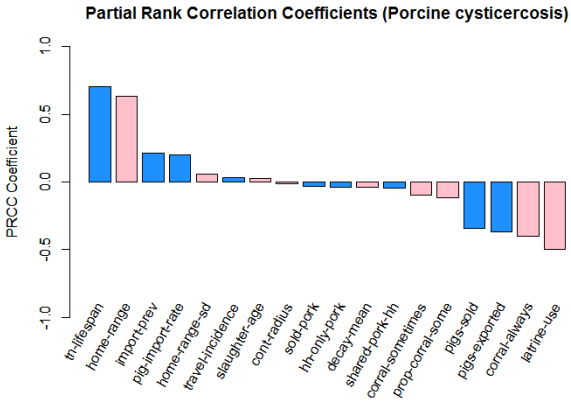


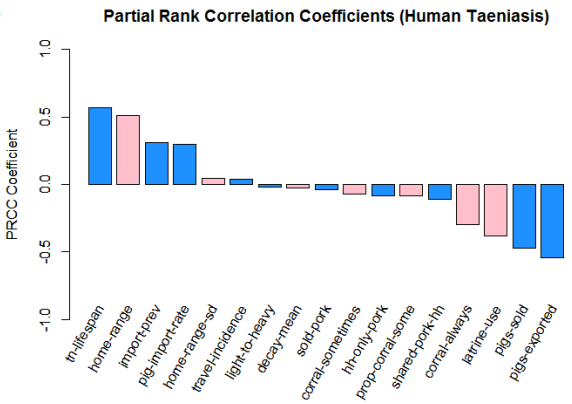
- Medium-density village





- High-density village





Appendix C: Supplemental material, tables, and figures for Chapter 5

Supplement C1. Model settings for simulations during CystiAgent validation.

Participation in interventions

When available, participation levels for humans and pigs were set separately for each village and each intervention activity (Table 3). Village-specific participation levels for humans included the proportion of eligible humans that participated in stool screening and presumptive treatment (NSM, 1 vs. 2 doses). Post-screening treatment positively identified humans was fixed at 91.8%, the average observed across all villages. For pigs, village-specific participation levels defined the proportion of eligible pigs receiving anti-helminthic treatment (OFZ), and the proportion participating in the first and second round of vaccination (TSOL18) in applicable villages. Pigs were eligible to participate in treatment, vaccine, and/or tongue-screening if they were ≥ 10 weeks old. For repeated interventions, participation levels observed in the field trials were averaged across all rounds for that village.

Drug efficacy and sensitivity of stool screening

A variety of other intervention settings were applied uniformly to all study villages. For both trials, the sensitivity of the CoAg-ELISA for detecting *T. solium* taeniasis was set to 96.4% [156], and the efficacy of NSM for treatment of human taeniasis was set at 76.6% for one dose, 86.6% for two doses, and 93.3% for post-screening follow-up. These values were based on results from the screening arms of the Ring Strategy interventions, and are generally in agreement with prior reports of NSM efficacy [93]. Treatment of pigs with OFZ was assumed to have an efficacy

of 100% [170], and render cysts non-viable within 1 week [34,170]. For infected pigs, treatment with OXF conferred protection against future infections for a period of 18 weeks [97].

Other intervention-specific settings

Additional settings required specifically for the Ring Strategy Trial included participation of pigs in tongue screening (77% applied uniformly to all villages and rounds), and the sensitivity and false-positive rate of tongue screening for detecting heavy cyst infection in pigs (90.9% and 2.1%, respectively) [43]. Human participation in the final round mass treatment/screening applied in Ring Strategy was set to 73.6%, the average observed across all villages. For the CEDP trial, the efficacy of the TSOL-18 pig vaccine was set to 99% for pigs receiving two doses [99,100]. Pigs receiving only one dose of TSOL-18 were not assumed to receive any protective benefit, as there is limited immunological evidence for single-dose vaccine protection [161]. Lastly, participation of pigs in the final round and post-intervention necroscopic examinations in CEDP were set to 48.2% and 38.4% of seropositive pigs. With the notable exceptions of pigs < 10 weeks old and humans traveling at the time of intervention applications, participation in interventions was randomly applied to members of the eligible population.

Methods used to compare model outputs with observed data

For each village evaluated, we compared the model-predicted prevalence of human taeniasis and porcine cysticercosis with the prevalence observed in the corresponding field trial. Given that the CystiAgent model represents actual infection, and field studies are limited by imperfect diagnostics, we made a variety of adjustments to the observed field statistics in order to compare them with model predictions.

Ring Strategy Trial. The prevalence of human taeniasis was directly measured at the conclusion of the study; however the baseline prevalence was not measured. Therefore, we predicted the baseline prevalence of human taeniasis using a regression equation developed from

a prior cross-sectional study of human and pig prevalence in the region [75]. For this, seroprevalence of (2+ EITB bands) in pigs was used a predictor for the log-prevalence of human taeniasis, and a 75% prediction interval was generated for each Ring Strategy village (Fig C1-1). Because necroscopic examination of pigs was not performed in this study, the prevalence of cyst infection at baseline and study-end was also estimated based on pig seroprevalence. For both measures, we estimated that 30.1% of seropositive pigs (2+ EITB bands) would have light cyst infection (<100 cysts), and 12.8% of seropositive pigs would have heavy cyst infection (≥100 cysts). For these proportions, we averaged the results from two large necropsy studies conducted in Peru – one in the Piura region [43], and the other in the highland of Huancayo [162]. For serological outcomes, we compared the incidence of pig seroconversion (2+ EITB) measured at seven time-points throughout the study with model-predicted sero-incidence.

CEDP. The CEDP trial directly measured the prevalence of human taeniasis in all villages at months 2 and 5 of the trial, thus no statistical prediction was needed. For porcine cysticercosis, the baseline prevalence of cyst infection was estimated based on the seroprevalence as above. Due to volatility in the seroprevalence observed in this study, we based prevalence predictions on the average of the pre-baseline and baseline sero-surveys. For this study, we assumed that 19.5% and 6.5% of seropositive pigs (2+ EITB bands) would have light and heavy cyst infection, respectively. These lower proportions were based on available necropsy data from this region [122], and averaged with data from the Piura region [43]. For the prevalence of porcine cysticercosis at study-end and in the post-intervention follow-up (12 months later), we assumed that *T. solium* elimination was achieved in all villages. This was based on necroscopic examinations conducted in all villages that detected cyst infection in only 6/4019 (0.15%) pigs directly following the interventions and 7/3073 (0.23%) pigs in the one-year follow-up [94]. The serological outcomes compared throughout the intervention comprised of seroprevalence (2+ EITB bands) across eight time-points.

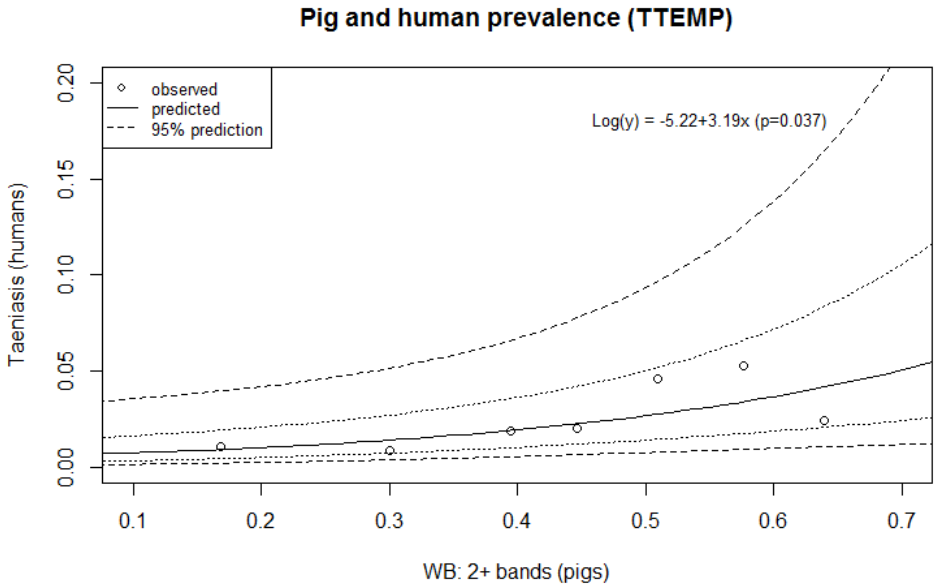
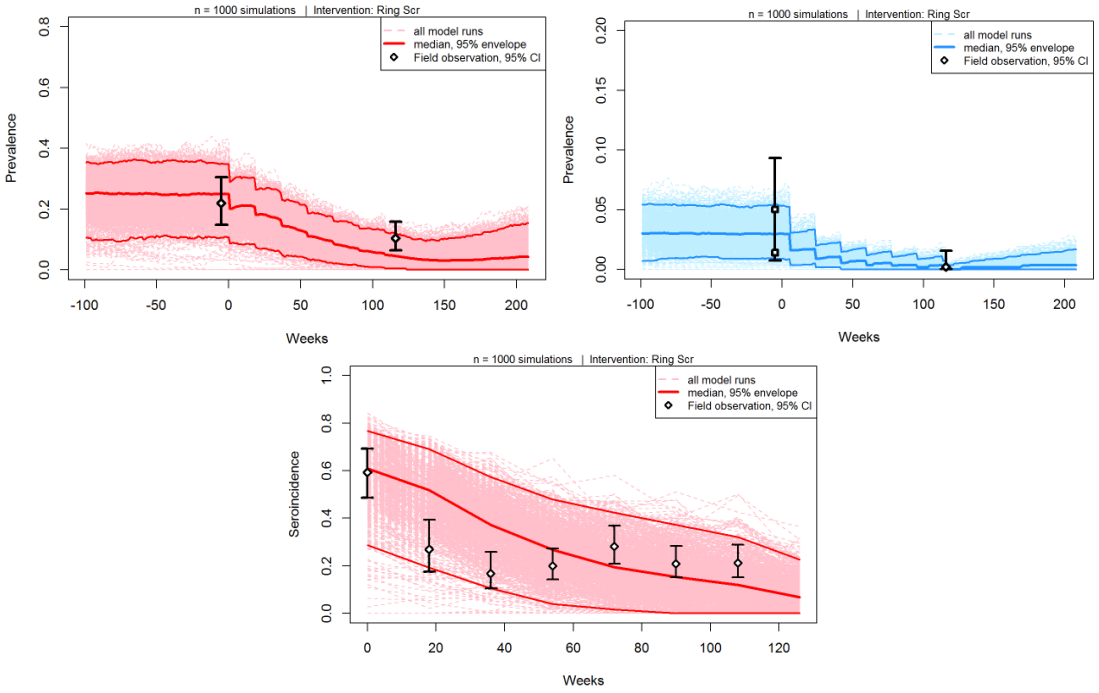


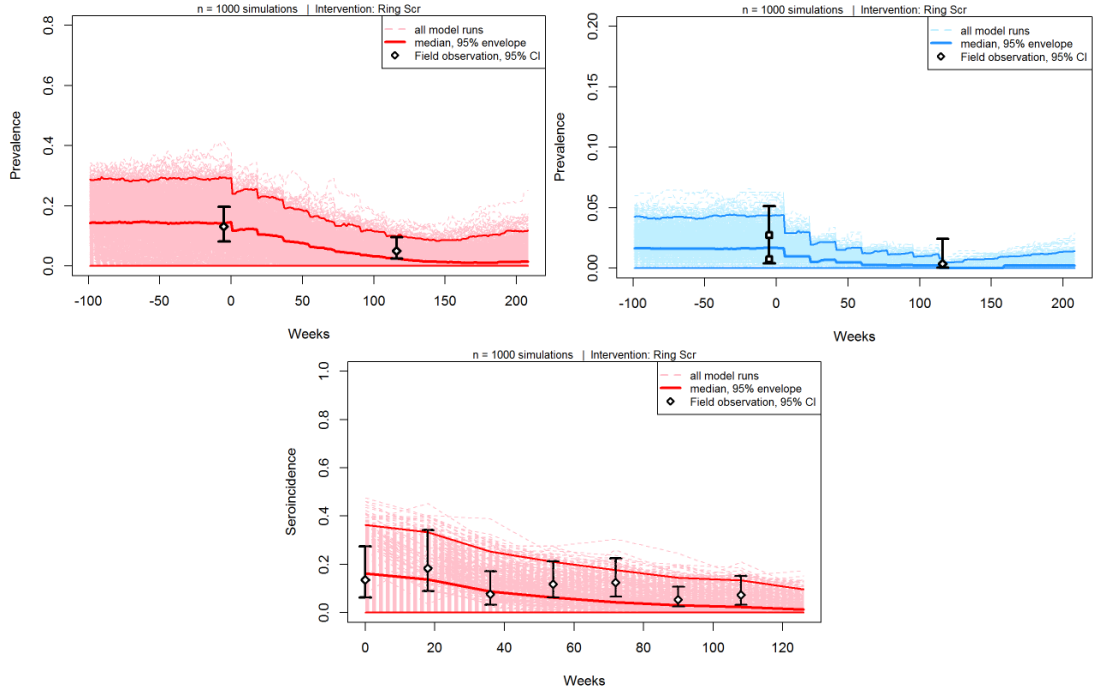
Fig C1-1. Regression model used to predict baseline prevalence of human taeniasis from pig seroprevalence. Data were extracted from a cross-sectional survey of seven villages in the Piura region of northern Peru [75].

