Summer 2020

Incorporating Characteristics of Gene Drive Engineered Ae. aegypti as Methods to Reduce Dengue and Zika Virus into the Bayesian Network – Relativ esian Network – Relative Risk Model, Using P e Risk Model, Using Ponce, Puer once, Puerto Rico as a Case Study

Steven R. Eikenbary
Western Washington University, s.eikenbary@gmail.com

Summer 2020

Incorporating Characteristics of Gene Drive Engineered Ae. aegypti as Methods to Reduce Dengue and Zika Virus into the Bayesian Network – Relative Risk Model, Using Ponce, Puerto Rico as a Case Study

Steven R. Eikenbary
s.eikenbary@gmail.com

Incorporating Characteristics of Gene Drive Engineered *Ae. aegypti* as Methods to Reduce Dengue and Zika Virus into the Bayesian Network – Relative Risk Model, Using Ponce, Puerto Rico as a Case Study

By

Steven R. Eikenbary

Accepted in Partial Completion
of the Requirements for the Degree
Master of Science

ADVISORY COMMITTEE

Dr. Wayne G. Landis, Chair

Dr. John McLaughlin

Dr. Katherine von Stackelberg

GRADUATE SCHOOL

Dr. David L. Patrick, Dean

Master's Thesis

In presenting this thesis in partial fulfillment of the requirements for a master's degree at Western Washington University, I grant to Western Washington University the non-exclusive royalty-free right to archive, reproduce, distribute, and display the thesis in any and all forms, including electronic format, via any digital library mechanisms maintained by WWU.

I represent and warrant this is my original work, and does not infringe or violate any rights of others. I warrant that I have obtained written permissions from the owner of any third party copyrighted material included in these files.

I acknowledge that I retain ownership rights to the copyright of this work, including but not limited to the right to use all or part of this work in future works, such as articles or books.

Library users are granted permission for individual, research and non-commercial reproduction of this work for educational purposes only. Any further digital posting of this document requires specific permission from the author.

Any copying or publication of this thesis for commercial purposes, or for financial gain, is not allowed without my written permission.

Steven R. Eikenbary

10 August 2020

Incorporating Characteristics of Gene Drive Engineered *Ae. aegypti* as Methods to Reduce Dengue and Zika Virus into the Bayesian Network – Relative Risk Model, Using Ponce, Puerto Rico as a Case Study

A Thesis
Presented to
The Faculty of
Western Washington University

In Partial Fulfillment
Of the Requirements for the Degree
Master of Science

by
Steven R. Eikenbary
August 2020

Abstract

This study proposes the use of the Bayesian network relative risk model (BN-RRM) to estimate the risk associated with the release of gene drives as vectors to control disease, using Ponce, Puerto Rico as a case study. Bayesian networks are an appropriate risk assessment tool for quantitatively and probabilistically examining complex systems involving multiple stressors acting on multiple endpoints in a wide variety of situations. The emerging field of synthetic biology has the capacity to drastically alter ecological systems with the use of gene drive engineered organisms as a method to alter population dynamics. The purpose of the release of a gene drive organism is for the introduced genetic material to propagate within the wild type population and persist within the environment. There are many proposed gene drive designs and no regulatory framework that quantitatively assess the risk associated with the use of gene drive engineered organisms released to the environment. The risk assessment describes how the gene drive may spread through the populations of wild type mosquitoes and decrease rates of disease. The Bayesian network relative risk model can perform the risk assessment of gene drive engineered *Ae. aegypti* for vector control and as part of an adaptive management strategy to reduce dengue and Zika transmission. This study illustrates how the BN-RRM can integrate gene drive related information within a risk assessment framework suitable for adaptive management of these novel stressors.

Acknowledgements

I would like to thank Jared Bennet and Héctor Manuel Sánchez Castellanos of the Marshall Lab at U.C. Berkeley for their help and clarification with “MGDrivE”. Ethan A. Brown helped with numerous discussions of everything related to gene drive risk assessment. I would like to Acknowledge April Markiewicz for all the technical and logistical support at the Institute of Environmental Toxicology and Chemistry. Thank you to my thesis committee for the guidance throughout this project: Dr. Wayne G. Landis, Dr. John McLaughlin, and Dr. Katherine von Stackelberg.

Contents

List of figures

List of Tables

List of abbreviations

BN-RRM	**Bayesian network relative risk model**
Cas9	**CRIPSR associated**
CRISPR	**Cluster regularly interspersed palindromic repeats**
EPA	**U.S. Environmental Protection Agency**
GD	**Gene drive**
HDR	**Homology-directed repair**
MGDrivE	**Mosquito gene drive explorer**
NEPA	**National Environmental Policy Act**
RR1	**Risk region 1**
RR2	**Risk region 2**
RR3	**Risk region 3**

1. Introduction

1.1 Introduction to risk assessment for synthetic biology

Given the current state of gene drive technology and the lack of regulatory oversight, it is clear that risk assessment is necessary in order to understand any adverse effects that the release of gene drive organisms may pose (NASEM 2016). Therefore, the goal of my study is to use Bayesian network tools to model the risk associated with the applications of synthetic biology products and how human health endpoints are impacted as a result. Specifically, I will apply the Bayesian network relative risk model (BN-RRM) to examine the risk of releasing gene drive engineered *Aedes aegypti* mosquitoes as vectors to control both the dengue and Zika virus using Ponce, Puerto Rico as a case study.

1.2 Gene drives

The use of gene drives to control pest species has been speculated since the late sixties (Curtis 1968; Foster et al. 1972). Gene drives are genetic constructs that represent themselves at rates greater than that of the Mendelian 50/50 ratio. Using a gene drive a trait deemed beneficial for some purpose could be introduced into a target population and driven to fixation in a relatively quick timeframe (Webber et al. 2015). Developments in biotechnology and synthetic biology engineering have made gene drives no longer speculation but a reality. The discovery and use of clustered regularly interspaced short palindromic repeats (CRISPR) and CRISPR associated protein 9 (Cas9) as a genetic engineering tool has led to the ability to precisely edit the genetic code of almost any living organism (Webber et al. 2015; Knott & Doudna 2018).

1.3 CRISPR/Cas9

Discovered in bacteria, CRISPR/Cas9 acts as an immune response to foreign viral DNA within bacterial cells (Knott & Doudna 2018). The CRISPR/Cas9 complex re-writes segments of genetic code by cutting a strand of DNA at a specific site and either replacing or removing a

genetic sequence. Using a complementary guide-RNA, the CRISPR/Cas9 complex homes to a specific sequence of genetic code. Once this site is located on the genome, a double strand break is made and the new sequence of base pairs is inserted into the target site. The CRISPR/Cas9 complex then disassociates and the cells own DNA repair mechanisms complete the process. By attaching a cargo sequence to the CRISPR/Cas9 complex, new traits can be written into the genetic code of nearly any organism (Webber et al. 2015; Knott & Doudna 2018).

1.4 Inheritance

Gene drives can and do occur naturally as selfish genetic elements (Burt and Trivers 2006). These naturally occurring drive systems achieve superseded Mendelian inheritance rates without any positive implications to individual fitness (Burt and Trivers 2006). These naturally occurring drive systems include: transposable elements (McClintock 1951; McClintock 1956; and Wicker et al. 2007), meiotic drives (McDermott and Noor 2010; Hickey and Craig 1966; Sweeny and Barr 1978; Hiraizumi and Crow 1960; Silver 1993; Ardlie 1998; Rhoades and Dempsey 1985), underdominance (Altrock et al. 2011; Curtis 1968), maternal-effect dominant embryonic arrest (Beeman et al. 1992), and homing endonuclease genes (Fraser 2012; Jasin 1996)

The transposable elements are short sequences of DNA that can change location within the genome by randomly removing and reinserting themselves at another location. Transposable elements were discovered in 1952 in maize (McClintock 1951; McClintock 1956) and have since been described in other eukaryotes (Wicker et al. 2007).

Meiotic drive systems are interference drive mechanisms that lead to disproportional allele frequencies relative to Mendelian rates (McDermott and Noor 2010). These types of gene drives interfere with some aspect of the phenotypic expression as a result of reducing functionality at

certain loci. For example, in mosquitoes a meiotic drive has been observed that interferes with sex-linked chromosomes, resulting in a disproportionate male-to-female sex ratio, favoring males (Hickey and Craig 1966; Sweeny and Barr 1978). In addition to mosquitoes (Hickey and Craig 1966; Sweeny and Barr 1978), meiotic drive systems have been observed in *Drosophila melanogaster* (Hiraizumi and Crow 1960), *Mus musculus* (Silver 1993; Ardlie 1998), and maize (Rhoades and Dempsey 1985).

Other forms of naturally occurring gene drive systems include underdominance. Underdominance occurs when the heterozygous offspring of two homozygous individuals incurs a fitness reduction relative to the parent organisms (Altrock et al. 2011). Curtis (1968) described underdominance systems that could be used to reduce pest species by causing sterility through introduced genes.

Another gene drive system, maternal-effect dominant embryonic arrest, has been described in the flour beetle (*Tribolium castaneum*) where offspring that lack the associated chromosome are nonviable (Beeman et al. 1992). Only offspring that are homozygous for the maternal-effect dominant embryonic arrest chromosome develop properly, thereby superseding Mendelian inheritance rates (Beeman et al. 2012).

Although gene drives occur naturally, they can now be created in the lab using techniques available through CRISPR/Cas9 (Gantz et al. 2015; Hammond et al. 2016; Champer et al. 2017; Buchman et al. 2018; Buchman et al. 2019). When CRISPR/Cas9 inputs a genetic sequence into the germline of an organism, the associated trait is inheritable (Webber et al. 2015). Regular Mendelian inheritance rates can then be superseded by incorporating the sequence calling for the entire CRISPR/Cas9 and guide-RNA complex and any attached cargo sequence into each chromosome during cellular development (Leftwich et al. 2018). In a rapidly reproducing target population with short generational times it may be possible that an introduced trait can be driven to fixation over a relatively short temporal scale, despite any

potential fitness implication due to the presence of the gene drive (Webber et al. 2015; Leftwich et al. 2018).

1.5 Proposed applications for environmental management

The broad applicability of genetic manipulation using CRISPR/Cas9 has led to numerous speculative proposals for the use of gene drives as ecosystem engineering catalysts (Webber et al. 2015; Gantz et al. 2015; Hammond et al. 2016; Champer et al. 2017; Buchman et al. 2018; Buchman et al. 2019). The majority of these potential uses fall under two main categories: population replacement or population suppression (Leftwich et al. 2018). Population replacement gene drives introduce a desired trait into a target population and over time replace the existing wild-type population with the new population that carries the introduced trait. Population suppression gene drives introduce a deleterious trait into a target population that spreads regardless of the magnitude of fitness costs at super Mendelian rates, potentially crashing populations within several generations (Windbichler et al. 2011; Adelman & Tu 2016; Hammond et al. 2016; Champer et al. 2017; Marcias et al. 2017; Buchman et al. 2018; Kyrou et al. 2018).

1.6 Pest management

Population suppression gene drives have been described for use as pest management tools in the context of invasive species (Prowse et al. 2017), agricultural pest species (Scott et al. 2018), and as methods of reducing mosquito borne illness (Windbichler et al. 2011; Hammond et al. 2016; Champer et al. 2017; Buchman et al. 2018; Kyrou et al. 2018). Suppression drives have been modeled for use in invasive mouse populations by introducing sex specific traits that lead to populations becoming infertile (Prowse et al. 2017). Synthetic CRISPR/Cas9 based gene drives that target agricultural pest species that are normally controlled through the use of pesticides have been described (Scott et al. 2018). Perhaps the most commonly proposed application of population suppression drives is the control of species that vector disease,

especially in mosquitoes (Windbichler et al. 2011; Hammond et al. 2016; Champer et al. 2017; Buchman et al. 2018; Kyrou et al. 2018).

1.7 Mosquito borne disease

Worldwide the burden of mosquito borne disease is staggering. Even with the best attempts at management, malaria alone kills on average over 400,000 people each year (WHO 2019). Many of these deaths occur in Sub-Saharan Africa and Asia. In the Americas, dengue is more prevalent, with 10,160,612 reported cases and 5,163 deaths over the last 7 years (PAHO 2020). Newly introduced diseases, like Zika virus, when introduced into an immunologically naïve population, have the capability to spread at epidemic proportions (Barrera et al. 2017). Current management strategies center on breeding habitat reduction and the use of pesticides when removal of breeding habitat is not feasible. Other management strategies include aerial spraying of insecticides indoors and outdoors, and the use of structural barriers, like mosquito nets and window screens (EPA 2016a, b, c). There are no current vaccinations for either dengue or Zika, though their development is being studied (Noyd and Sharp 2015; Barrett 2018; CDC 2018).

1.8 Current synthetic biology regulations

The risks posed by the use of synthetic biology products are not adequately covered by the current regulatory framework in the U.S. (NASEM 2016). The current regulatory laws are not specific to synthetic biology and therefore do not cover the unique risks that synthetic biology poses (NASEM 2016). Under the National Environmental Policy Act (NEPA) there are no requirements for quantitative assessment of risk nor the probability of effects. Current NEPA regulations do not require uncertainty to be reported. The lack of testable hypotheses of the environmental impact statement, that NEPA does require, makes the process less geared toward developing future research. Ecological risk assessment gives the probability of the desired outcomes as well as the indirect effects. Though NEPA does not require an ecological risk assessment to be conducted, other various regulations do (e.g. Federal Insecticide,

Fungicide and Rodenticide Act, Comprehensive Environmental Response, Compensation and Liability Act, and Toxic Substances Control Act). The general guidelines that the U.S. Environmental Protection Agency (EPA) requires for ecological risk assessment are: problem formulation, analysis, and risk characterization. These EPA guidelines do not require quantitative, probabilistic outcomes (NASEM 2016).

Many questions about the use of gene drive organisms are related to probabilities of the outcomes both desired and undesired (NASEM 2016). The desired outcome of the release of a gene drive is to introduce a genetic element that persists in the environment through preferential inheritance and has an effect on population dynamics. The current field is lacking in studies that describe the risk associated with the introduction of a gene drive engineered species. As currently framed by the EPA, ecological risk assessments are not well suited to mathematically describe the multiplicity of interactions between stressors and endpoints over broad spatial scales. The BN-RRM is designed for such complex interactions and are able to quantitatively represent the interactions among multiple stressors and multiple endpoints across large spatial and temporal scales (NASEM 2016).

1.9 The relative risk model for ecological risk assessment

Landis and Wiegers (1997) developed an ecological risk assessment design that considered the interactions between the stressors and endpoints, the variation of the different stressors, and the spatial interests of stakeholders. This ecological risk assessment approach was applied to a real-world scenario by Wiegers et al. (1998), where risks and associated impacts of anthropogenic stressors were compared and evaluated for the marine environment of Port Valdez, Alaska. The model aided managers in developing testable hypotheses to further reduce impacts (Wiegers et al. 1998). Further iterations led to the development of the relative risk model that has been successfully applied to many very different risk assessment scenarios

(Hayes and Landis 2004; Landis and Wiegers 2005; Colnar and Landis 2007; Anderson and Landis 2012).

Hayes and Landis (2004) used the relative risk model to assess the regional coastal marine area near Cherry Point, WA. The goal was to develop a model that estimated the risk associated with anthropogenic sources to biological endpoints in order to aid the decision-making process of the Washington Department of Natural Resources. The results of the model depicted the risk of habitats to stressors, which stressors were likely to result in stronger impacts, and the endpoints most at risk (Hayes and Landis 2004).

In Landis and Wiegers (2005), the outline of a regional ecological risk assessment was presented. In the assessment, habitats and sources of stressors within the study area are classified. Ranks were given weighting the relative importance of each per location within the study area to give levels of possible risk. If a succinct pathway between components of the RRM was found, then there was a risk to the associated endpoint.

In Colnar and Landis (2007), the previously developed Relative Risk Model was applied to determining the risk associated with an invasive species (*C. maenas*) to the Cherry Point, WA, region. The authors investigated the probability of introduction of green crabs to the study area, as well as the effects of the introduction green crabs to the overall region. The risk to the region was determined using two scenarios: 1) established conditions, and 2) conditions associated with El Niño events which may influence dispersal patterns. The objective for the authors was to develop a system which incorporated conditions and management goals to develop testable hypotheses which could lead to better allocation of resources by management to attain their goals (Colnar and Landis 2007).

In Anderson and Landis (2012), the relative risk model was applied to an entire watershed which consists of the Upper Grande Ronde area of eastern Oregon. The area of study was an

area of concern as outlined by the U.S. Forest Service. The goals brought forward by the U.S. Forest Service included management strategies aimed at predicting the outcomes of disturbance regimes and other processes (Anderson and Landis 2012).

1.10 The Bayesian network relative risk model

Cain et al. (1999) described the application of Bayesian networks to natural resource management. Cain et al. (1999) found that Bayesian networks provide quantitative methods for analyzing ecological systems and highlight uncertainties in a way that managers and the public can easily understand. In a key publication, Marcot et al. (2006) developed a set of guidelines for constructing Bayesian networks specifically to be used in ecological scenarios. Marcot et al. (2006) argued that although Bayesian networks had been used in the past to model complex ecosystem interactions, the models themselves had not been developed using a standard process, nor had they been consistently evaluated after more data became available.

Using the guidelines constructed by Cain et al. (1999) and Marcot et al. (2006), various BN-RRMs were developed that showed their effectiveness in modeling risk over large spatial scales (Ayre and Landis 2012), climate change (Landis et al. 2013), contaminants and urbanization (Hines and Landis 2013), the spread of a novel disease in trout species (Ayre et al. 2014), and risk to benthic estuarine populations (Graham et al. 2019). Ayre and Landis (2012) first applied the Bayesian network approach to ecological risk assessment while studying a large-scale terrestrial ecosystem. In Landis et al. (2013), the BN-RRM was used as an approach to assessing risk to systems from global climate change. Hines and Landis (2013) assessed risk of prespawn mortality in Coho salmon as a result of contaminants associated with urbanization and land-use. Ayre et al. (2014) applied the BN-RRM to the spread of whirling disease in native populations of cutthroat trout in the Colorado and Rio Grande Rivers. Graham et al. (2019) showed another case where the BN-RRM is an appropriate risk assessment tool by modeling benthic population risk in estuary systems in Australia.

The use of gene drives creates a potential for long term effects and impacts in a variety of different ways and ecosystems (NASEM 2016). Any design of a gene drive will impact not just the target species, but potentially all other species within an ecosystem to more or less of a degree. Gene drives have been proposed for use in terrestrial as well as aquatic systems, leaving the potential for interactions almost endless. The use of BN-RRMs can provide the foundation for evaluating the risk associated with gene drives because information regarding the genetics and ecological interactions can included in the BN-RRM ecological risk assessment approach. There are components of BN-RRMs that make them well suited to assessing risk related to the deployment of gene drives.

The BN-RRM requires clear dialog between stakeholders, the public, managers, and researchers to define goals and endpoints. For gene drive use, there will be stakeholders represented from every socio-economic background with differing values that need to be addressed. When a mechanistic understanding of interactions is known, causal pathways are clearly represented within the BN-RRM and the outcomes are given as probabilities of relative risk. This poses a potential source of uncertainty for gene drives because there are lots of areas of uncertainty. For example, greater understanding of how the gene drive spreads in lab trials does not necessarily translate into field trials.

As gene drive research progresses, more information gained at the cellular level may produce more questions at the organismal level. The world that gene drives will ultimately be introduced is inherently an open system, meaning that the interactions taking place as a result of the drive will change within a project as the timescale changes, as well as from project to project as the location changes. Alternative management strategies are easily assessed as part of the design and uncertainty is easily examined. The trade-offs of alternative management options can be represented within the BN-RRM. This means that non-gene drive interventions can be modeled

alongside the use of a gene drive or in conjunction with the gene drive, with outcomes represented probabilistically.

1.11 Ten steps of the Bayesian Network-Relative Risk Model

Ten steps have been described in the development of the relative risk model (Landis and Wiegers 2005) and have been further adapted to incorporate Bayesian networks (Ayre and Landis 2012). These steps will be used to establish the BN-RRM framework for and calculations of risk associated with the release of gene drive engineered mosquitos in Ponce, Puerto Rico.

Step 1: Explicitly establish the management goals, defining specific endpoint criteria and ecological services important for decision makers (Landis and Wiegers 2005). It is imperative for the risk assessment process that endpoints be chosen through clear dialogue with all applicable stakeholders. Communicating with current management and others through personal contact will be essential to this step.

Step 2: Develop a map that includes the locations of sources and habitats of the areas to be treated. This includes the spatial extent of the release of the gene drive engineered mosquitos, the land use and land class attributes of the area, suitable mosquito habitat, the distribution of human dwellings, as well as any landscape characteristics that may influence migration rates and dispersal of the mosquito populations (Landis and Wiegers 2005).

Step 3: Develop risk regions based on spatial connectedness of sources, stressors, habitats, and endpoints. Risk regions will be discerned by the ability of the gene drive to persist and disperse throughout the study area. This includes all of the landscape characteristics that were determined in step 2 (Landis and Wiegers 2005). The determination of risk regions will depend on the factors within the landscape that influence mosquito population dynamics and migration rates of the organisms, as determined through GIS analyses, and the dispersal rates of the

gene drive throughout the wild type *Ae. aegypti* population as determined through population and gene drive propagation simulations.

Step 4: Develop a conceptual model that describes the causal linkages between sources, stressors, habitats, effects, and impacts. The criteria for inclusion into the conceptual model are clearly understood mechanistic properties from exposure to effect (Landis and Wiegers 2005). This conceptual model serves as the cause-effect framework for the BN-RRM (Ayre and Landis 2012).

Step 5: Use the linkages derived in step 4 to create a Bayesian network, including risk rankings to calculate risk to specified endpoints. The software Netica™ is used to develop the Bayesian network (Norsys 2014). Discrete states (zero, low, medium, or high) are assigned to all nodes within the Bayesian network. Higher tiered nodes (parent nodes) that influence lower tiered nodes (child nodes) are connected through causal linkages, creating causal chains from sources through endpoints. Probability distributions of outcomes for each state, which are determined through observations, derived from the literature, or from model output, are assigned to the nodes within the Bayesian network. Conditional probability tables are used to quantify the causal linkages. The conditional probability tables describe the interactions of all possible combinations of each state of the parent nodes that connect to a child node. The information built into step 5 is derived from the peer-reviewed literature, government sources, expert elicitation and site-specific data (Landis and Wiegers 2005).

Step 6: Calculate risk based on the model output. Risk will be calculated by running the model. (Landis and Wiegers 2005).

Step 7: Conduct model sensitivity and uncertainty analyses. Sensitivity analysis will be done by determining the input classifications that are most important changes in risk calculations through mutual information. Mutual information is the amount of information that is shared between two

nodes in the network. Larger amounts of mutual information between two variables indicates that an input variable has more influence on the probability distribution seen in an endpoint node (Norsys 2014). Outcomes of sensitivity analysis will illustrate the factors most influential to the risk calculations and can identify priorities for increasing accuracy of risk estimates and reducing uncertainties in distributions used in step 5 (Landis and Wiegers 2005).

Step 8: Use the model outputs to develop testable hypotheses to enhance knowledge of the system. The questions developed will clarify areas of the model needing more information (Landis and Wiegers 2005).

Step 9: Use the hypotheses from step 8 to conduct further research. Testing hypotheses from step 8 will create data that can be used to further parameterize the model, resulting in less model sensitivity and uncertainty (Landis and Wiegers 2005).

Step 10: Communicate results established through the BN-RRM calculations. Providing results of the modeling to decision makers and publishing results in peer-reviewed scientific literature will be the top priorities in communicating the outcomes (Landis and Wiegers 2005).

1.12 Study objectives

The risk assessment approach proposed by the National Academies of Science, Engineering, and Medicine (NASEM 2016) to model the risk posed by gene drives is used here. The goal is to show that a quantitative risk assessment of gene drive application as a method to reduce the transmission of mosquito borne disease is attainable using realistic landscape characteristics in a large metropolitan area and can be easily reproduced for other scenarios involving the use of gene drives, given that sufficient data exist regarding: the life history traits of the target species, the interactions between the host and vector populations, and the landscape characteristics that influence the productivity of the landscape in terms of vector population numbers.

1.13 Case study - Ponce, Puerto Rico

The municipality of Ponce, Puerto Rico is the second most densely populated area in Puerto Rico following the greater San Juan area (Figure 1). The *Ae. aegypti* mosquito is the primary vector of concern for mosquito borne disease on the island and within Ponce. This mosquito is responsible for the spread of endemic dengue as well as introduced Zika virus (CDC 2017). The adaptive management strategies currently in place center on eliminating the larval stages from the environment (CDC 2016a, EPA 2016a, b). When larval eradication is not attainable, the use of pesticides to control adults is leveraged (EPA 2016a, b). Other measures used are physical barriers that limit the exposure of humans to mosquitoes, such as bed nets and window screens (EPA 2016a, b).