Fig C2. Full validation results from Ring Strategy Trial (n=21 villages). Left: Porcine cysticercosis; Right: Human taeniasis; Lower: Porcine seroincidence

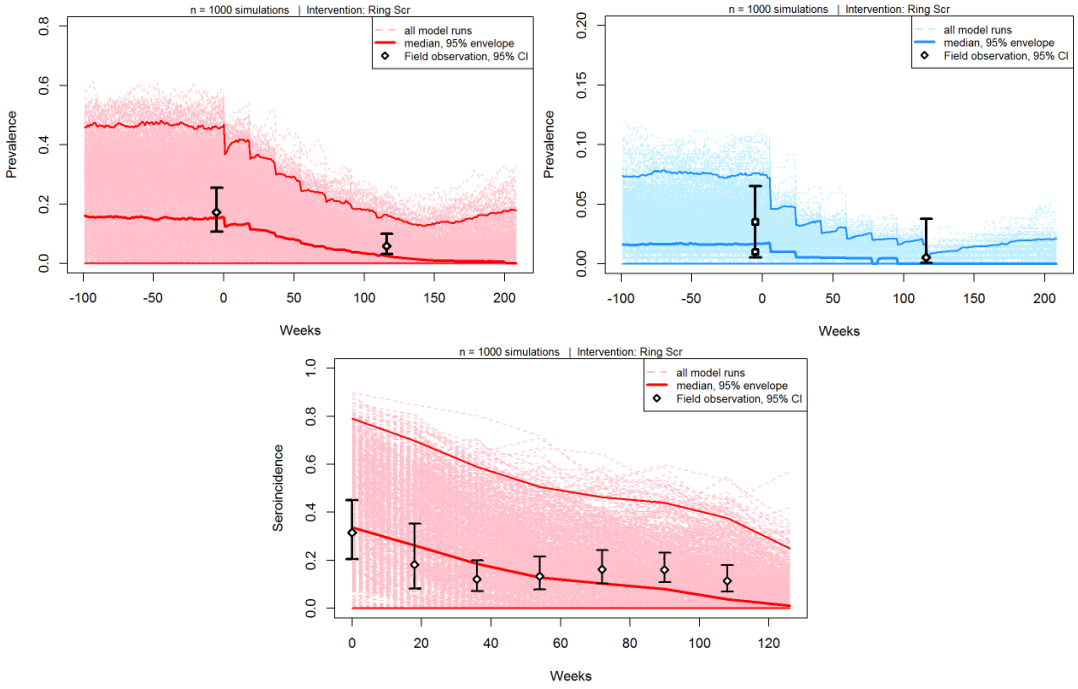
Ring Screening w/o pig treatment (RST, village 1 of 4)



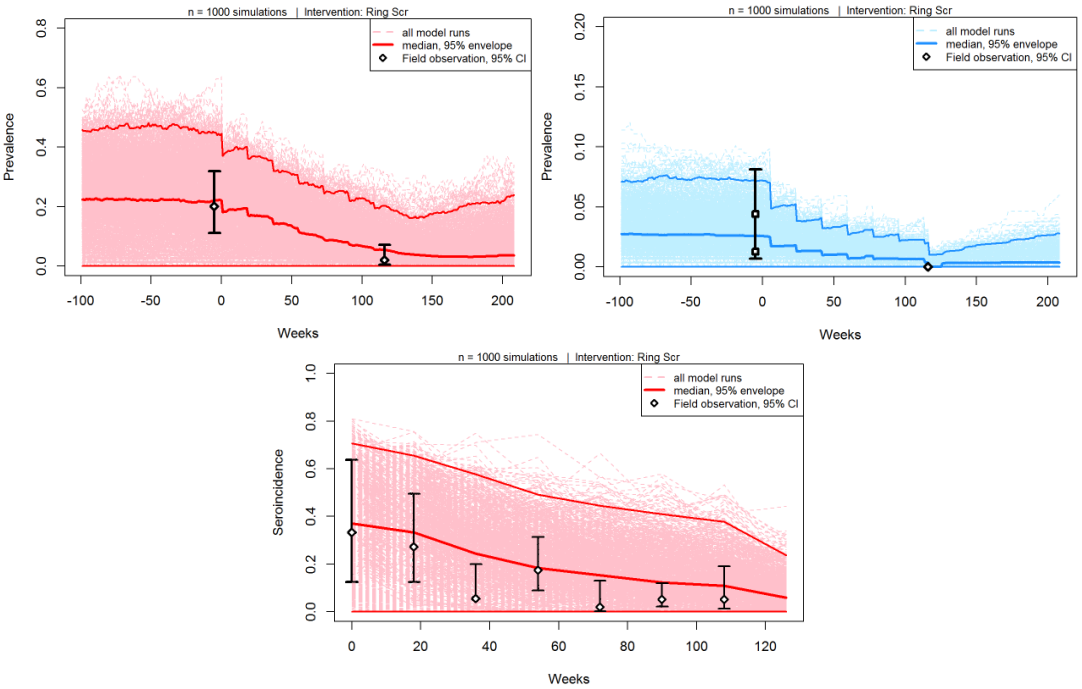
Ring Screening w/o pig treatment (RST, village 2 of 4)



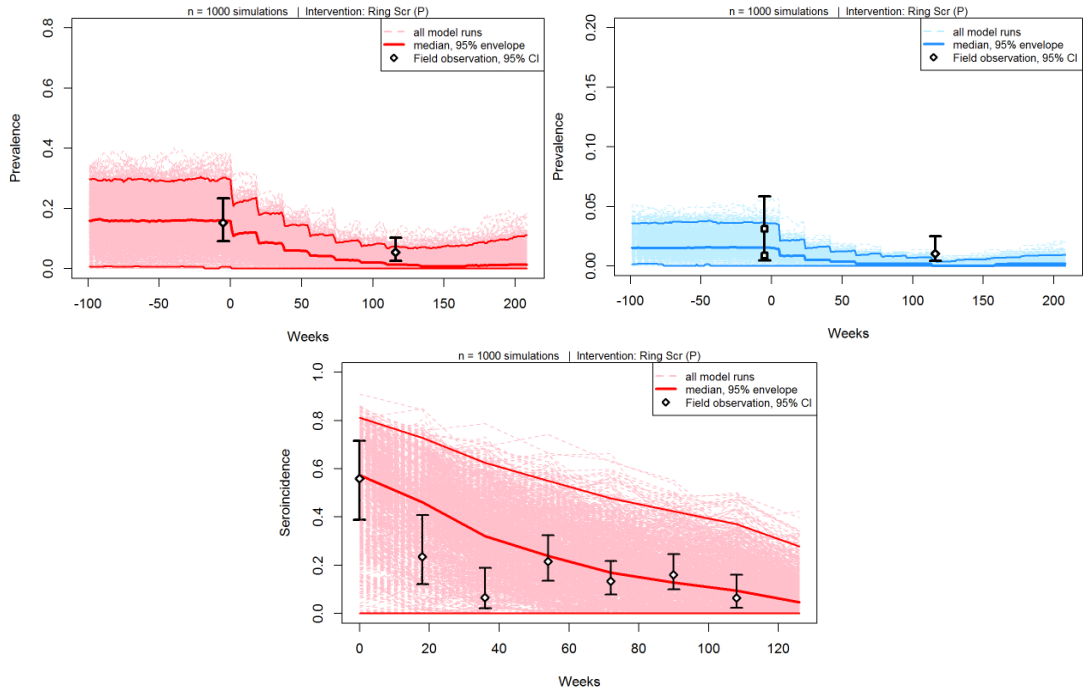
Ring Screening w/o pig treatment (RST, village 3 of 4)



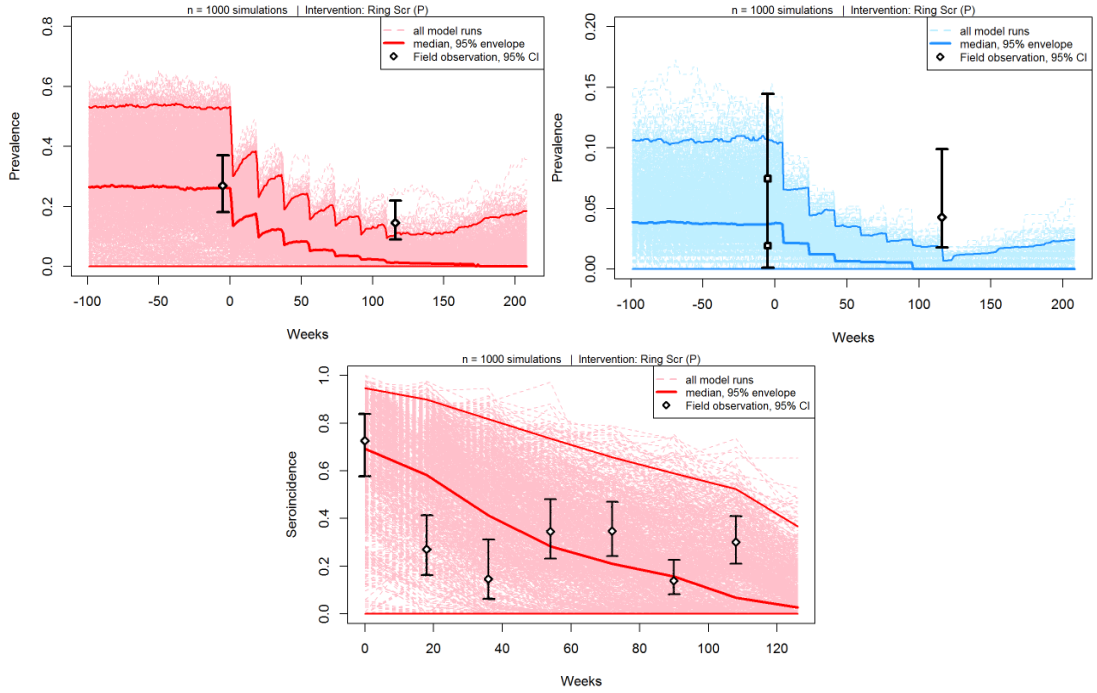
Ring Screening w/o pig treatment (RST, village 4 of 4)



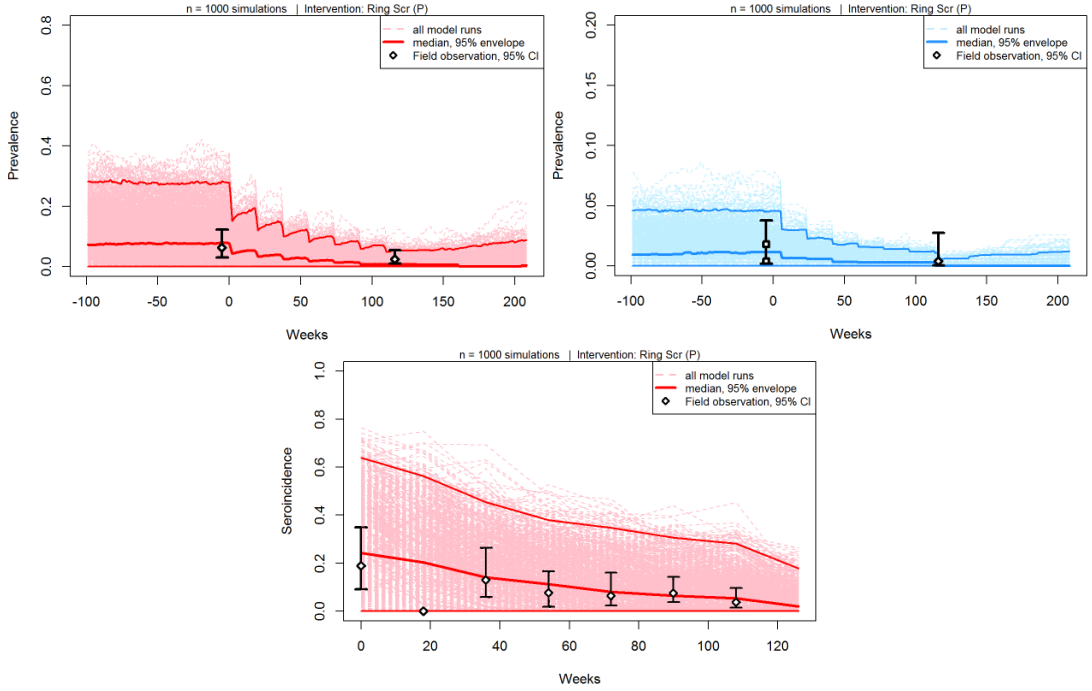
Ring Screening w/ pig treatment (RST, village 1 of 4)



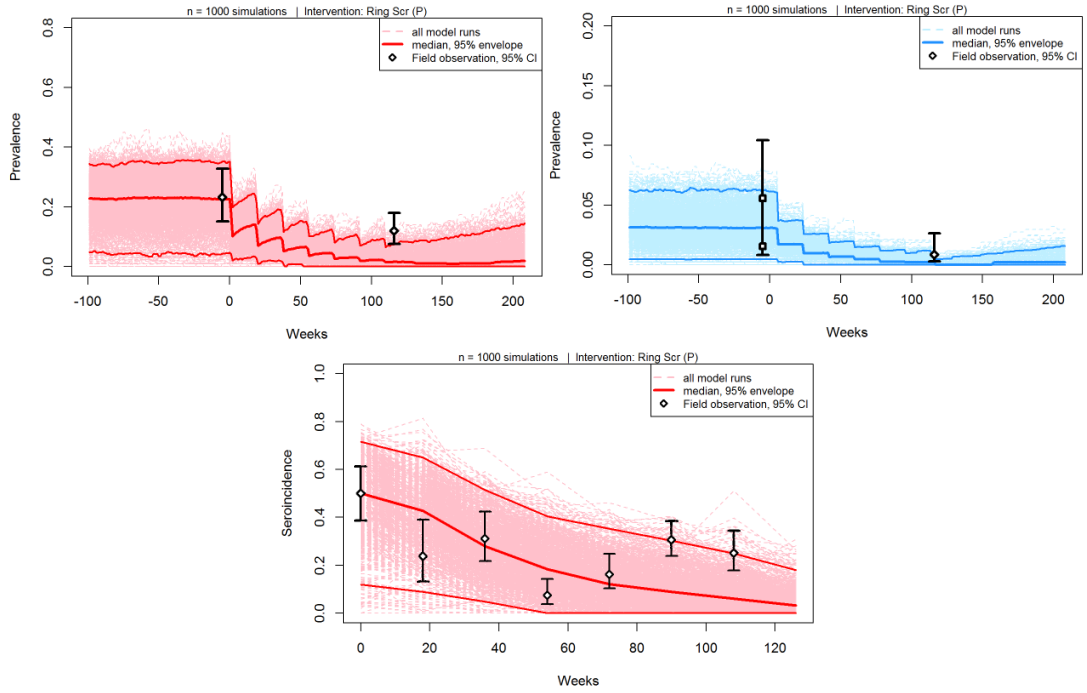
Ring Screening w/ pig treatment (RST, village 2 of 4)