Here I model the release of gene drive engineered *Ae. aegypti* and incorporate the findings to quantitatively assess the direct and indirect effects of their use. The landscape characteristics of Ponce, Puerto Rico provide the information regarding the habitat parameters of the wild and subsequent hybrid populations. The connectivity of the habitat describes the movement of the gene drive throughout the population. Habitat connectivity as determined through published mosquito dispersal patterns and GIS analyses of distances between habitat patches will influence the spread of resistance to the drive that develops in one patch and the rates of the developed resistance spreading to other patches, until a new resistant population has replaced the wild type population.

The emergence of resistance is a highly advantageous trait in the presence of the extreme selective pressure posed by the gene drive and almost assuredly occurs (Unckless et al. 2017). Homology-directed repair (HDR) is the process where a double strand break in a segment of

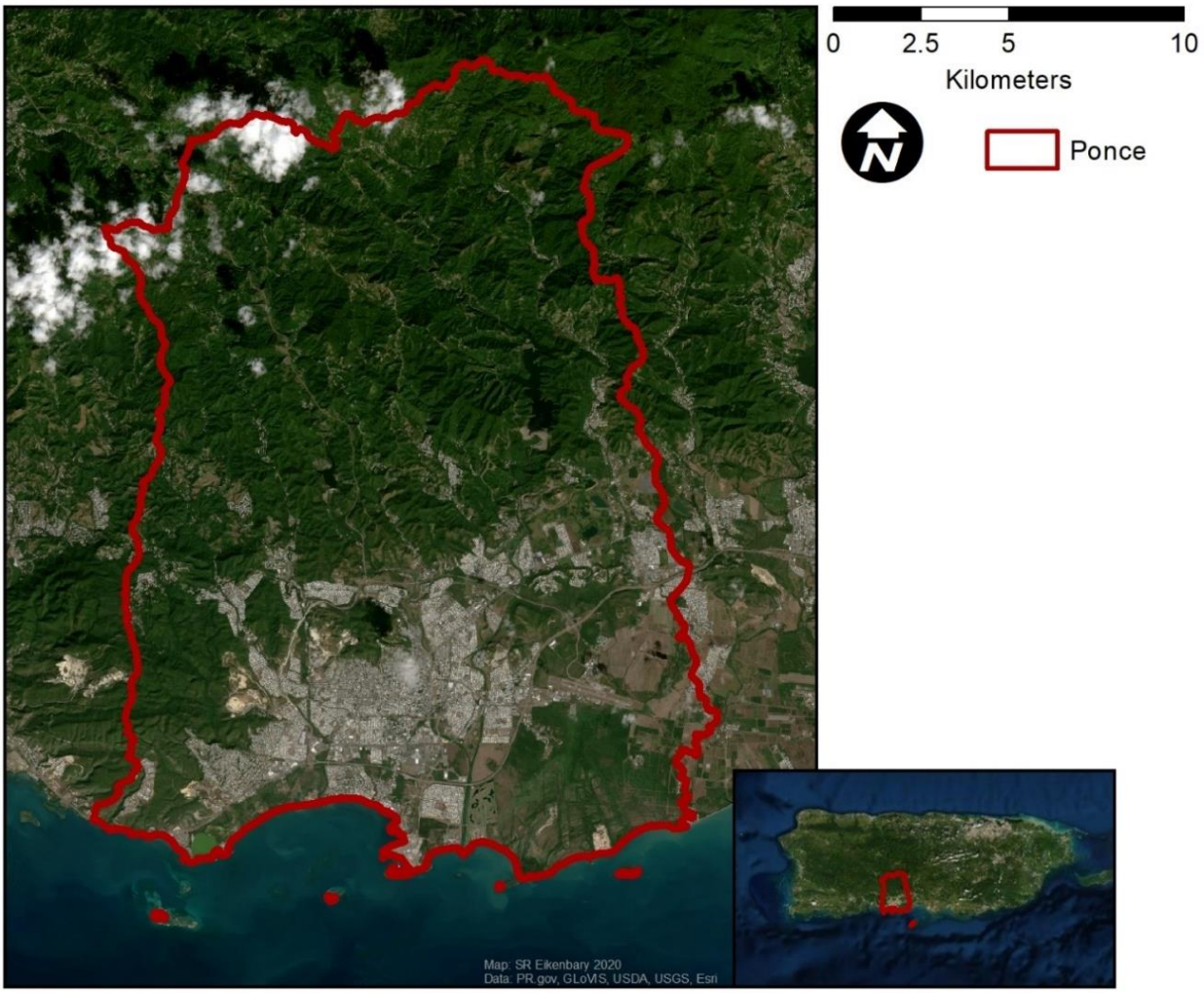

Figure 1. Map showing the boundary of the municipality of Ponce, Puerto Rico.

DNA is repaired using an identical DNA sequence as a template (Gratz et al. 2014). By using an identical sequence as a repair template, the breakage can be repaired very precisely. In contrast, non-homologous end-joining occurs when a double strand break in a segment of DNA is repaired by the addition of compatible nucleotides to each end of the break that are then fused together (Weterings & Chen 2008). This can result in the loss of nucleotides at a specific locus. These inaccurate repairs due to non-homologous end-joining can give rise to the formation of alleles that give resistance to the gene drive (Unckless et al. 2017).

The habitat patch where resistance develops is a source of further resistance, given that the initially resistant organisms successfully breed and reach the minimum viable population size required within that patch (Deines et al. 2005). Large areas of suitable habitat for the resistant individuals become available for recolonization as the gene drive decreases the wild-type population because the habitat requirements between the two are identical. The entire study area represents a patchy metapopulation because of the spatial heterogeneity of the environment (Deines et al. 2005). Resources are located in discrete locations correlating with distributions of human habitations that subdivides the existing mosquito population into discrete patches (Lutambi et al. 2013) and dispersal of the resistant individuals is able to occur through migration between patches (Deines et al. 2005).

1.14 Current management strategies

The current mosquito management strategy in Puerto Rico, implemented by the Puerto Rico Vector Control Unit, is a three-pronged integrative pest management approach based on reduction of breeding habitat, introduction and maintenance of barriers to movement, and the use of pesticides to control larval and adult stages when reduction of breeding habitat is not feasible (CDC 2016a, CDC 2016b). The United States Environmental Protection Agency (EPA) recommends the control of mosquito populations primarily at the larval stage when populations are most vulnerable. The use of larvicides focuses on treating breeding habitat in order to reduce maturation into adult life stages and dispersal (EPA 2016a). Currently, the EPA and the World Health Organization approve the use of organophosphate insecticides, insect growth inhibitors, and bacterial insecticides for use as mosquito larvicides (EPA 2016a; WHO 2017).

Organophosphate larvicides cause mortality in mosquito larvae by interfering with nerve transmission (CDC 2016a). Organophosphates are recommended as larvicides over pyrethroids in order to avoid the development of pyrethroid resistance in larval mosquitos. Due to the wide variety of suitable breeding habitat for the *Aedes* spp. mosquitos and the difficulty in eliminating

all sources of standing water, use of larvicides is recommended as an integrated part of a more comprehensive mosquito control program (CDC 2016a).

Insect growth inhibitor larvicides obstruct developmental processes in larval mosquitos, effectively reducing adult mosquito populations. This class of chemicals mimics juvenile insect hormones and interferes with larval development (EPA 2001).

Bacterial insecticides include *Bacillus thuringiensis iraelensis* and *Bacillus spaericus*. Spores of the bacterium *Bacillus thuringiensis iraelensis* create toxins that are detrimental to mosquito, blackfly, and fungus gnat larvae. *Bacillus sphaericus* is a bacterium that produces spores that inhibits the digestive processes in the gut of mosquito larvae (EPA 2016a).

Mosquito adulticides target adult mosquito populations and can be applied through aerial or truck mounted spraying. The EPA lists 2 classes of adulticides for use: organophosphates and synthetic pyrethroids (EPA 2016b).

The organophosphate insecticides approved by the EPA are malathion, temephos and Naled (1,2-dibromo-2,2-dichloroethylphosphate). These chemicals are neurotransmitter inhibitors that work by interfering with the acetylcholinesterase receptor. Malathion is primarily applied with truck-mounted sprayers (EPA 2016c). Naled is primarily applied aerially with ultra-low volume sprayers. Temephos is applied directly to water sources to control the larval stage. Unlike malathion, resistance to Naled or temephos in *Ae. aegypti* has not been detected (Del Rio-Galvin et al. 2016; EPA 2016c; EPA 2018).

The pyrethroid class of insecticides approved by the EPA for use as mosquito adulticides are applied by ultra-low volume aerial spraying (EPA 2019). Pyrethroids approved for use in Puerto Rico include: alpha-cypermethrin, bifenthrin, deltamethrin, etofenprox, lambda-cyhalothrin, permethrin, phenothrin, and tetramethrin (CDC 2017). These chemicals kill adult mosquitos through paralysis via neurotransmitter disruption (Soderlund et al. 2002). Synthetic pyrethroid

adulticides are used in conjunction with synergistic compounds (e.g. piperonyl butoxide) that enhance toxicity (EPA 2019).

Sterile insect technique is another management option that results in non-viable embryos produced from wild-type females (Soderlund et al. 2002). Sterile insect technique releases laboratory reared sterile males that are treated with radiation to induce chromosomal breakages. When released into wild pest populations the embryos produced when sterile insect technique males are mated with wild-type females are non-viable (Soderlund et al. 2002).

2 METHODS

2.1 Scenario development

Amid the Zika epidemic of 2015-2016, the Centers for Disease Control and Prevention (CDC) and the Puerto Rico Science, Technology and Research Trust established the Puerto Rico Vector Control Unit as the implementation arm of the integrative pest management strategy for Puerto Rico (CDC 2016b). The Puerto Rico Vector Control Unit is tasked with the implementation of control strategies that will reduce mosquito populations with the goals of eliminating the *Ae. aegypti* mosquito and reducing mosquito borne disease transmission (CDC 2016b). This private entity was largely born out of a workshop hosted by the Puerto Rico Science, Technology and Research Trust that concluded the *Ae. aegypti* mosquito, the primary vector for dengue, Zika, and chikungunya, could be eliminated through an integrated pest management strategy that leverages reducing breeding habitat, the usage of window screens and bed nets, and the applications of pesticides (CDC 2016b). This study models the existing management options plus the addition of simulated gene drive releases to model the outcomes in relation to management goals.

2.2 Site conditions

The municipality of Ponce, Puerto Rico, covers an area of 297 km^2 (Figure 1). The U.S. Census lists the population as 140,859 people. The population density of the municipality of Ponce is 474 persons/km^2. Ponce, Puerto Rico is composed of rural, urban, and suburban landscapes (U.S. Census Bureau 2019). The Caribbean Sea borders the southern section of the municipality and there is a large global shipping port within the Ponce Bay. The Köppen-Geiger climate classification of Ponce, Puerto Rico, is tropical, savannah (Beck et al. 2018), the region receives on average 935 mm of rainfall per year and has an average temperature of 26.2°C, with seasonal highs of both temperature and precipitation in the late-summer and fall months (SERCC 2007). Since 2015 there have been 2,791 reported cases of Zika in Ponce, Puerto Rico (CDC 2019b; ArboNET 2019). Of these cases, 2,788 cases have been locally transmitted and 3 cases have been due to patients returning from travel. Puerto Rico has experienced dengue outbreaks since 1963 (CDC 2019b). There have been 1,287 reported cases of dengue in Ponce, Puerto Rico since 2010, all cases have been reported as locally transmitted (ArboNET 2019).

2.3 Gene drive applications: Population replacement and population suppression

Although several applications of gene drives have been proposed, they nearly all fall under two main categories, population replacement or population suppression (Leftwich et al. 2018). Population replacement gene drives introduce a trait into the target population that would be driven to dominance over successive generations, effectively replacing the wild population with a subsequent population exhibiting the introduced trait. Population suppression drives would introduce a trait into the wild population that would incur a fitness reduction in the individuals that carried that trait, leading to overall population crashes and eventual extinction of the target population (Leftwich et al. 2018). The gene drive constructs discussed in greater detail within this research fall under the population suppression category and are modeled after the construct

developed by Hammond et al. (2016). The theoretical improvements to the resistance allele generation rates suggested through the use of multiple guide-RNA sequences are also modeled (Marshall et al. 2017).

Multiplexing uses multiple guide-RNA sequences within the CRISPR/Cas9 complex to target multiple locations on the genome as introduction sites of the gene drive (Marshall et al. 2017). Gene drives that use multiplexing would target multiple locations on the genome thereby reducing the probability of resistance generating to the drive because resistance would have to develop simultaneously at each location. As the number of locations that the multiplexed gene drive targets increases, the theorized rate of resistance generation is decreased exponentially (Marshall et al. 2017; Noble et al. 2017).