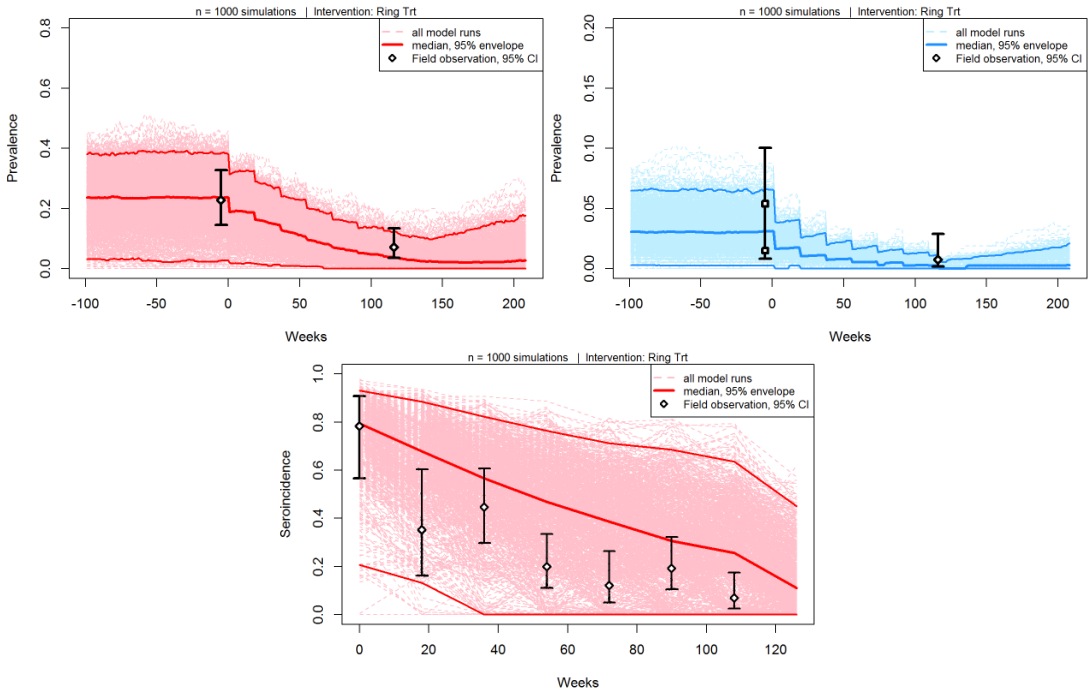
Ring Screening w/ pig treatment (RST, village 3 of 4)



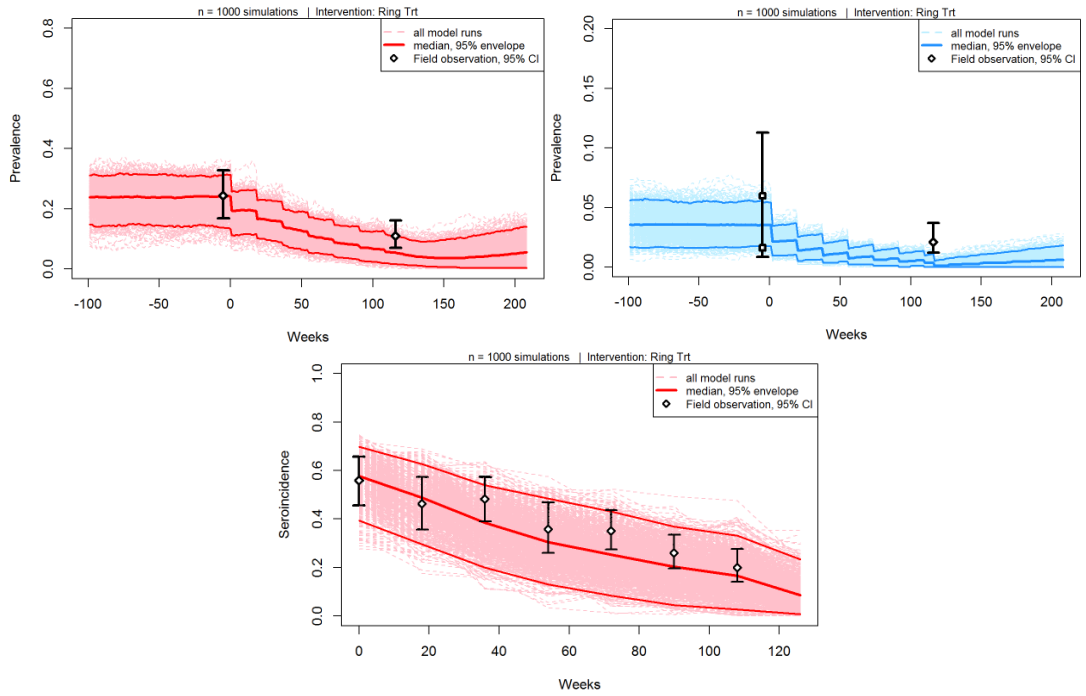
Ring Screening w/ pig treatment (RST, village 4 of 4)



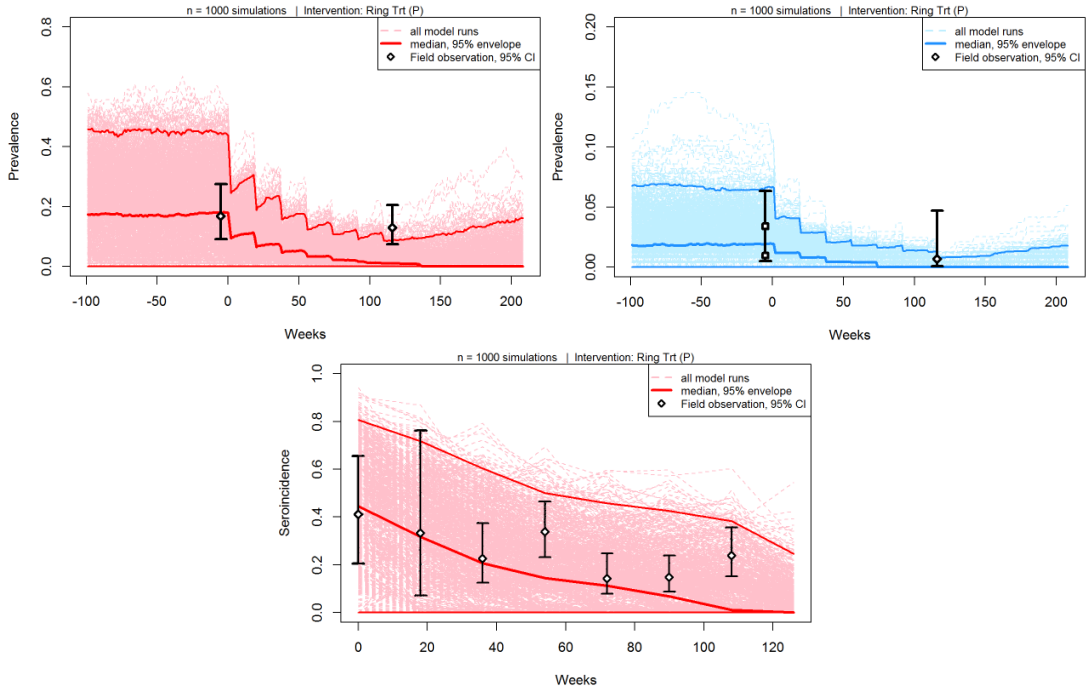
Ring Treatment w/o pig treatment (RST, village 1 of 2)



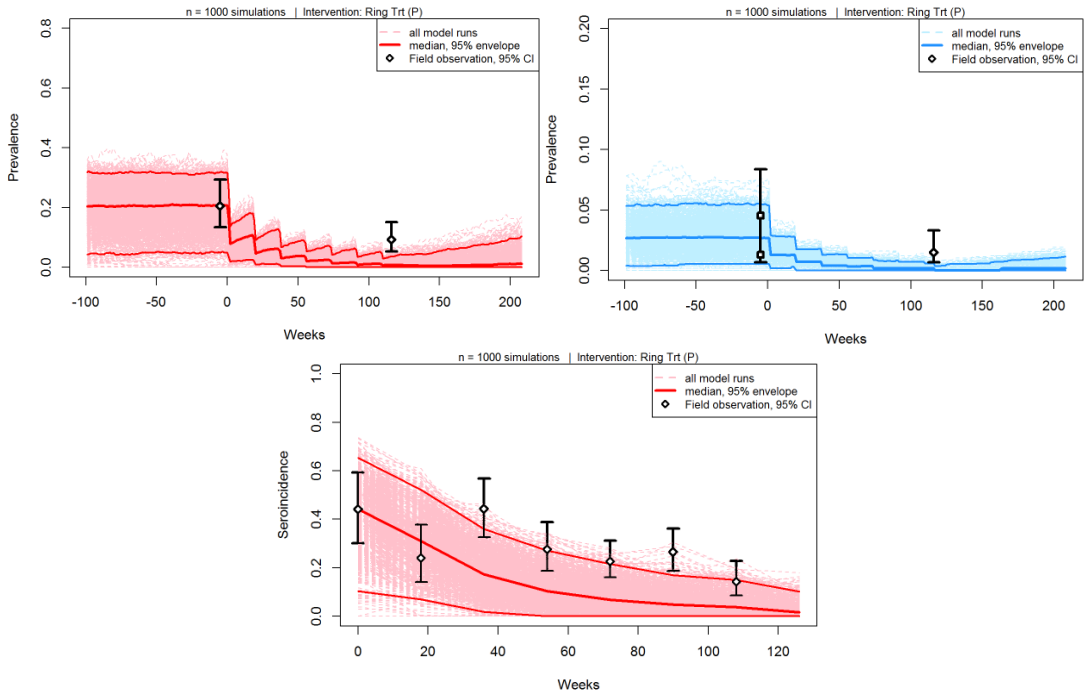
Ring Treatment w/o pig treatment (RST, village 2 of 2)



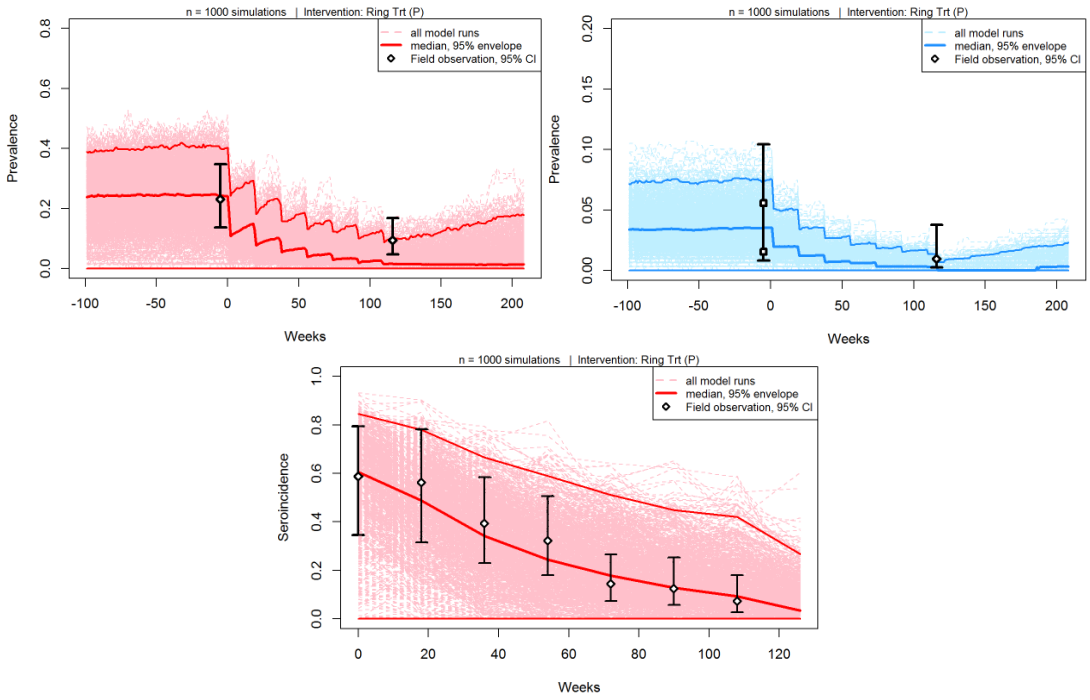
Ring Treatment w/ pig treatment (RST, village 1 of 4)



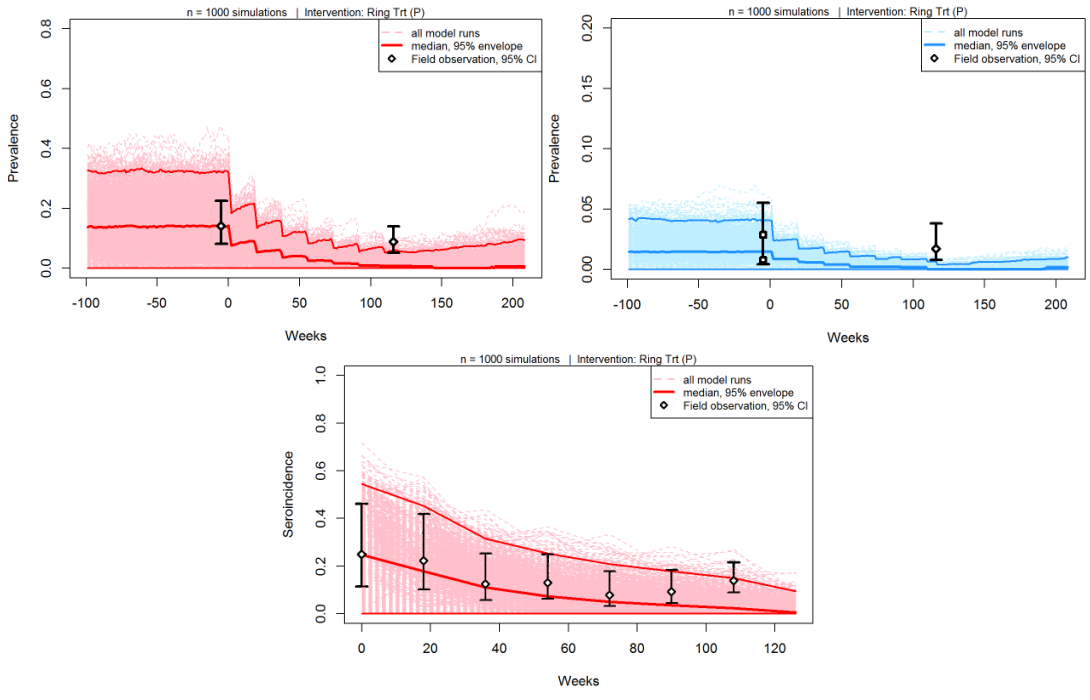
Ring Treatment w/ pig treatment (RST, village 2 of 4)



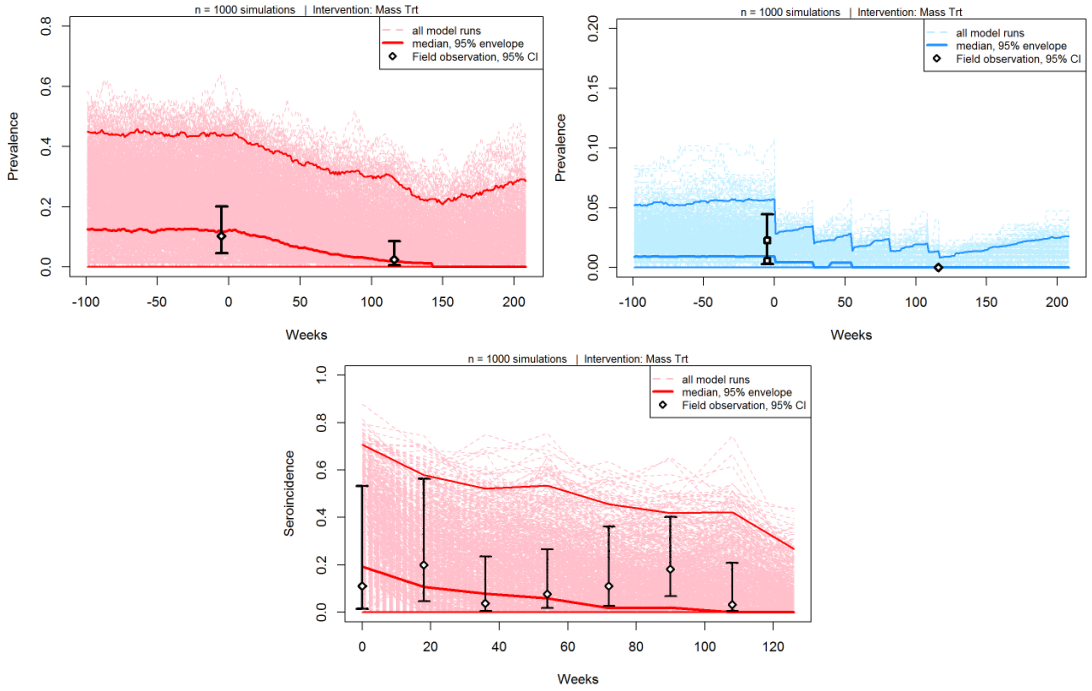
Ring Treatment w/ pig treatment (RST, village 3 of 4)



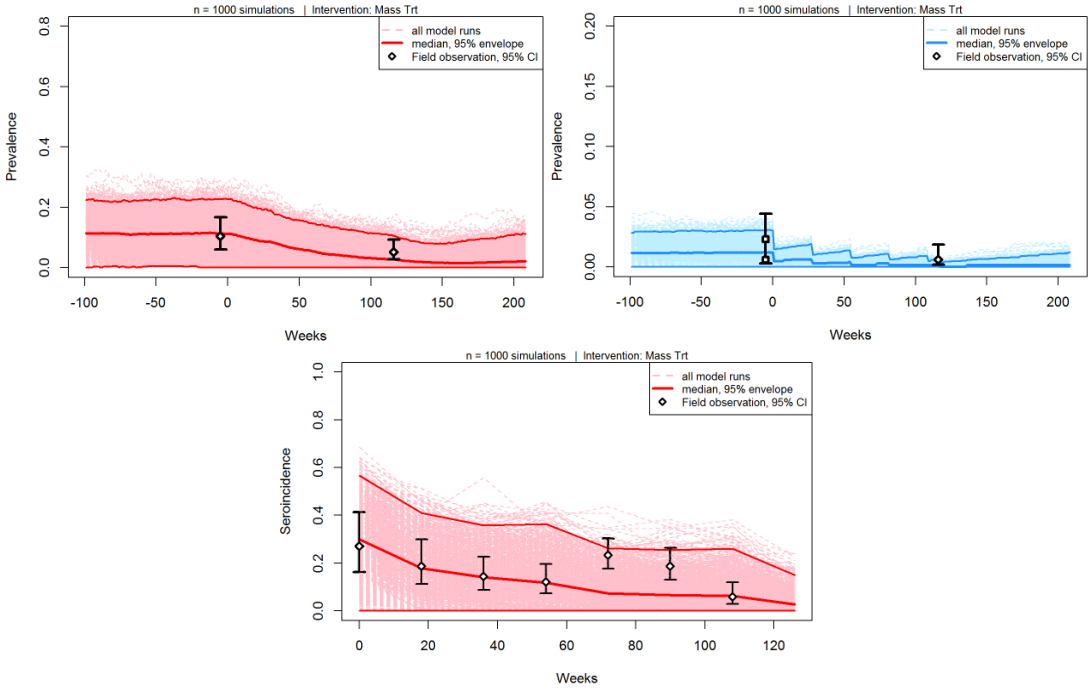
Ring Treatment w/ pig treatment (RST, village 4 of 4)



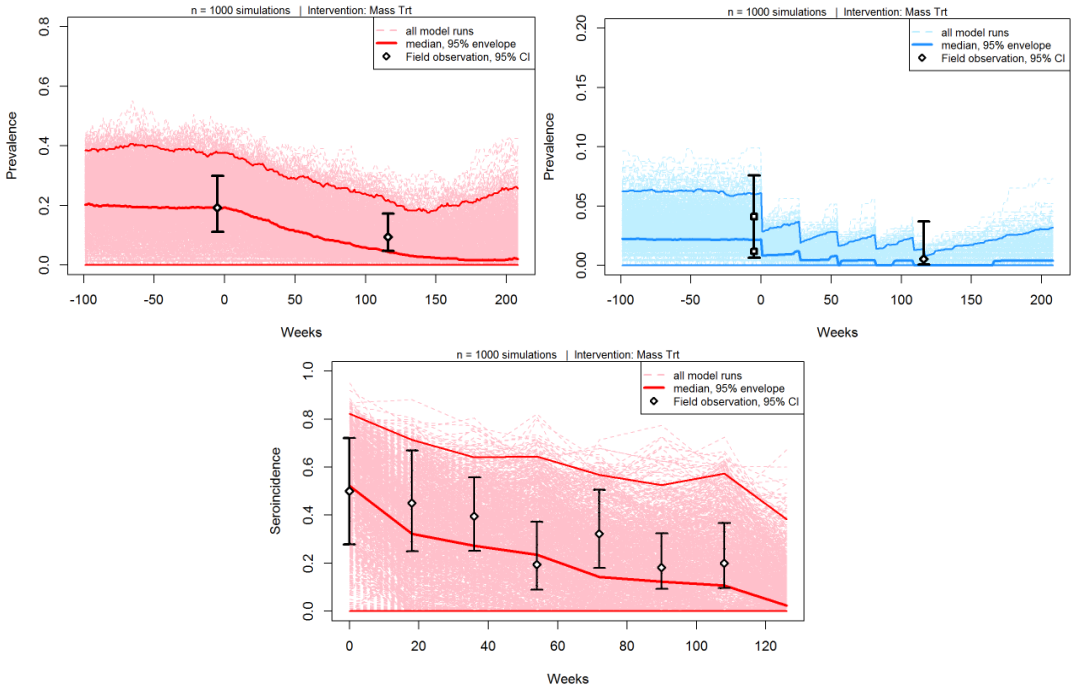
Mass Treatment w/o pig treatment (RST, village 1 of 4)



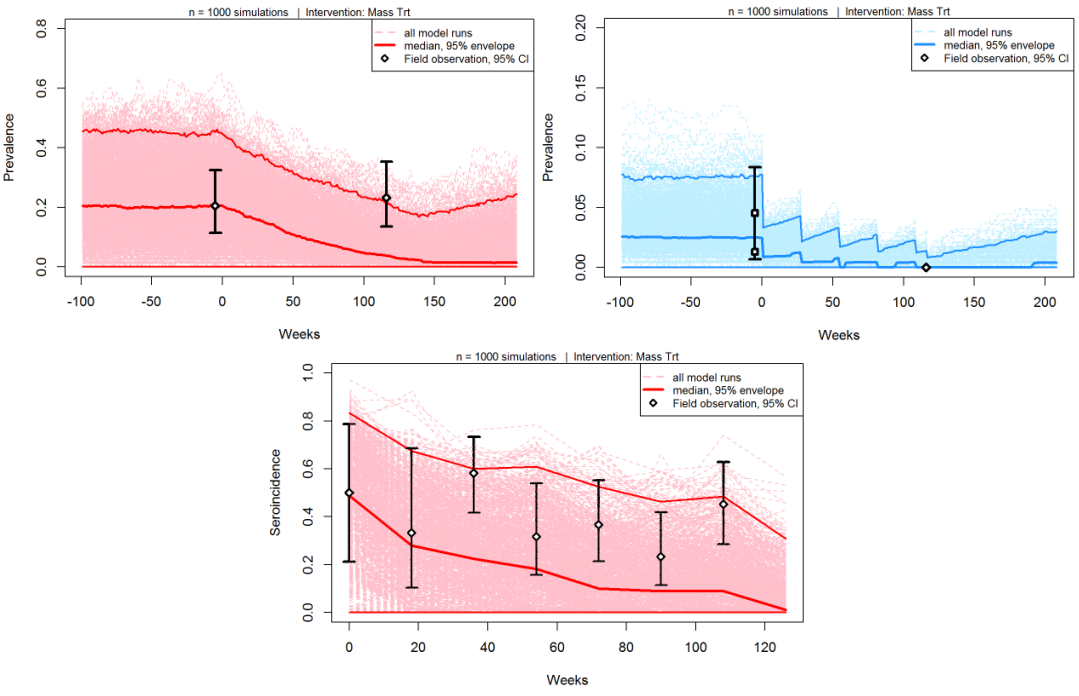
Mass Treatment w/o pig treatment (RST, village 2 of 4)



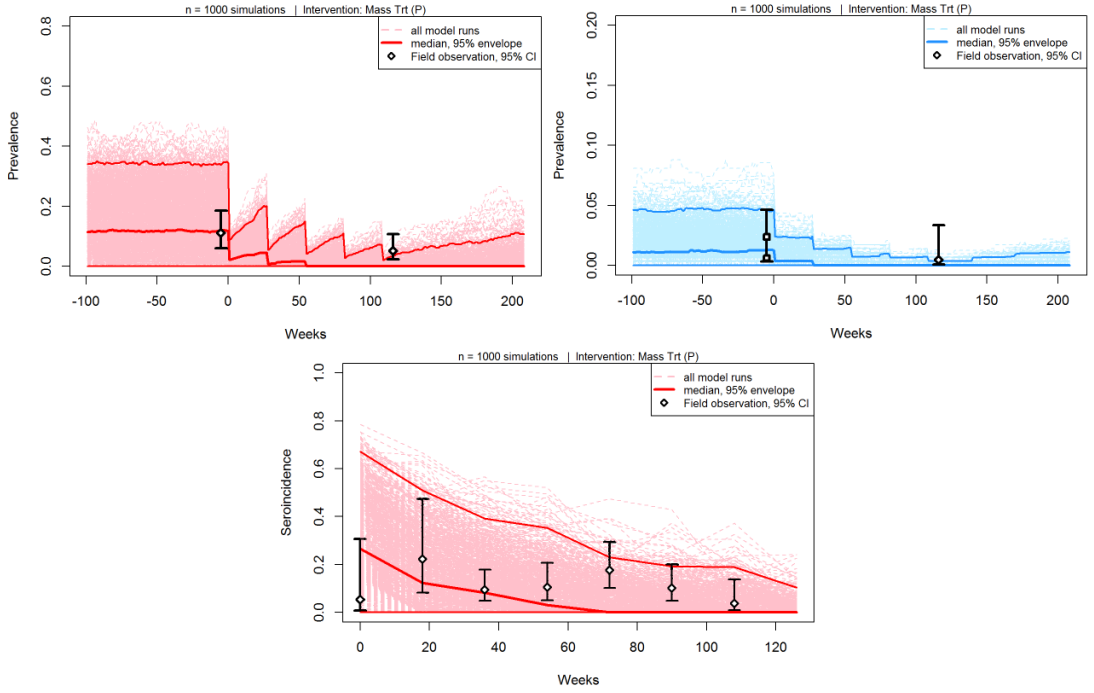
Mass Treatment w/o pig treatment (RST, village 3 of 4)