The population suppression gene drive system that provided the basis for the gene drive modeled within this study was developed by Hammond et al. (2016). This system was first engineered for use in *Drosophila melanogaster* and successfully demonstrated in the malaria vector *Anopheles gambiae*. The Hammond et al. (2016) gene drive selected haplosufficient female sterility genes and disrupted their normal expression through gene knockout. Haplosufficiency occurs when a single copy of a functional gene is sufficient to maintain the same normal functions as those seen in individuals carrying two copies of the gene. Females that were homozygous for the gene drive were sterile and failed to lay eggs. The heterozygous females did not incur any fitness reduction (Hammond et al. 2016).

2.4 Endpoint selection

The BN-RRM requires the selection of endpoints that includes the entity being examined and a specific attribute or characteristic of the entity that can be measured (Landis & Wiegers 1997; Landis & Wiegers 2005). Endpoints selected are the likelihood of reductions in dengue and Zika transmission as a result of population reductions due to the simulated releases of the gene drive

engineered *Ae. aegypti* and the development of resistance to the drive. The first endpoint is the eradication of the wild type mosquito and is defined as zero remaining *Ae. aegypti* within the study area. Eradication is shown as an intermediate node in the Bayesian network because eradication proved to be unlikely depending on the resistance allele generation rate. The second endpoint, development of resistance to the drive is defined as the proportion of remaining mosquitoes surviving after the releases of the gene drive have ceased in terms of relative proportionality from 0 to 1. The third endpoint, the likelihood of reduction in dengue transmission is defined as the probability of being below the dengue transmission threshold of 1.42 pupae/person at a temperature of 28°C (as developed in Focks et al. 2000). The fourth endpoint node, Zika transmission, is the likelihood of the remaining *Ae. aegypti* densities being below the transmission threshold of 0.2 females/trap/day (as developed in Barrera et al. 2016).

2.5 Wild-type mosquito population numbers

The relationship between landscape characteristics and existing mosquito population numbers were incorporated from a previous study conducted in San Juan, Puerto Rico (Barrera et al. 2019). Barrera et al. (2019) found that the land class of the areas surrounding San Juan were significantly correlated with the abundance of existing *Ae. aegypti* populations. They based their study off of CDC trapping data using the BG-2 Sentinel mosquito traps at 59 locations in San Juan city from October to November, 2017. Although San Juan city is on the northern edge of Puerto Rico and Ponce is on the southern edge, the two areas do share similarities. The municipality of Ponce is the second most populated area of Puerto Rico, following the greater San Juan area. In both regions there exists a strong north-south urban concentration gradient, including elevation gains found as distance from the coast is increased. The five types of land class examined were high-density housing, low-density housing, forest, non-forest vegetation, and wetlands (Barrera et al. 2019). The linear modeling estimates for *Ae. aegypti* abundance that were found using the BG-S Sentinel trapping data per land class were: high-density

housing = 3.4 ± 0.4097 (females/trap/day), low-density housing = 9.8 ± 1.1 (females/trap/day), forest = 2.6 ± 1.0 (females/trap/day), non-forest vegetation = 2.7 ± 0.5 (females/trap/day), and wetlands = 2.1 ± 0.3 (females/trap/day) (Barrera et al. 2019).

To obtain existing mosquito population estimates within Ponce, I evaluated several landscape characteristics: the extent of the patchiness of the landscape, the existing wild type mosquito populations, and the human population of the study area. Since the *Ae. aegypti* mosquito preferentially feeds on humans, it is found primarily in and around human habitations, although the importance of non-dwelling buildings as sources of mosquito breeding habitat has been noted (Barrera et al. 2006). Because of the relative importance of any building type as potential for breeding and the ability of the *Ae. aegypti* mosquito to use highly cryptic containers as breeding habitat, I analyzed a GIS layer that included all of the buildings within the municipality of Ponce, available from the Humanitarian OpenStreetMap Team (2020). This shapefile was converted from polygons to points to show the locations of all buildings within Ponce. The flight and dispersal distance of *Ae. aegypti* within Puerto Rico was suggested to be within a maximum range of 100m (Harrington et al. 2005; Cox et al. 2007). However, the *Ae. aegypti* mosquito has been demonstrated as exhibiting leptokurtic dispersal. This right-tailed dispersal distribution means that although most mosquitoes do not travel very far over their lifetime, long distance dispersal has been seen (Harrington et al. 2005). To determine the extent of the habitat patches, I measured the distances between the major groupings of building points and characterized patches based on buildings being separated by either a major barrier (e.g. highways, or water channel) or by being separated by distances of over 100m. This resulted in there being 36 distinct habitat patches within the municipality of Ponce. The distances between patches were also used in conjunction with mark-recapture data gathered in Puerto Rico to estimate the probability that migration will occur between adjacent patches (Harrington et al. 2005).

Risk regions (RR1, RR2, and RR3) were developed by examining the flight distance between patches and the dominant landcover classification for each set of landscape patches that were included in each region (Figure 2). RR1 was predominantly high-density housing and was bounded by a navigable channel to the East, and a major highway to the North. RR2 was mixed high-density housing and forest and was bounded by a major highway to the South, the border of the municipality of Ponce to the East, and mountains to the North. RR3 was composed of mainly forest and patches of low-density housing and was characterized by the strong North-South series of valleys extending to the Northern border of the municipality.

Existing mosquito population estimates were obtained by combining the *Ae. aegypti* abundance estimates of the Barrera et al. (2019) study with my GIS analysis of the existing landscape characteristics of Ponce (Table 1). A shapefile available from the National Land Cover Database (NLCD 2016) was used to quantify the amount of the 5 major land classes used in the abundance models of the Barrera et al. (2019) study: high-density housing, low-density housing, forest, non-forest vegetation, and wetlands. A 100m buffer surrounding the building points was created and used to clip the landcover dataset so that the dominant land cover classification of each distinct patch was defined. Once each patch was defined by its dominant landcover classification, the amount of buildings per patch and the abundance equations per land class were used to estimate the maximum number of existing wild mosquito population for each patch (Barrera et al. 2019).

Table 1. Human population and maximum mosquito population estimate per risk region.

Risk Region	Human population (persons)	*Ae. aegypti* maximum mosquito population estimate (individuals)
RR1	113,430	80,200
RR2	38,629	23,200
RR3	14,170	10,200

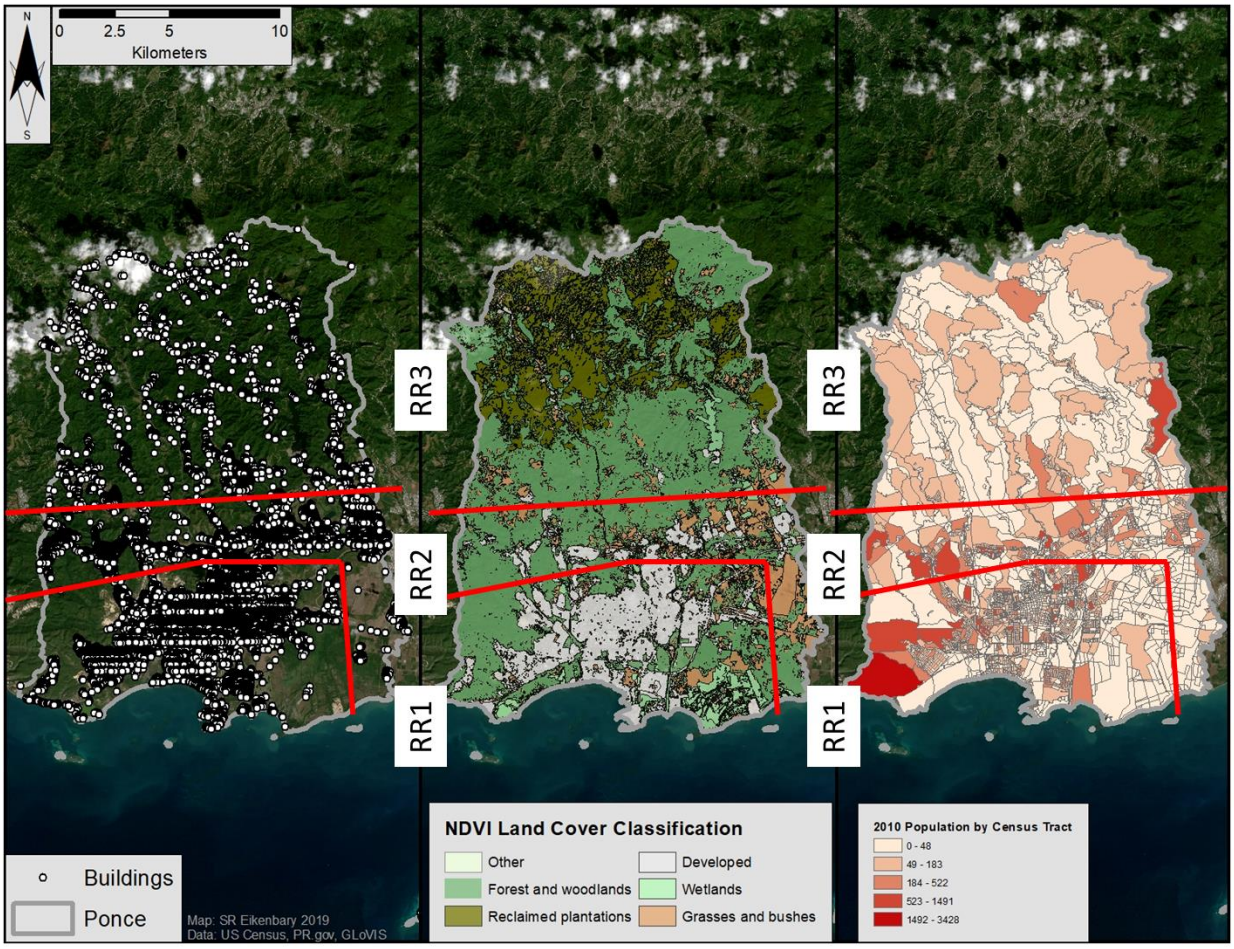

Figure 2. Map showing the locations of all buildings, habitat classification, and human populations within each risk region (RR1, RR2, and RR3) within the study area.

2.6 Disease transmission thresholds

Mathematical epidemiological models date back to the 1700s (Bernoulli 1760). This early work was further developed by Hamer (1906) to describe the mass action principle, that the rate of the spread of a contagion is dependent on the number of immunologically susceptible individuals within a population and the number of infected individuals. Kermack and McKendrick (1927) expounded upon these early models to describe the idea that a population density threshold within a population exists that limits the spread of a contagion, relative to the infectivity rate of the contagion. With the idea of threshold theory, an increase in contagious individuals or

vectors into a population does not lead to an epidemic unless the overall density of the immunologically naïve population is above a critical threshold (Kermack and McKendrick 1927; Focks et al. 2000).

2.7 Dengue transmission threshold

Focks et al. (2000) coupled the mathematical threshold theory with computer simulation models to develop threshold levels for dengue as a function of pupae per person, herd immunity (initial seroprevalence), temperature, and the magnitude of viral introduction (Table 2). Focks et al. (2000) describe the dengue thresholds in terms of pupae per person for 3 main reasons: pupae are readily available to count, the *Aedes* subgenus *Stegomyia* can be easily separated from other container inhabiting organisms, and because of the research correlating the number of pupae to the number of adult individuals.

Table 2. Dengue transmission thresholds (pupae/person) per temperature and level of herd immunity. Adapted from Focks et al. (2000).

Temperature	Herd Immunity	Dengue Transmission Threshold (pupae/person)
22	0	9.57
24	0	2.92
26	0	1.42
28	0	0.53
30	0	0.13
32	0	0.07
22	33	14.1
24	33	4.47
26	33	2.03
28	33	0.75
30	33	0.19
32	33	0.1
22	67	30.55
24	67	9.22
26	67	4.26
28	67	1.69
30	67	0.38
32	67	0.26

The temperature of the environment influences the number of adult female mosquitoes and the number of pupae (Focks et al. 2000). In general, as temperature increases the length of the gonotrophic cycle decreases, resulting in higher rates of emergence from pupae to adult (Focks et al. 2000). Temperature also influences the incubation period of the dengue virus. Infected mosquitoes become able to transmit dengue faster at higher temperatures (Focks et al. 2000).

2.8 Zika threshold

Barrera et al. (2017) developed transmission threshold levels for Chikungunya in Puerto Rico. Chikungunya is another viral disease that is vectored by the *Ae. aegypti* mosquito. The first reported case of chikungunya was confirmed in Puerto Rico in May of 2014. The first confirmed case of Zika virus in Puerto Rico occurred in 2015 (Barrera et al. 2017). Though chikungunya and Zika viruses are different families of viral infections they are both primarily spread by the same vector, both share similar transmission cycles, and both were introduced into a serologically naïve population (Barrera et al. 2017). The co-occurrence of dengue, chikungunya, and Zika virus suggests that the areas that have been impacted by dengue and chikungunya are also at risk of being impacted by Zika virus. Therefore, the use of chikungunya transmission thresholds would be prohibitive to the spread of Zika and an appropriate proxy for Zika transmission thresholds (Barrera et al. 2017). The Zika transmission thresholds estimated in Barrera et al. (2017) were developed through comparing mosquito trapping data in several areas in Puerto Rico and comparing the rates of antibody presence in the human populations. Although thresholds for Zika will vary with temperature similarly as with dengue thresholds, the Zika threshold values in the Barrera et al. (2017) study were developed and reported with no specificity to temperature (Focks et al. 2000; Barrera et al. 2017).