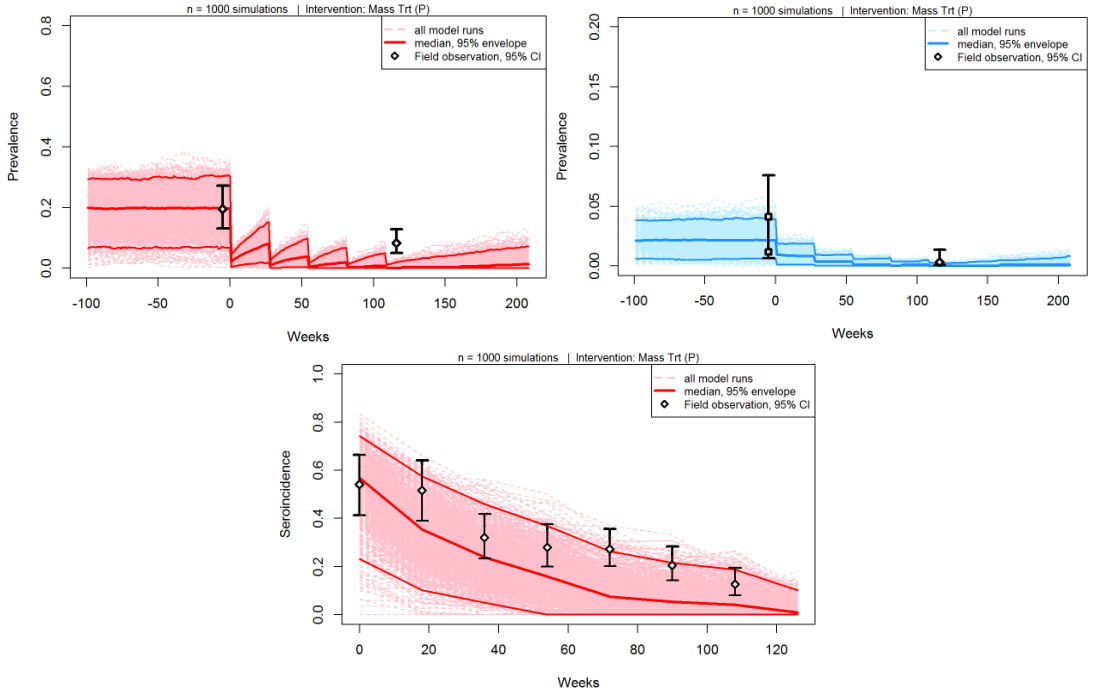
Mass Treatment w/o pig treatment (RST, village 4 of 4)



Mass Treatment w/ pig treatment (RST, village 1 of 3)



Mass Treatment w/ pig treatment (RST, village 2 of 3)



Mass Treatment w/ pig treatment (RST, village 3 of 3)

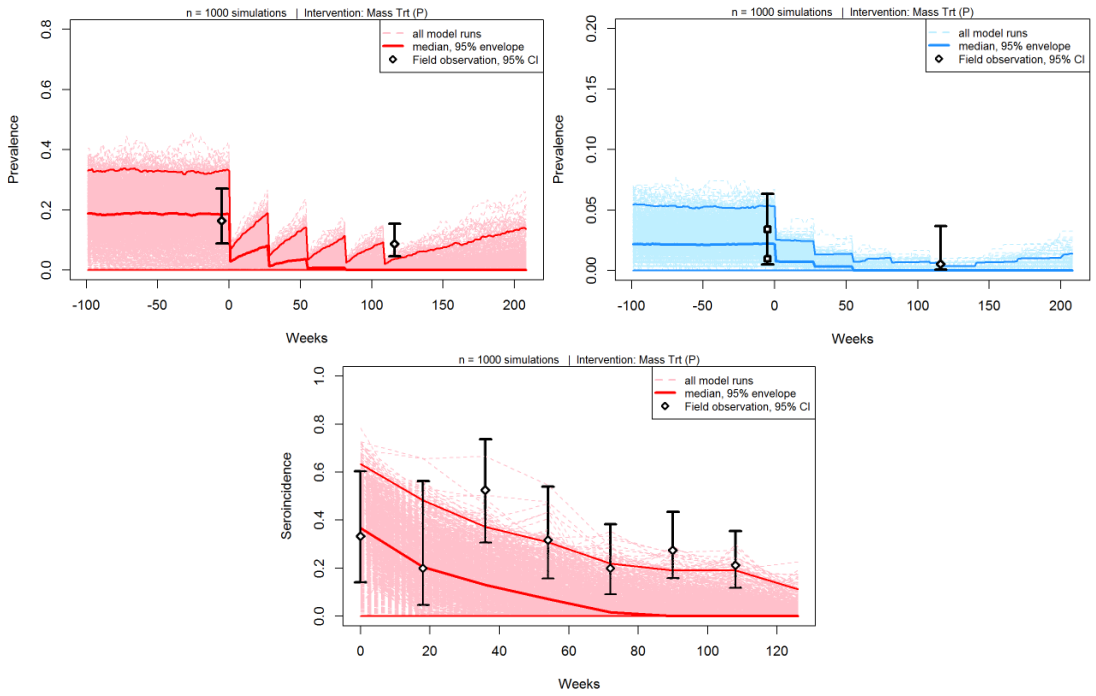
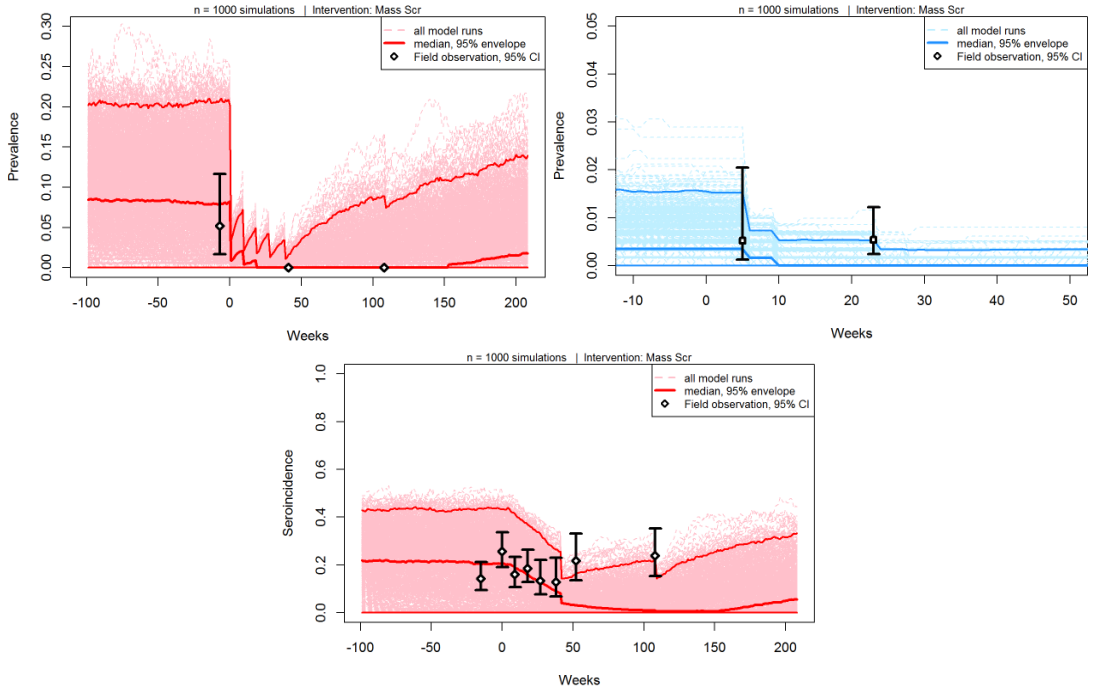
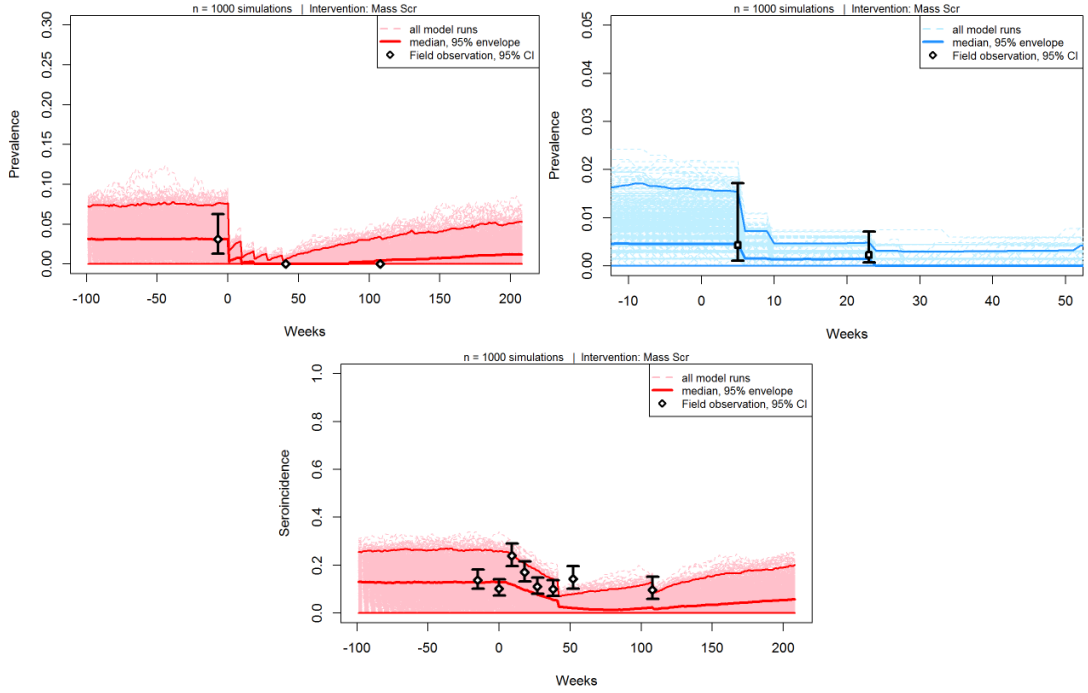


Fig C3. Full validation results from CEDP Trial (n=15 villages). Left: Porcine cysticercosis; Right: Human taeniasis; Lower: Porcine seroprevalence

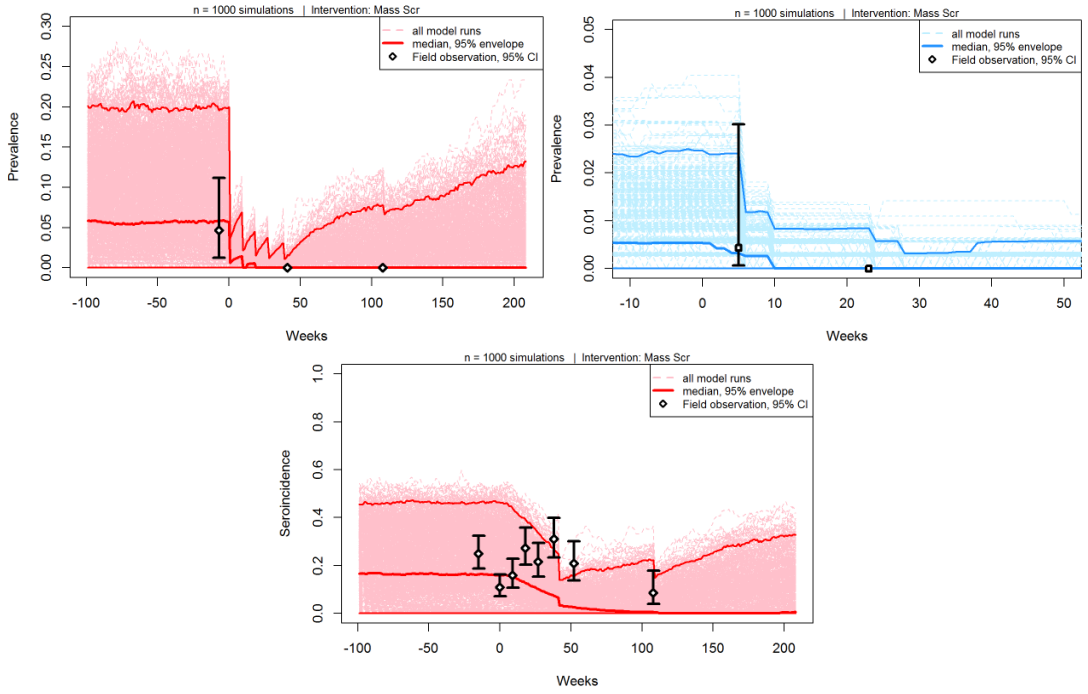
Mass Screening w/o vaccine (CEDP, village 1 of 4)



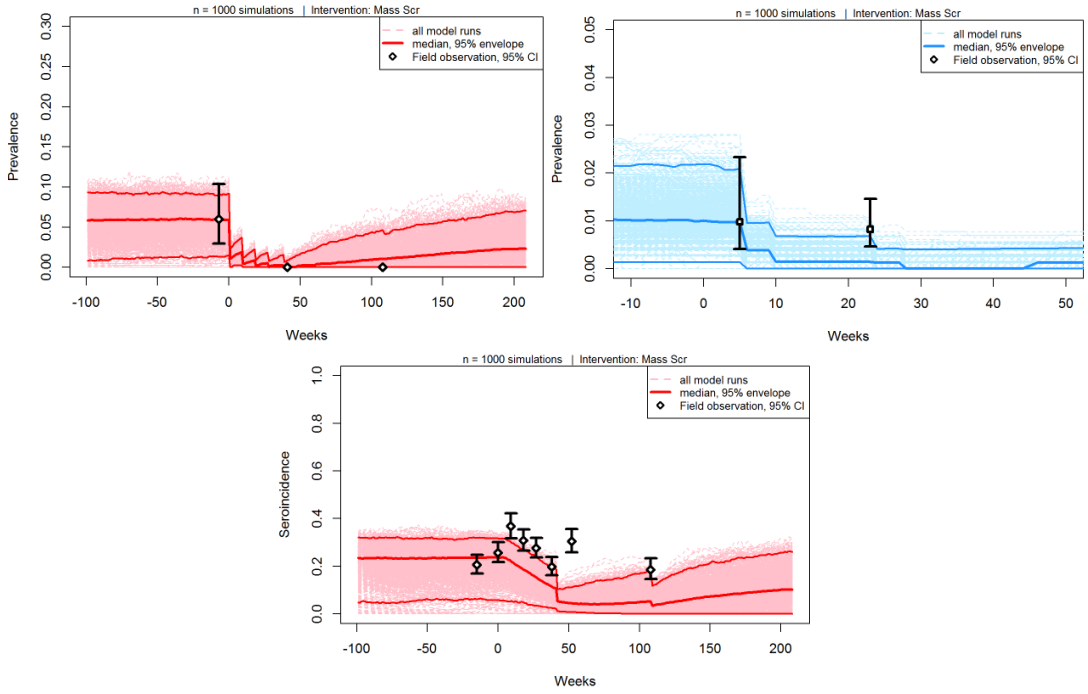
Mass Screening w/o vaccine (CEDP, village 2 of 4)



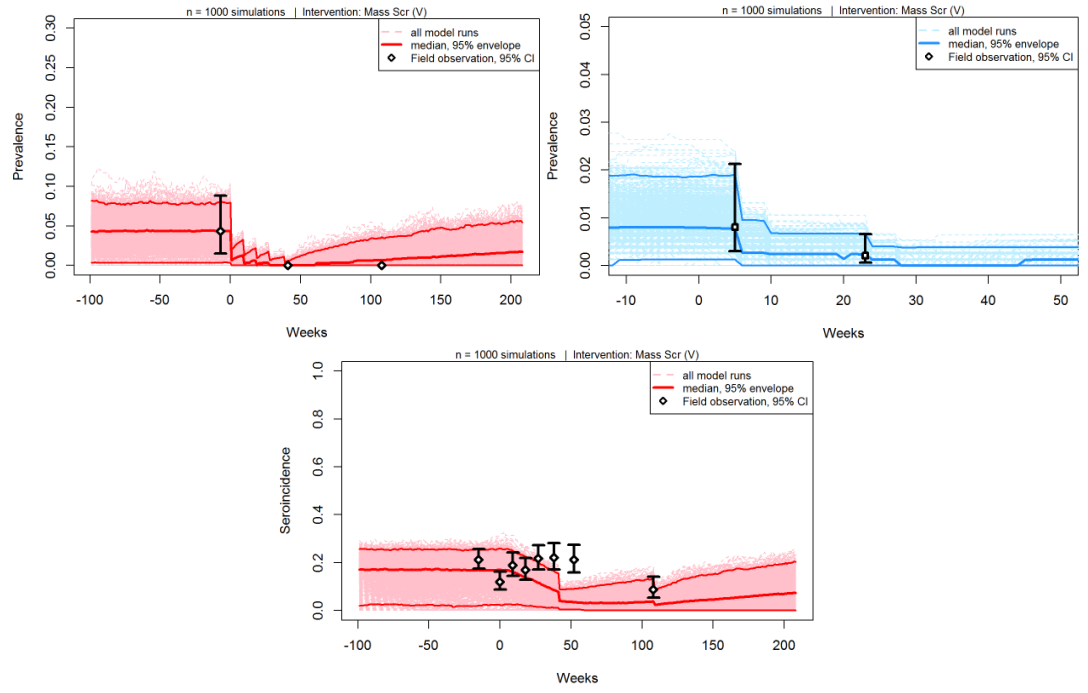
Mass Screening w/o vaccine (CEDP, village 3 of 4)



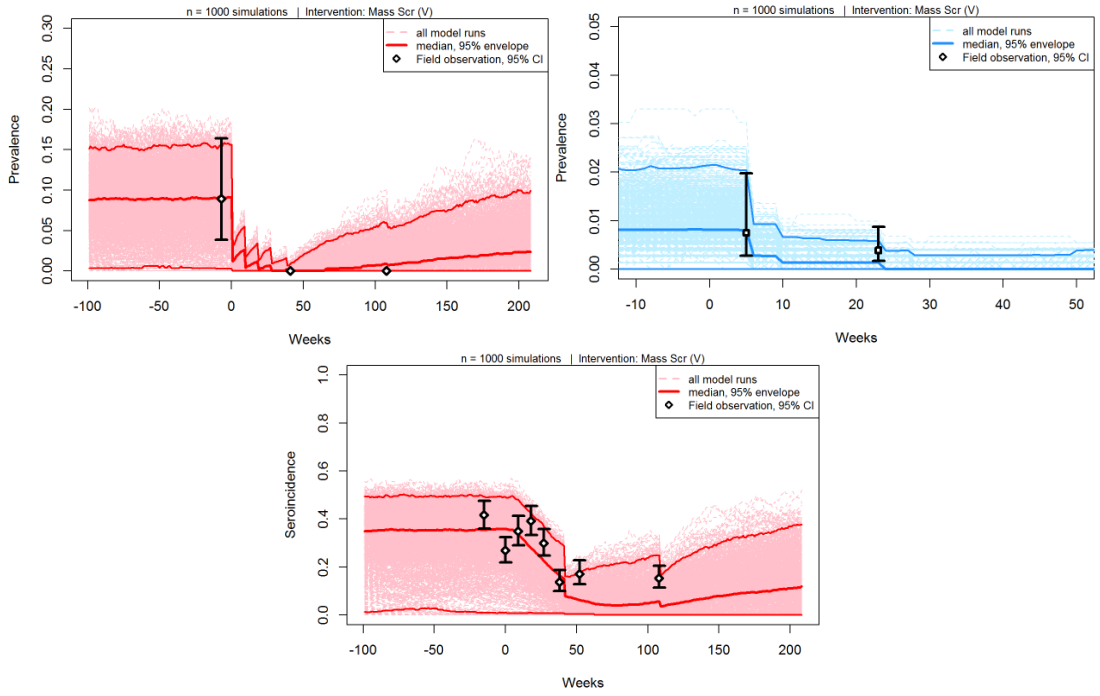
Mass Screening w/o vaccine (CEDP, village 4 of 4)



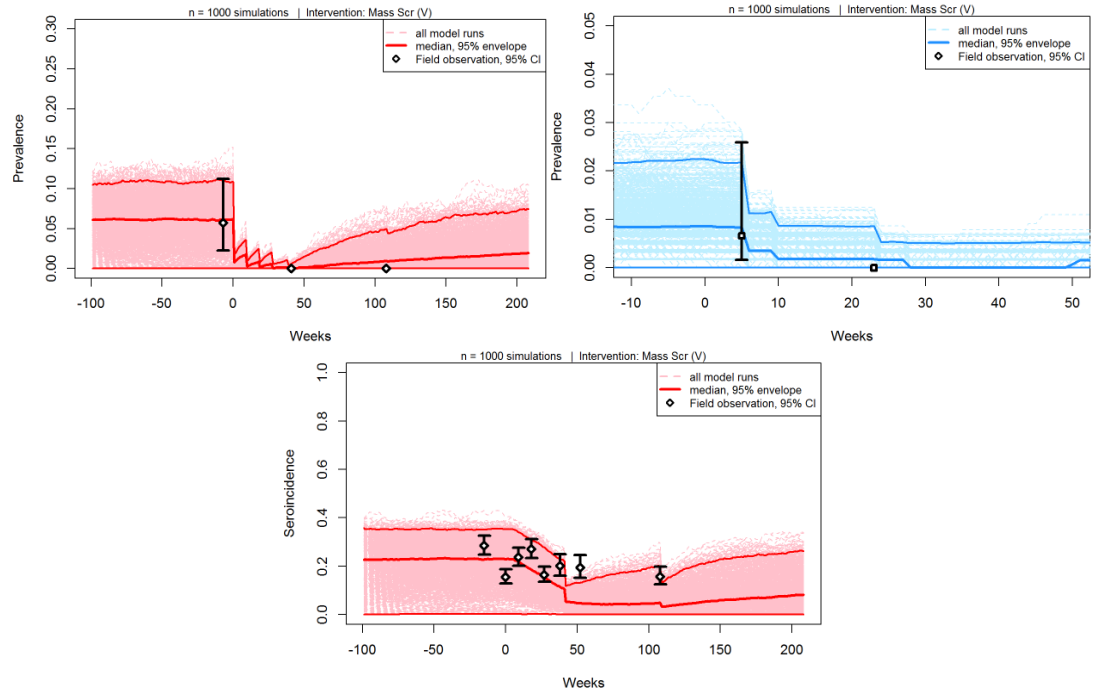
Mass Screening w/ vaccine (CEDP, village 1 of 3)



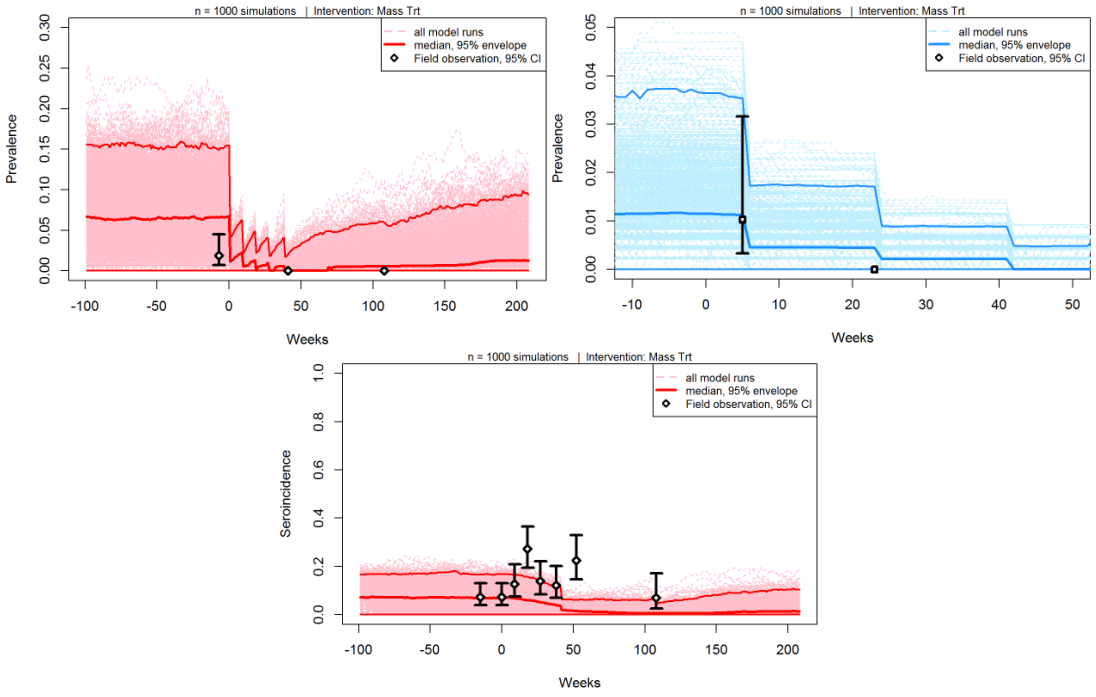
Mass Screening w/ vaccine (CEDP, village 2 of 3)



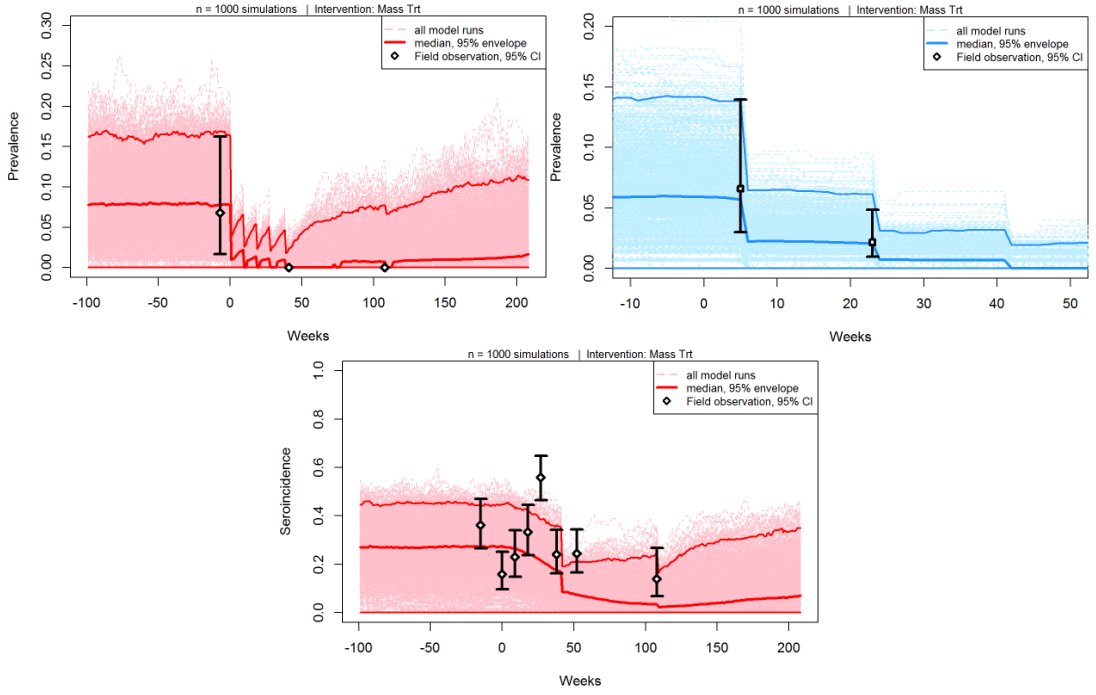
Mass Screening w/ vaccine (CEDP, village 3 of 3)