2.9 Case learning

Case learning can be used within the Bayesian networks created with Netica™ to automatically construct the conditional probability tables that describe the likely occurrences of a specific

variable (Norsys 2014). Case learning in Netica™ is the process where the software performs probabilistic inference on a set of case data where each entry (case) describes one outcome of an event. This process automatically generates the values for the conditional probability tables that describes the outcomes of the set of variables (nodes) included in the Bayesian network that cause the event (Norsys 2014). To parameterize nodes within the model using case learning, case files must be created that include each measurement of interest for each of the variables (nodes) that are being described. The case files I used for case learning within my Bayesian networks describe the outcomes of each of the gene drive release simulations. Input parameters include the resistance generation rate, the duration of releases (weeks), and the release ratio. The nodes that the conditional probability tables are parameterized using case learning are the 'Resulting Pop. After GD Release' and 'Proportion of GD Resistance' nodes. The case files included the allele frequencies and population numbers for each patch per time point.

2.10 Interaction with pesticide applications

The effects of pesticide application are modeled within the Bayesian network through the incorporation of dose-response modeling within the pesticide application node. Four pyrethroid pesticides and DDT were incorporated. The likely impacts of these pesticides on adult *Ae. aegypti* after the release of the gene drive can be shown by selecting the specific pesticide and amount to be applied (mols). The impacts on remaining populations are shown via the percent reduction due to the application of the various pesticides.

3. Model Structure

3.1 Conceptual model

A general cause-effect conceptual model adapted from Landis (2004) by NASEM (2016) was developed to map out the causal pathways that there is a mechanistic understanding of, that link the release of the gene drive engineered mosquitoes and the incidence of disease within

the study area (Figure 3). Each component of the conceptual model is linked to the study area through a spatial component. The conceptual model is composed of the sources of the stressors, the stressors themselves, the characteristics of the habitat, the direct and indirect effects, and the impacts to the defined endpoints.

Once the general conceptual model was created a simplified version was created that lists the specific causal pathways which includes only the items that I was able to quantify and were to be included within the Bayesian networks (Marcot et al. 2006; Chen and Pollino 2012; Landis et al. 2020). The simplification was done to only represent the cause-effect relationships of stressor to endpoint where data and research regarding the mechanistic interactions was available (Figure 4). A table of items included in the simplified conceptual model and the resulting Bayesian networks, as well as data sources, is included in appendix 5.

Source. The source column of the conceptual model lists the spatially explicit parameters of the release of the gene drive engineered *Ae. aegypti* mosquitoes. To introduce the gene drive engineered organism there will be multiple release points throughout the study area that are based on the migration rates between patches as estimated through GIS analyses in conjunction with known dispersal rates (Harrington et al. 2005). Other components describing the release of the gene drive into the target population include the number of organisms to be released, the timing of release, the interval of time between releases, and the overall amount of releases of the gene drive engineered *Ae. aegypti*.

Stressor. The stressor column includes the characteristics of the specific genetic construct that makes up the gene drive. The homing rate of the drive mechanism (the rate at which the CRISPR/Cas9 complex successfully finds the target sequence), the cutting rate of the drive mechanism (the rate of the CRISPR/Cas9 complex successfully causing a double-strand break in the segment of DNA), and the resistance allele generation rate are all components that will impact the rates of inheritance of the synthetic genetic construct and the outcomes of the

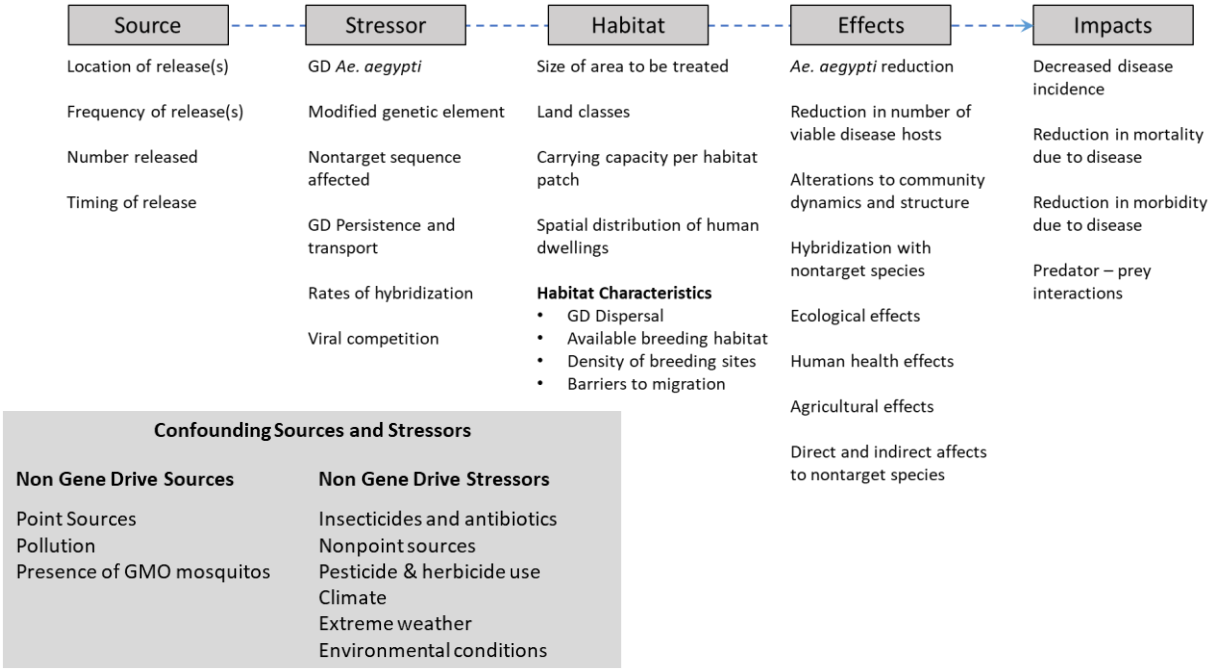

Figure 3. Conceptual model for the BN-RRM assessment of gene drive engineered *Ae. aegypti* as vectors to control dengue and Zika virus.

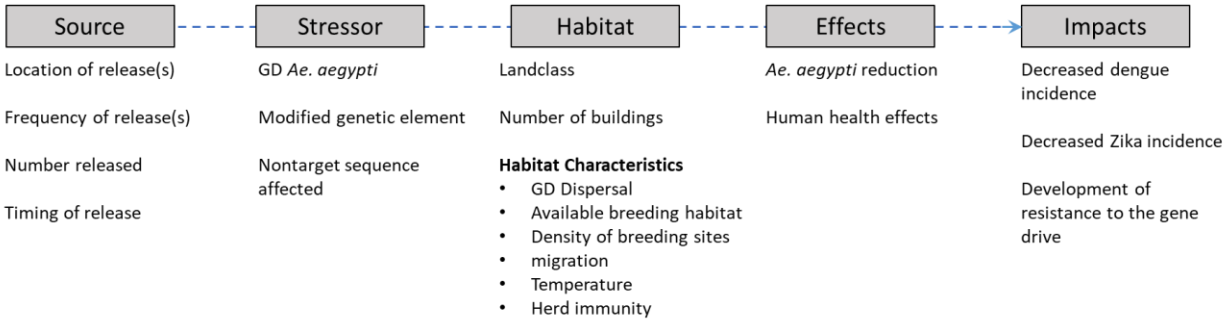

Figure 4. Simplified version of the conceptual model including only the items that are incorporated into the Bayesian network. Incorporation into the Bayesian network structure is dependent on mechanistic understandings of the causal interactions taking place between items on the conceptual model (Chen and Pollino 2007; Marcot 2017; Landis et al. 2020).

release of the gene drive into the target population. Other stressors within the study area include the pesticides that are currently approved for use and other current mosquito

management efforts being conducted that may impact the dispersal of the gene drive. For example, extensive pesticide application could result in the engineered mosquitoes being killed before they can successfully mate and pass on the gene drive to their progeny, or extensive reductions in the amount of standing water around habitations would result in a lack of breeding habitat, potentially hindering the breeding success and dispersal of the gene drive within the target population.

Habitat. The land classes and land use within the study area will be listed in the habitat column. The key component here is to understand the exposure of the target population to the gene drive. This will include the abundance of the existing wild type mosquitoes, the productivity of each associated land class, and the spatial extent of where the mosquitoes breed and forage.

Effect. The effects include the direct and indirect effects resulting from the release of the gene drive *Ae. aegypti* into the environment. Effects of the release will include the relative frequency of the gene drive construct in the population, any changes in population size and population dynamics, and any resulting fitness cost that may result from hybridization between wild and gene drive mosquitoes. Agricultural and ecological effects resulting from any reduction in mosquito populations as a result of the drive are not to be excluded and remain as items within this more comprehensive conceptual model.

Impact. The impact to our defined endpoints resulting from the release of the gene drive into the environment will include the number of viable disease vectors, rates of disease incidence, and any reduction in disease morbidity and mortality. Alterations to ecological structure or food web interactions, as well as any reduction of agricultural productivity resulting from release of the gene drive are included as potential impacts associated with the agricultural and ecological effects pathways that are generalized within the comprehensive conceptual model.

3.2 Development of the gene drive specific Bayesian network relative risk model

The conceptual model provides the basic framework for the development of our Bayesian network and reflects the cause-effect relationships between the variables that are included (Ayre and Landis 2012; see Figures 4 & 5). The underlying mathematical driving force of the Bayesian network is the relationships between variables that are described within the conditional probability tables for each distinct node within the network (Cain et al. 1999; Marcot et al. 2006). Each conditional probability table lists the likely outcomes of each combination of input variables in terms of probabilities.

The probability distributions defining the likelihood of outcomes to *Ae. aegypti* populations as a result of the simulated releases of the gene drive engineered mosquitoes will be developed from output from the Mosquito Gene Drive Explorer Model (MGDrivE) and is described in detail below (Sánchez et al. 2019). Landscape data resulting from GIS analyses were incorporated to fill out the nodes that represent source and habitat characteristics. Dengue transmission threshold levels from Focks et al. (2000) were used to develop the conditional probability table for the reduction in dengue node. The mosquito abundance estimates resulting from the gene drive dispersal model and the human population densities of the risk regions were used to determine the likelihood of dengue transmission in terms of pupae per person. Zika thresholds were developed in Barrera et al. (2017) in terms of female mosquitoes per building and were used with the mosquito abundance estimates resulting from the gene drive propagation modeling output to determine the relative number of female mosquitoes per building per region which resulted in the risk estimates for Zika transmission.