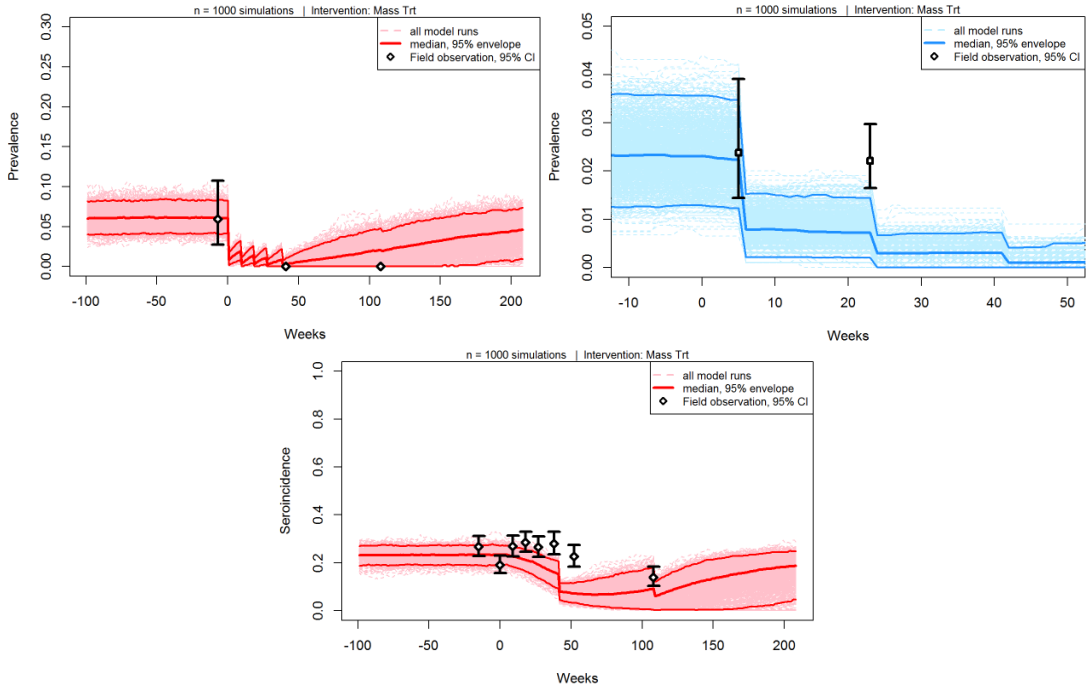
Mass Treatment w/o vaccine (CEDP, village 1 of 6)



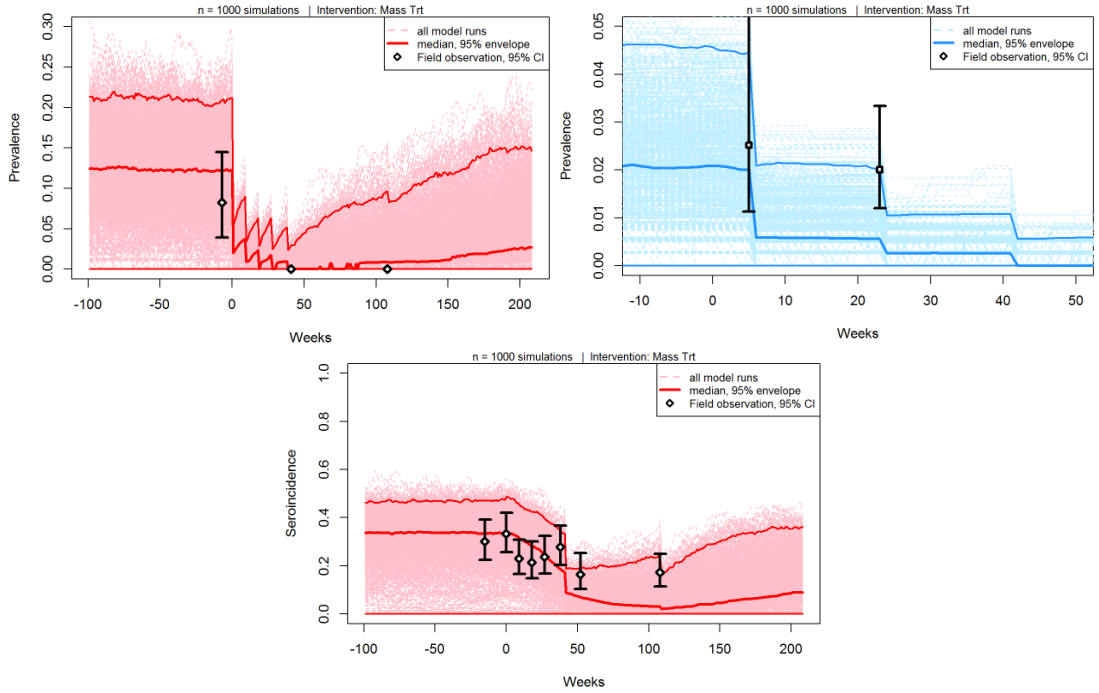
Mass Treatment w/o vaccine (CEDP, village 2 of 6)



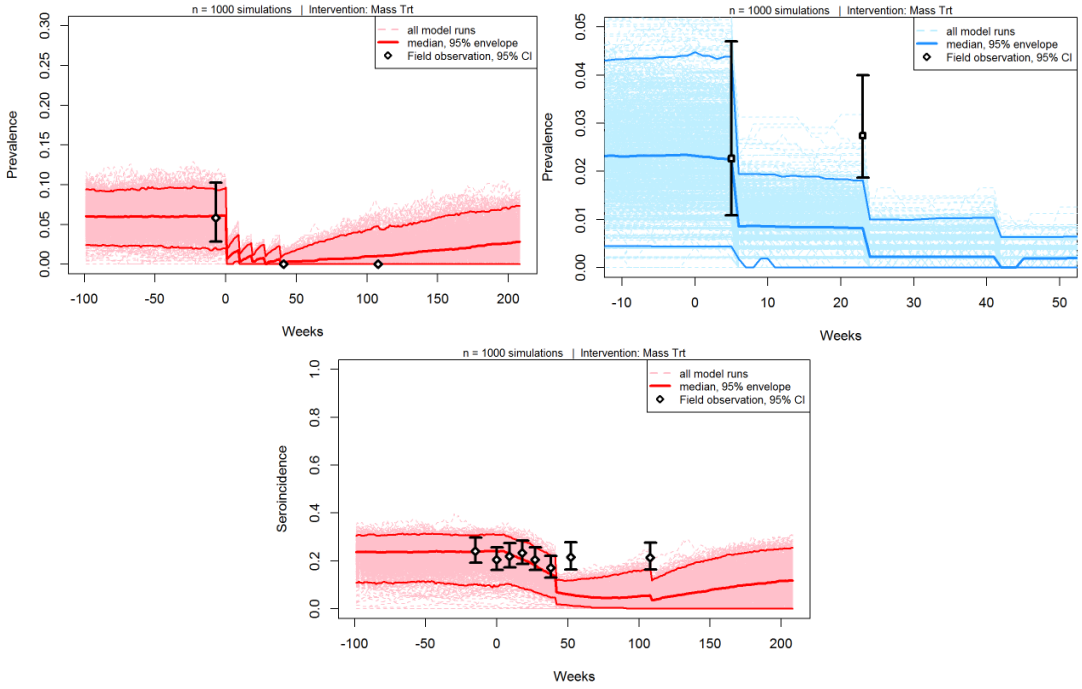
Mass Treatment w/o vaccine (CEDP, village 3 of 6)



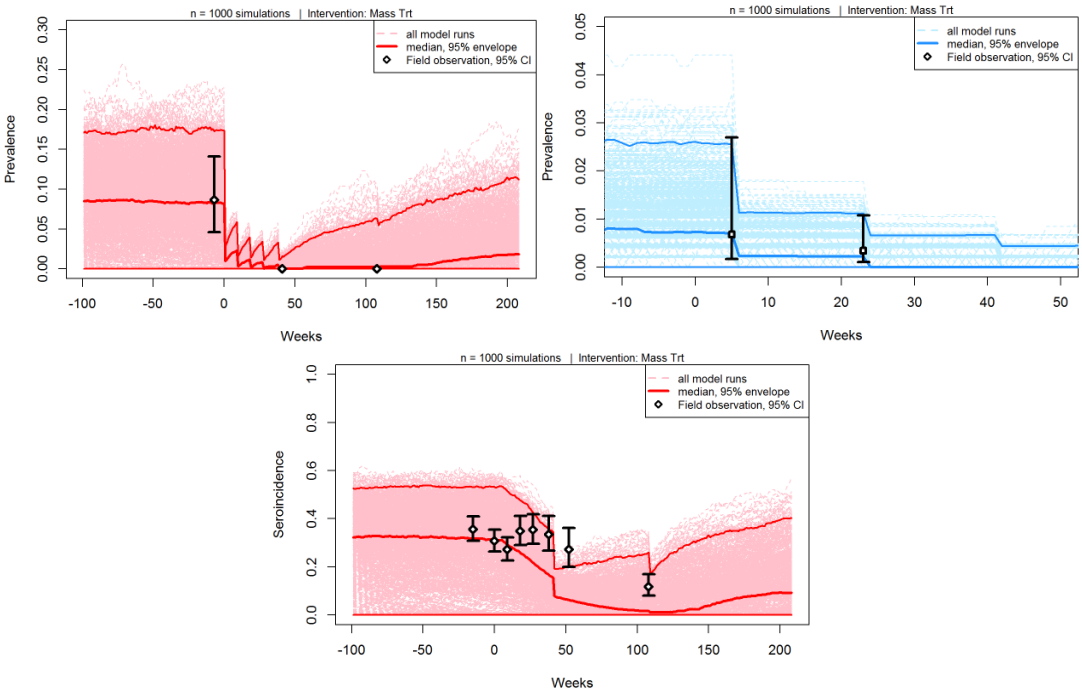
Mass Treatment w/o vaccine (CEDP, village 4 of 6)



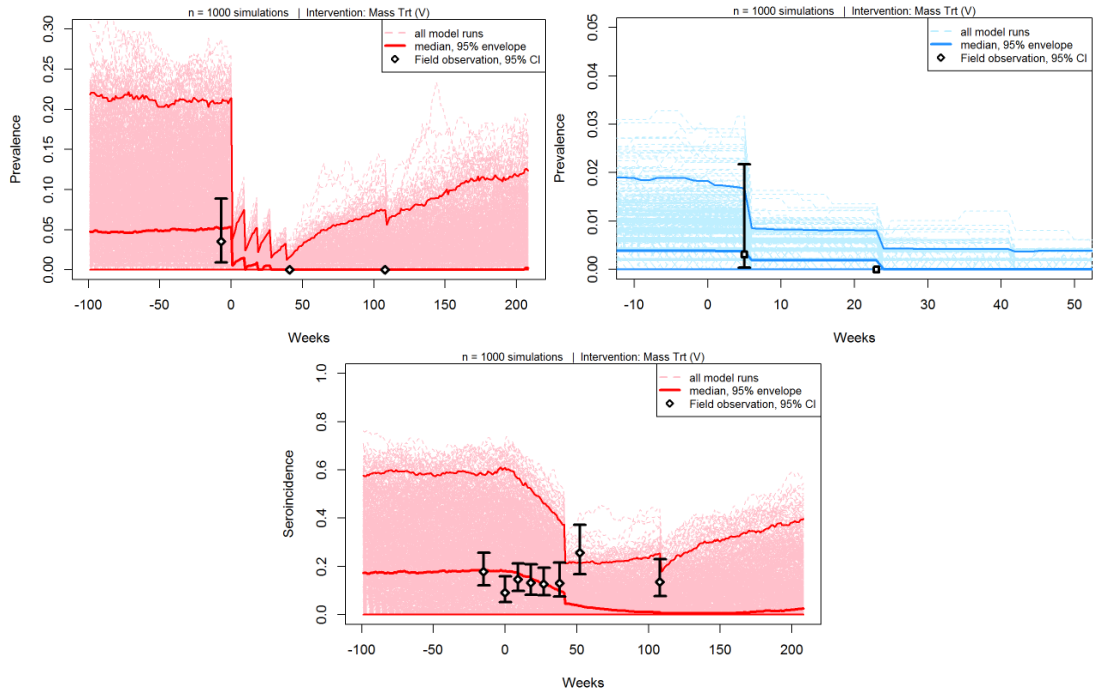
Mass Treatment w/o vaccine (CEDP, village 5 of 6)



Mass Treatment w/o vaccine (CEDP, village 6 of 6)



Mass Treatment w/ vaccine (CEDP, village 1 of 2)



Mass Treatment w/ vaccine (CEDP, village 2 of 2)

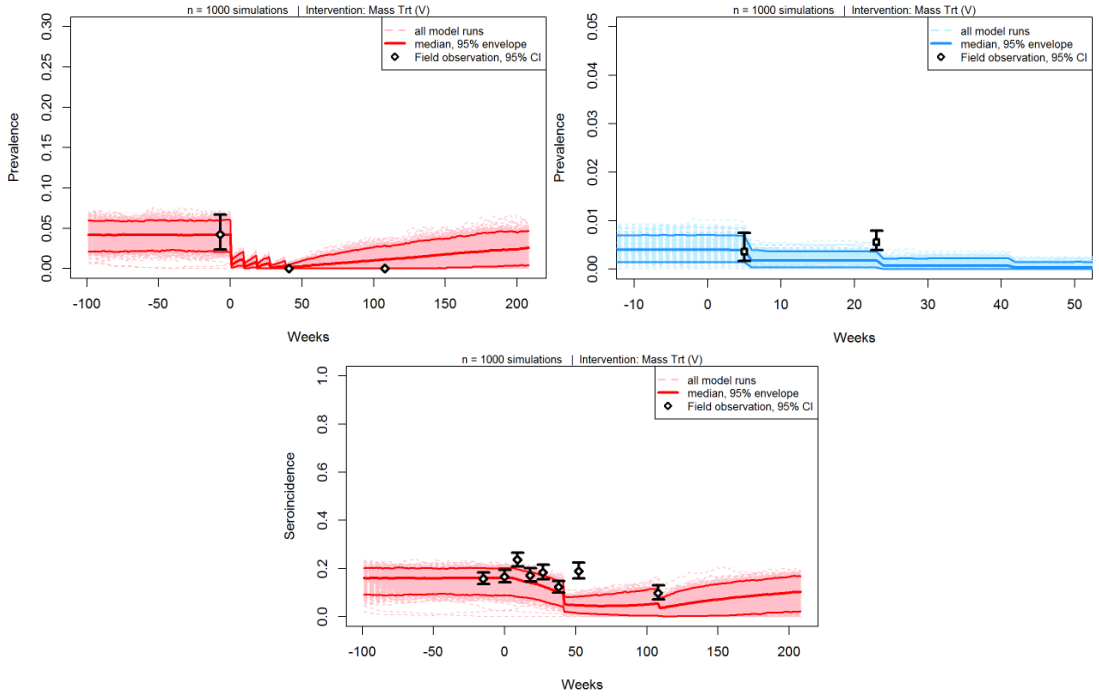
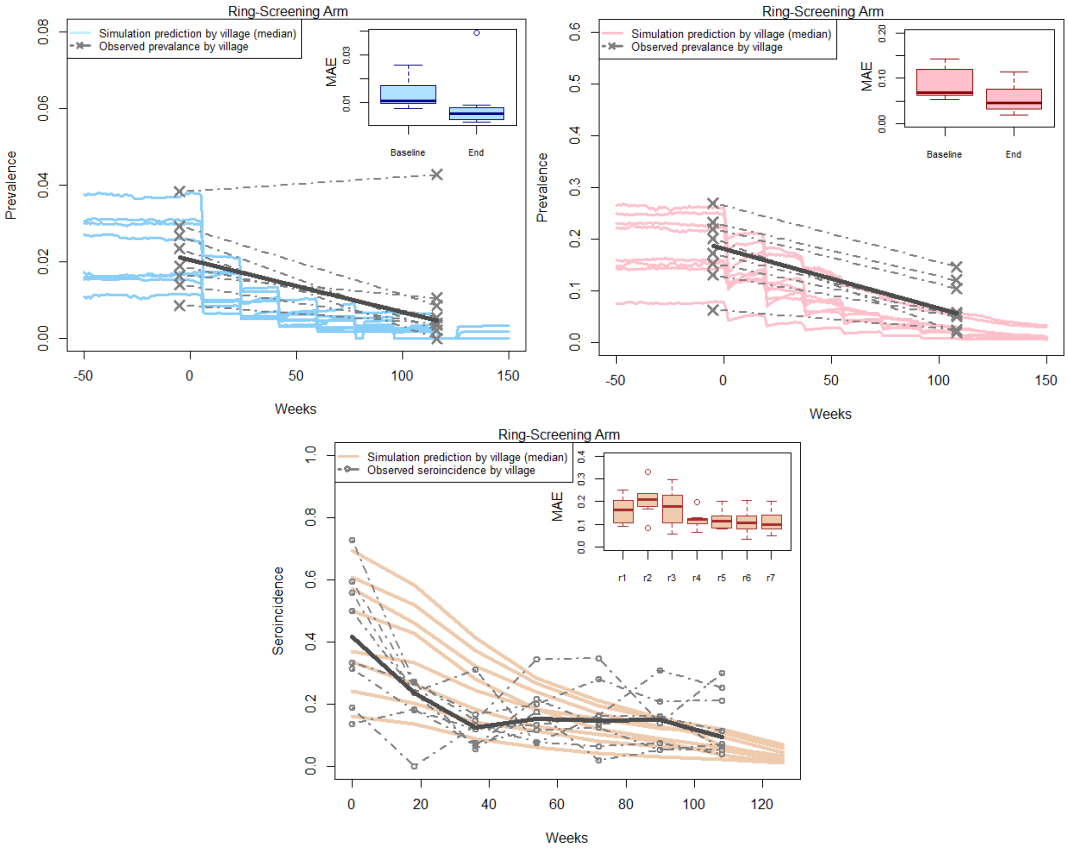
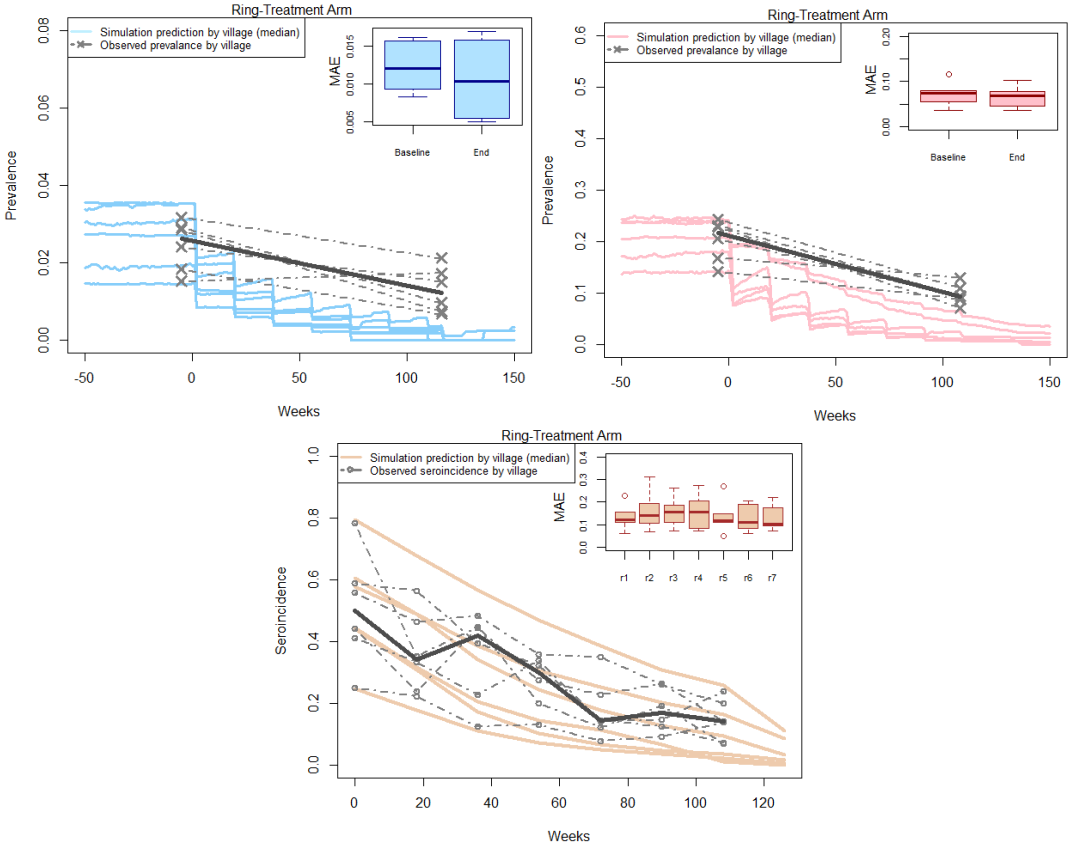


Fig C4. Summary of validation results by intervention-type (Ring Strategy Trial, n=21 villages)

<u>Ring Screening (RST, 8 villages)</u>



Ring Treatment (RST, 6 villages)



Mass Treatment (RST, 7 villages)

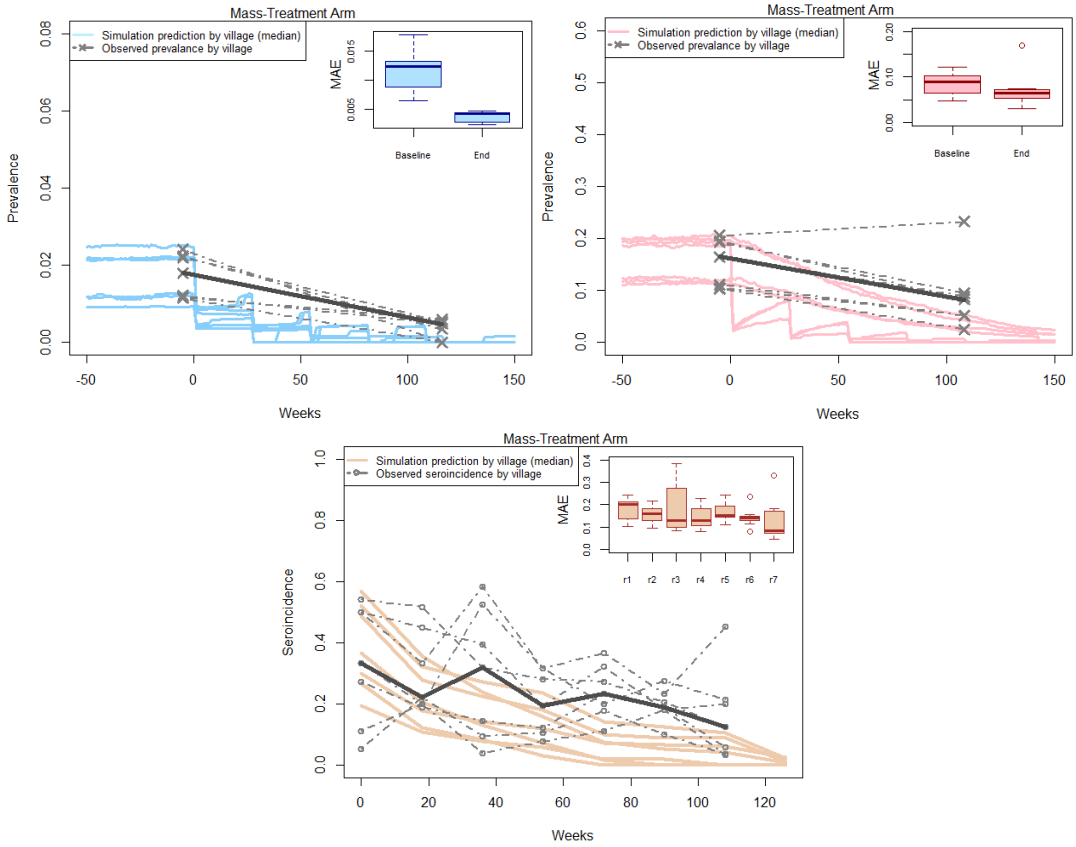
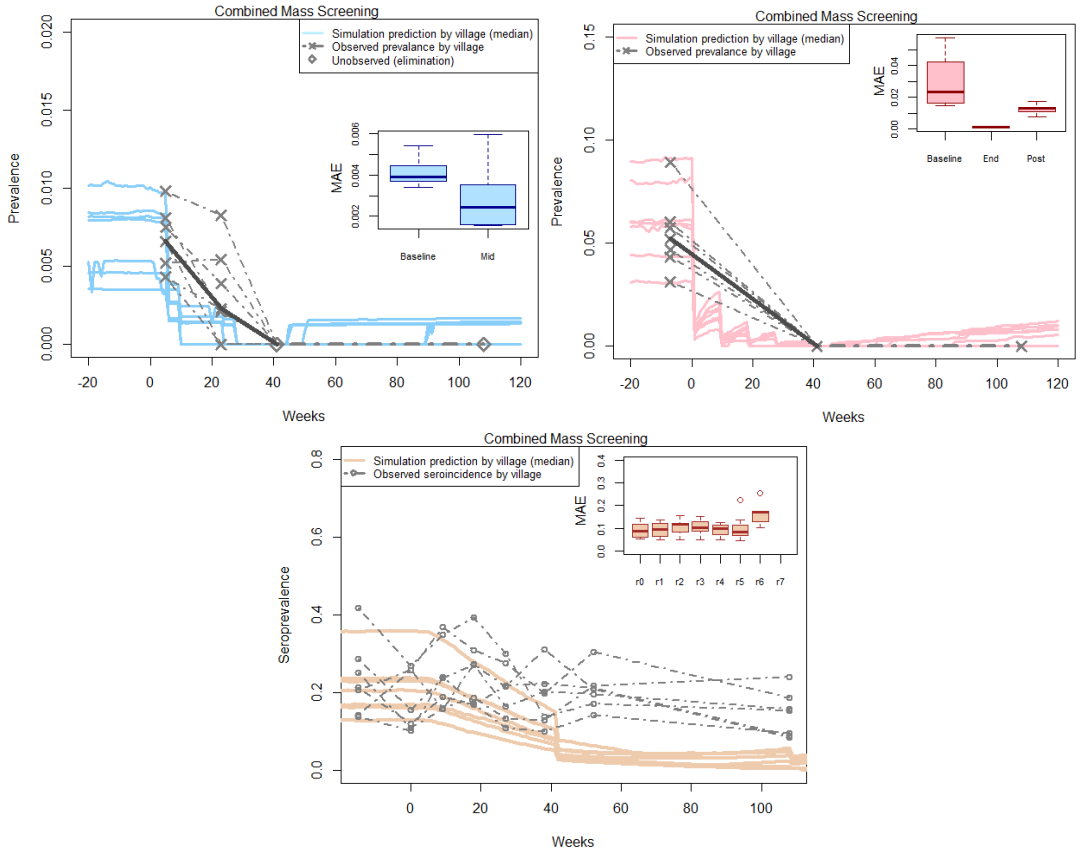


Fig C5. Summary of validation results by intervention-type (CEDP Trial, n=15 villages)

Mass Screening (CEDP, 7 villages)



Mass Treatment (CEDP, 8 villages)