3.3 Modeling of gene drive propagation

To model the release of gene drive engineered *Ae. aegypti* into Ponce, Puerto Rico I used the Mosquito Gene Drive Explorer (MGDrivE) model created by the Marshall Lab at UC Berkeley

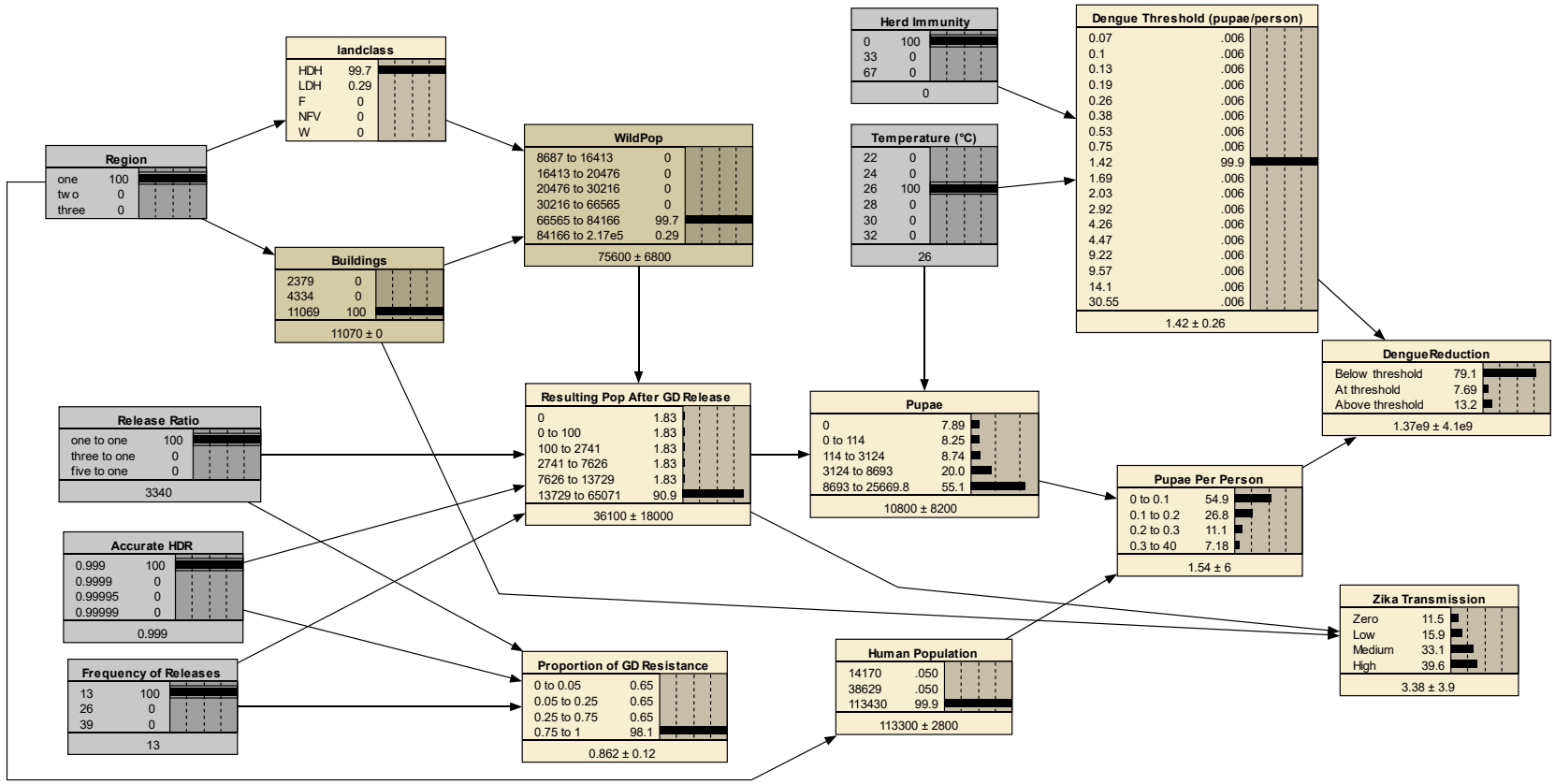

Figure 5. Bayesian network model structure for risk estimation of gene drive releases at the maximum simulation time (6.75 years). The parameters selected are for RR1, a release ratio of 1:1, rate of accurate HDR of 0.999, and 13 weekly releases. Dengue transmission thresholds are based on 0% herd immunity and a temperature of 26°C (Focks et al. 2000).

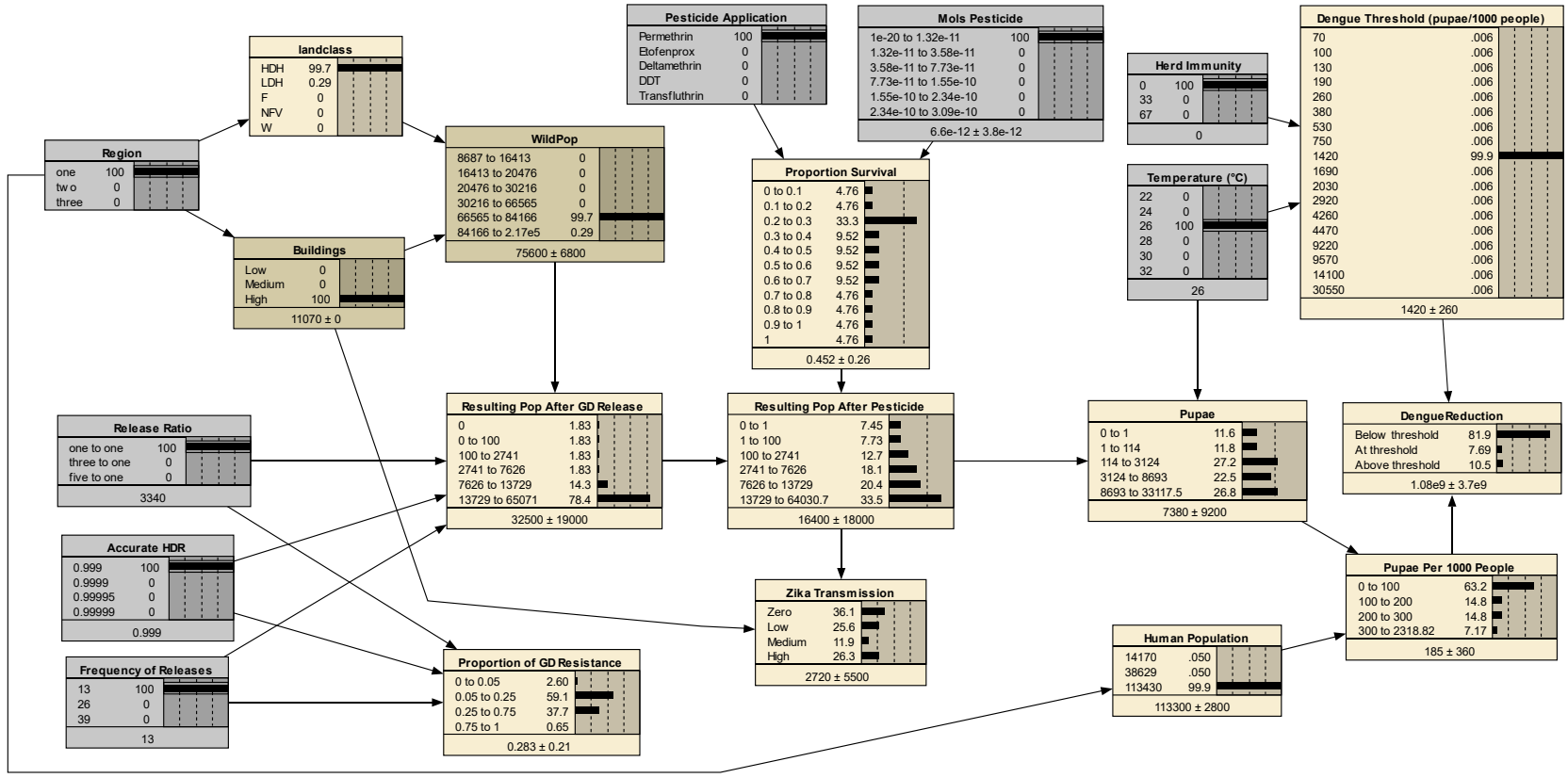

Figure 6. Bayesian network model structure for risk estimation of gene drive releases at the time point to minimum overall population per release scenario. The parameters selected are for RR1, a release ratio of 1:1, rate of accurate HDR of 0.999, and 13 weekly releases. Dengue transmission thresholds are based on 0% herd immunity and a temperature of 26°C (Focks et al. 2000). Pesticide application as an additional management strategy is included.

(Sánchez et al. 2019). MGDrivE is a spatially explicit mathematical simulation framework that models the deployment of gene drive engineered mosquitoes and their effect on population dynamics. The MGDrivE modeling framework was designed to model the spread of a variety of gene drive architectures through spatially explicit mosquito populations. MGDrivE incorporates lumped age-class ecological modeling with genetic information and site-specific spatial dynamics. MGDrivE allows the user to input specific life-history parameters of the target population including the developmental stages: egg, larvae, pupae, and adult. Patches within the simulated release landscape are represented with a metapopulation structure, with used defined rates of migration between patches. Within MGDrivE, genetic inheritance of gene drives is modeled using 'inheritance cubes' that describe the maternal and paternal genetic information which results in the genetic composition of any offspring. The MGDrivE modeling framework allows the user to define specific characteristics of the gene drive that is being simulated. These gene drive characteristics include the correct homing rate and the rates of resistance allele generation (Sánchez et al. 2019). Rates of both in-frame and out-of-frame resistance (Marshall et al. 2017; Unckless et al. 2017) are able to be modeled within MGDrivE. In-frame resistance is resistance generation that affects only the functionality of the gene drive, with no related fitness reduction as a result. Out-of-frame resistance imparts resistance to the gene drive while also affecting some other conserved portion of DNA resulting in a substantial fitness reduction (Sánchez et al. 2019).

Mosquito life history is modeled within MGDrivE by a lump age-class model, in which the different life stages of the mosquito as incorporated: egg, larvae, pupae, and adult (Sánchez et al. 2019). Density dependence is modeled in the larval stage to account for larval competition within breeding sites. Density independent mortality is modeled in the juvenile and adult life stages. Spatial dynamics of the simulated release area are incorporated through a migration matrix that lists the probability of an individual in one patch migrating to an adjacent patch at

each timestep (day). The population of each distinct patch is simulated as randomly mating (Sánchez et al. 2019).

MGDrivE can be run as a deterministic or stochastic modeling framework, with the stochastic simulations accounting for random chance events that may occur when patch densities become low. Monte Carlo iterations of each simulated release can be implemented to capture the full range of possible outcomes as a result of the simulated releases (Sánchez et al. 2019).

3.4 Simulation parameters

I ran 36 different stochastic simulation scenarios for various deployments of gene drives using the MGDrivE modelling framework (Sánchez et al. 2019). The parameters were primarily focused on the initial release parameters of the engineered organisms, the homing rate of the gene drive construct, and the rate of resistance allele generation. I accounted for the spatial uniqueness of the landscape of Ponce by incorporating estimated migration rates into the migration matrix within the model. Each stochastic simulation was repeated 50 times to ensure that the range of outcomes was accounted for.

To determine the patches that received the simulated weekly releases of gene drive engineered *Ae. aegypti*, I developed a matrix of migration rates based on GIS analyses of mosquito dispersal and flight distance relative to spatial locations of patches (Harrington et al. 2005). Patches with less than 4% migration probability to adjacent patches were chosen as release patches, as well as the patches that made up our primary “mainland” patch. Too account for the right-tailed skew in dispersal distances noted in previous studies (Harrington et al. 2005), patches that were beyond the 100m range but within a reasonably unimpeded flight path were assigned the minimum probability of migration (1%). This resulted in 24 patches designated as release patches. The study area represents a patchy metapopulation of wild *Ae. aegypti* (Deines et al. 2005). This release scheme is comparable to the introduction of an invasive

species and was chosen so that the released gene drive organisms would have a high probability of radiating throughout the study area. Because of the patchiness of the study area multiple release locations were chosen to increase the probability of the spread of the introduced gene drive engineered *Ae. aegypti*. This reduces the probability of chance occurrences that would result in certain patches being effectively cut off from the spread of the gene drive due to low migration rates (Deines et al. 2005). If the gene drive does not spread throughout the target population and reduce the vector densities there will be no impacts to disease transmission.

The parameters surrounding the initial release of the gene drive engineered *Ae. aegypti* were the number of engineered organisms released, the locations of releases, the interval of time between releases, and the duration of the scheduled releases (Table 3).

Table 3. MGDrivE release parameters. Each combination of parameters was modeled giving 36 individual release scenarios that were repeated 50 times per simulation.

Rate of accurate homology-directed repair	Release ratio	Duration of releases (weeks)
0.001	1:1	13
0.0001	3:1	26
0.00005	5:1	39
0.00001		

The release ratios were chosen to be similar to previous modeling efforts (Robert et al. 2014; Pham et al. 2019; and Sánchez et al. 2020). These previous studies modeled the release ratios of gene drive engineered organisms to wild type organisms. Ratios included ranged from 1:1 to 10:1. I chose an intermediate range of simulated release ratios, including 1:1 (3,340 males released), 5:1 (10,021 males released), and 5:1 (16,702 males released). These release ratios show the effect of the gene drive resulting from a range of inundation in terms of release ratio.

A model burn-in period of 1 year was used to allow for the simulated wild mosquito populations per patch to stabilize due to the various migration probabilities between adjacent patches. The

different release durations that I included were 13, 26, and 39 weeks, simulating 1 release every week starting on simulation day 366. These release durations were chosen based on previous modeling efforts which had similar numbers of weekly releases (Robert et al. 2014; Pham et al. 2019; and Sánchez et al. 2020).

The total number of simulated gene drive engineered mosquitoes was the product of variations in release ratio (1:1, 3:1, and 5:1), duration of weekly releases (13, 26, and 39 weeks), and the number of patches receiving releases (24 patches). This resulted in 9 separate values for total number of engineered male *Ae. aegypti* released over the simulation periods. Complete range of released engineered mosquitoes over entire simulation period by release scenario is given in table 4.

The successful homing rate of the simulated gene drive was assumed to be 95%, which is the rate demonstrated in Hammond et al. (2016). The rate of accurate homology directed repair (HDR) was simulated at levels of: 0.999, 0.9999, 0.99995 and 0.99999. The inverse of these rates was used to simulate the resistance allele generation developing as a result of the drive, giving us resistance generation rates ranging from 0.001 to 0.00001. The gene drive construct described in Hammond et al. (2016) designed for population suppression of the malaria vector *Anopheles gambiae* showed rates of resistance generation of approximately 0.0013. Marshall et al. (2017) modeled the effects of decreasing the rates of resistance allele generation through multiplexing (from 10^{-3} to 10^{-7}) to determine the maximum size of the wild populations of mosquitoes that could be suppressed for each rate of resistance generation. I chose to model the rates of resistance generation of gene drive constructs that have been lab tested (Hammond et al. 2016) and a moderate level of resistance proposed by Marshall et al. (2017) through the theoretical use of multiplexing.

To develop the amount of Monte Carlo iterations used, I ran preliminary MGDrivE simulations, increasing the value of repetitions. When comparing the results, I found that n=50 was suitable

to capture the range of possible outcomes and to stabilize the resulting outcome distributions to convergence. Time series plots of the results showed convergence was well established after approximately 1,500 simulation days. Specific statistical measures of the mean, standard deviation, and 95% confidence intervals of outcomes were not included and the value was based largely on visual interpretation of convergence within the preliminary results.

Table 4. Total numbers of gene drive engineered *Ae. aegypti* released over the entire simulation period for each combination of release parameters. A total of 24 out of 36 patches were chosen to receive simulated releases based on migration rates between patches. If a patch had less than 4% probability of migration it was chosen as a release patch.

Release ratio	Duration of releases (weeks)	# of patches receiving releases	Total GD released during simulation period
1:1	13	24	1,042,080
1:1	26	24	2,084,160
1:1	39	24	3,126,240
3:1	13	24	3,126,552
3:1	26	24	6,253,104
3:1	39	24	9,379,656
5:1	13	24	5,211,024
5:1	26	24	10,422,048
5:1	39	24	15,633,072

3.5 Model analysis

To understand the full range of outcomes associated with the use of population suppression drives I modeled 3 separate management strategies with Bayesian networks: 1) no gene drive releases to model the maximum mosquito population estimate and current risk estimates of the

landscape, 2) the absolute outcome of the gene drive release (Figure 7) at the end of the simulation period, & 3) the time point associated with the largest decrease in population size resulting from the various gene drive release simulations that could be used in conjunction with other adaptive management strategies (Figure 8). Modeling the maximum mosquito population estimate without releases provides a scenario in which I can compare with the remaining two release scenarios. The maximum time scenario provided the likelihood of resistance to the drive developing, and the likelihood of population rebounds as a result of only the releases of the gene drive. The third Bayesian network was built to simulate how non-eradication resulting from the gene drive releases could still be used as part of an integrative pest management strategy within an adaptive management scheme.

3.6 Maximum mosquito population simulations

The maximum mosquito population estimate release scenario was set using the landscape characteristics of Ponce to determine the wild type population size, as well as the rates of migration between patches. Fifty stochastic simulations were run, each simulating 6.75 years, to evaluate the range of possible wild type *Ae. aegypti* populations within the study area under non-intervention scenarios. This modeling effort represents the base mosquito population levels that would occur using realistic landscape characteristics. The population densities of mosquitoes per region, coupled with the human population densities described the current likelihood of disease transmission (Focks et al. 2000; Barrera et al. 2017).

3.7 Gene drive deployment simulations

The second Bayesian network created shows the effectiveness of the release of the gene drive after 6.75 years (Figure 7). This illustrates how the various release parameters interact to impact mosquito populations within the study area. This modeling effort shows the outcomes of the gene drive release, including: the overall efficacy of the gene drive in terms of extinction

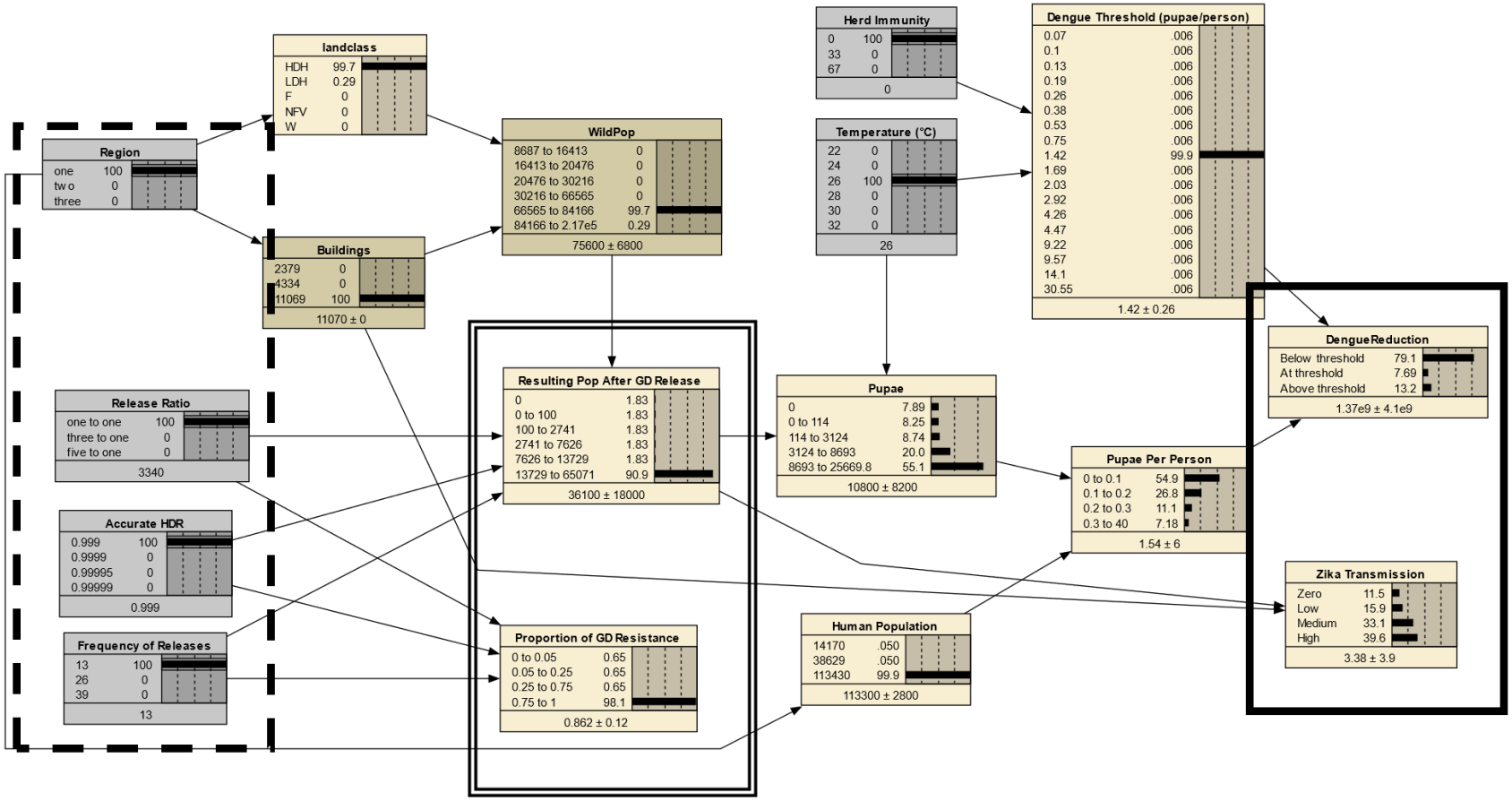

Figure 7 Bayesian network showing the results of one of the 36 MGDrivE release scenarios within RR1 at the end of the simulation period. Release ratio is set to 1:1, the rate of accurate HDR is set to 0.999 to simulate the Hammond et al. (2016) gene drive, duration of releases is set to 13 weeks. The dashed line shows the release parameters and risk region selected. The double-solid line shows the outcomes of the gene drive release and the percentage of surviving mosquitoes that are resistant to the gene drive. The solid line shows the risk estimates to the transmission rates of dengue and Zika. Herd immunity and temperature are selected within this example to represent dengue transmission thresholds at 0% herd immunity within the human population, at 26°C.

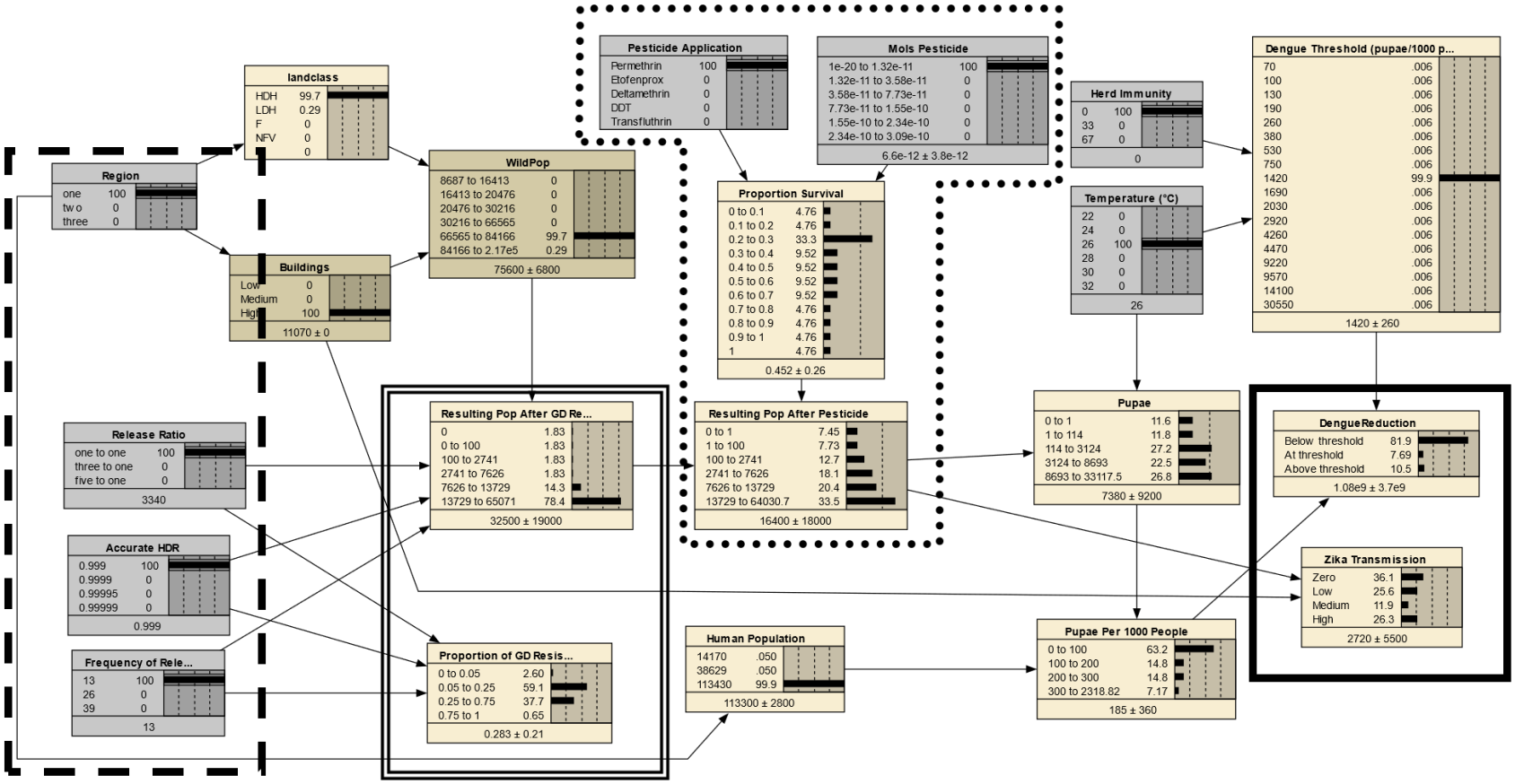

Figure 8. Bayesian network showing the results of one of the 36 MGDrivE release scenarios within RR1 at the time point associated with the average lowest possible *Ae. aegypti* population. Release ratio is set to 1:1, rate of accurate HDR is set to 0.999 to simulate the Hammond et al. (2016) gene drive, and the duration of releases is set to 13 weeks. The dashed line shows the release parameters and risk region selected. The double-solid line shows the outcomes of the gene drive release and the percentage of surviving mosquitoes that are resistant to the gene drive. The solid line shows the risk estimates to the transmission rates of dengue and Zika. The circular outline shows the inclusion of pesticide application as an additional management strategy. Herd immunity and temperature are selected within this example to represent dengue transmission thresholds at 0% herd immunity within the human population, at 26°C.

occurrence, the likelihood of wild population rebounds, the likelihood of gene drive resistance developing within the mosquito population resulting in either mixed (wild and resistant) or resistant population rebounds, and any impacts on disease transmission due to lowered abundance of the vector population. This means that the impact to mosquito population abundance can be modeled

3.8 Gene drive deployment with alternate management strategies

To show the applicability of the use of gene drives within an adaptive management framework, the third Bayesian network integrates a second management intervention taken at the average timepoint where the average mosquito populations were lowest within the landscape due to the release of the gene drive (Figure 8). To find the time to minimum population for each of the 36 different release scenarios, the model output from MGDrivE was aggregated per release scenario to estimate the timepoint that the average number of mosquitoes within the entire study area was at its lowest per release scenarios. Using this subset of the simulation results I was able to show the maximum effectiveness, in terms of overall population reduction, of each gene drive release and used this to develop the best-case scenario that could be used to guide the implementation of alternative mosquito control measures (e.g. pesticides). This approach was used because model results showed that extinction is not likely, resistance is highly likely, and population rebounds were suggested to occur with high probability for many of the MGDrivE release scenarios. Resistance is almost assured to develop as a result of the non-homologous end-joining repair mechanism and because of this the likelihood of eradication decreases with increased resistance generation rates. If I only examined the results at the end of the simulation period, I would only have developed binary results (either the gene drive caused extinction or did not) and these results would have had little to no relevance to the human health endpoints (either the disease was eradicated along with the mosquitoes or not).

3.9 Case learning simulation output

Case learning was used to construct the conditional probability tables for the nodes within the Bayesian networks that describe the propagation of the gene drive throughout the mosquito population. The MGDrivE model output was restructured in R to show the numbers of wild and resistant mosquitoes per region for each of the different release scenarios. This data was then used as case learning data set that Netica™ (Norsys 2014) used to determine the likely outcomes of each different combination of release parameters. The case learning dataset incorporated each of the fifty Monte Carlo simulations for each of the 36 simulated release combinations and were used to model the three management strategies: 1) no intervention, 2) the release of gene drive engineered *Ae. aegypti* only, and 3) release of gene drive *Ae. aegypti* and a secondary management intervention. The MGDrivE model output and case learning approach were used to build the conditional probability tables for the nodes in the Bayesian network that are associated with wild type populations, the resulting populations after the gene drive release, and the rates of resistance developing as a result of the gene drive release.

With case learning, the software uses the observed output and automatically determines the relationships between the different input variables to give the probability of different outcomes for each combination of input states (Norsys 2014). With large sets of simulation output, like I developed with MGDrivE, Netica™ can use case learning to perform probabilistic inference on the model output, where each entry describes one outcome. Netica™ used the observed MGDrivE output to determine the relationships between the input variables (Norsys 2014). This process automatically calculates the values for the conditional probability tables that describe the probability distributions for the nodes relating to the release of the gene drive and the impacts to the mosquito population after the releases.

3.10 Sensitivity Analysis

Model sensitivity (mutual information or entropy reduction) describes the amount of influence that changes in one node can assert on the other nodes within the Bayesian network (Norsys 2014). This means that the probability distributions of inputs nodes can have more or less influence on the probability distributions of the endpoint nodes. The built-in model sensitivity analysis within Netica™ provides a framework to understand how the various input nodes affect the endpoint nodes. The entropy reduction for each endpoint node can vary depending on the location of each node in the Bayesian network being analyzed (Norsys 2014). The nodes I examined for entropy reduction were the input nodes describing the release parameters of the gene drive, the endpoint nodes for resistance development, dengue and Zika transmission, the intermediate node for resulting population after the release of the gene drive, and temperature as it is strongly related to the dengue transmission thresholds developed by Focks et al. (2000).

4. RESULTS

4.1 Maximum simulation time

To examine what the end result of the release of gene drive engineered *Ae. aegypti* within the study area would be, I analyzed the results after our 6.75-year simulations (Table 5). The simulation outcomes gave results for each release scheme. Out of all of the simulations only 1 resistance generation rate resulted in zero extinctions. The simulations modeling the release of gene drive mosquitoes modeled after the Hammond et al. (2016) construct with a resistance generation rate of 0.001 never resulted in extinction for any combination of release scenarios. All other release scenarios showed some level of extinction occurring, but this varied with the combinations of the resistance rates, number of weekly releases, and the wild to engineered release ratios. Generally, as rates of resistance decreased or ratio and duration increased, the likelihood of extinction increased, but not always in a monotonic relationship (Table 5). Of all the simulations examined for the maximum simulation time that did not result in extinction, the

surviving mosquito population was composed of individuals resistant to the gene drive due to resistance developing through non-homologous end-joining (Unckless et al 2017). General trends show that increasing the resistance rate, number of releases, or the release ratio resulted in a higher likelihood of extinction occurrence (Table 5).

For a modeled resistance rate of 0.00001, extinction occurred 4 out 50 times at a release ratio of 1:1 and 26 weekly releases and 3 out of 50 for 39 weekly releases. Thirteen weekly releases resulted in extinction in only one of 50 simulations at a 3:1 release ratio. Both of the 26 week and 39-week release scenarios of a 0.0001 resistance rate and 3:1 ratio resulted in 15 extinction events out of 50. Increasing the release ratio to 5:1 was most successful at 39 weekly releases and caused extinction in 50% of simulations (Table 5).

When the resistance rate was decreased to 0.00005 the scenarios with 26 weekly releases and a 1:1 or 3:1 release ratio were more successful than the 13 or 39 weekly releases. Scenarios modeled with a resistance rate of 0.00005 and a 5:1 release ratio caused extinction in 20 simulations. The highest degree of successfully causing extinction at a resistance rate of 0.00005 occurred with 26 weekly releases at a 3:1 release ratio (50%) (Table 5).

Further reducing the resistance generation rate to 0.00001 resulted in the highest likelihood of extinction occurring overall. Increasing each release parameter increased the likelihood of extinction with the highest occurrence of 47/50 simulations with a 5:1 release ratio and 39 weekly releases. The shortest duration of 13 weekly releases and the lowest release ratio of 1:1 resulted in extinction occurring in 29 out of 50 simulations (Table 5).

4.2 Time to minimum population size

The results of aggregating the data for the time point coinciding with the lowest overall population levels show that all release scenarios drastically reduced the population of mosquitoes. Resistance was strongly selected for in the population and the simulations had

Table 5. Results of MGDrivE maximum simulation time release scenarios. In general, a decrease in resistance generation or an increase in release ratio or duration of releases increased the probability of extinction occurring as a result of the gene drive.

Release ratio			Rate of accurate HDR			Duration of Release		
1:1 (3,340)			0.001			13 weeks		
3:1 (10,021)			0.0001			26 weeks		
5:1 (16,702)			0.00005			39 weeks		
			0.00001					
Resistance	Ratio	Duration	Resistance	Ratio	Duration	Resistance	Ratio	Duration
0.0001	1:1	13 weeks	0.0001	1:1	26 weeks	0.0001	1:1	39 weeks
Extinct	0		Extinct	4		Extinct	3	
Rebound	50		Rebound	46		Rebound	47	
0.0001	3:1	13 weeks	0.0001	3:1	26 weeks	0.0001	3:1	39 weeks
Extinct	1		Extinct	15		Extinct	15	
Rebound	49		Rebound	35		Rebound	35	
0.0001	5:1	13 weeks	0.0001	5:1	26 weeks	0.0001	5:1	39 weeks
Extinct	4		Extinct	10		Extinct	25	
Rebound	46		Rebound	40		Rebound	25	
0.00005	1:1	13 weeks	0.00005	1:1	26 weeks	0.00005	1:1	39 weeks
Extinct	6		Extinct	15		Extinct	13	
Rebound	44		Rebound	35		Rebound	37	
0.00005	3:1	13 weeks	0.00005	3:1	26 weeks	0.00005	3:1	39 weeks
Extinct	13		Extinct	25		Extinct	22	
Rebound	37		Rebound	25		Rebound	28	
0.00005	5:1	13 weeks	0.00005	5:1	26 weeks	0.00005	5:1	39 weeks
Extinct	20		Extinct	20		Extinct	20	
Rebound	30		Rebound	30		Rebound	30	
0.00001	1:1	13 weeks	0.00001	1:1	26 weeks	0.00001	1:1	39 weeks
Extinct	29		Extinct	34		Extinct	42	
Rebound	21		Rebound	16		Rebound	8	
0.00001	3:1	13 weeks	0.00001	3:1	26 weeks	0.00001	3:1	39 weeks
Extinct	31		Extinct	39		Extinct	44	
Rebound	19		Rebound	11		Rebound	6	
0.00001	5:1	13 weeks	0.00001	5:1	26 weeks	0.00001	5:1	39 weeks
Extinct	35		Extinct	43		Extinct	47	
Rebound	15		Rebound	7		Rebound	3	

varying degrees of resistance development. General trends in the data show that increasing the release ratio from 1:1, 3:1, and 5:1 for each combination of resistance development rate and number of weekly releases gave an increased reduction in overall mosquito population numbers (Table 6).

The target reduction in *Ae. aegypti* populations for the Bayesian network modeling only the release of the gene drive was 100%. Complete eradication because of the gene drive deployment would be necessary to avoid a population rebound of the disease vector with newly developed resistance to the gene drive deployed. The target reductions for the Bayesian network modeling the use of the gene drive and pesticide application as an additional management strategy is based on the percent reduction required for zero Zika transmission. The target reduction was 94% for RR1, 93% for RR2, and 92% for RR3. The percent reductions would eliminate the spread of any novel disease, like Zika virus, as well as be prohibitive of dengue transmission.

The lowest percent reduction resulting from the parameters resembling the Hammond et al. (2016) construct for all levels of scenarios was 76.75% (release ratio = 1:1, weekly releases = 13). The largest percent reduction of 97.76% followed 26 weekly releases at a 5:1 release ratio. Increasing the release ratio increased the percent reduction in population numbers for scenarios that included 13 and 26 weekly releases. The opposite was true for scenarios simulating 39 weekly releases; 97% reduction at 1:1 to 91.26% reduction at 5:1. Of the surviving mosquitoes, the % remaining that were resistant to the gene drive varied but was most consistent when grouped by weekly releases. Percent resistant at 13 weekly releases were 35.97% at 1:1, 33.06% at 3:1, and 32.95% at 5:1. The percent resistant at the 26 weekly release level were 37.76% at 1:1, 33.80% at 3:1, and 36.97 at 5:1. The percent of the total surviving mosquito population at the 39 weekly release level were 38.18% at 1:1, 44.04% at 3:1, and 45.12% at 5:1 weekly releases (Table 6).

When decreasing the resistance generation rate of simulations to 0.0001, the lowest percent reduction in populations of 88.97% occurred with 13 weekly releases and a 1:1 release ratio. This release scenario of 13 weekly releases at a 1:1 ratio also resulted in the highest percentage of surviving mosquitoes being resistant to the gene drive; 44.57% of remaining *Ae. aegypti* were resistant. No extinction events resulted from this release scheme. The largest reduction in simulated populations of 99.43% of the maximum mosquito population estimate resulted from 26 weekly releases at a 5:1 release ratio. The release scenario that resulted in the lowest percent of remaining mosquitoes being resistant to the gene drive (23.38% resistant) occurred with 13 weekly releases of a 5:1 release ratio (Table 6).

When further decreasing the resistance generation rate to 0.0005, the scenario that resulted in the lowest percent reduction of mosquito populations (90.75%) and highest percent of surviving mosquitoes being resistant to the gene drive occurred with 13 weekly releases at a 1:1 release ratio. The largest reduction in mosquito populations occurred with 26 weekly releases at a 5:1 release ratio. The lowest percentage of surviving mosquitoes being resistant to the drive was 29.04% and resulted from 26 weekly releases at a 1:1 release ratio (Table 6).

The lowest resistance generation rate simulated was 0.00001. Thirteen weekly releases and a 1:1 release ratio at this level of resistance generation resulted in a 95.51% reduction in *Ae. aegypti* populations. The largest reduction in population levels of 99.74% occurred with 26 weekly releases at a 5:1 release ratio. This was the largest percent reduction for any simulation across all release parameters. Not surprisingly, this release scenario with the lowest resistance generation rate also resulted in the least percentage of surviving mosquitoes being resistant to the gene drive; 12.64% of surviving mosquitoes were resistant following 13 weekly releases of a 5:1 release ratio. The largest release ratio and longest duration of releases at this resistance generation rate resulted in 58.33% of surviving mosquitoes to be resistant to the gene drive but also had the highest probability of extinction at 94% of simulations (Table 6).

Table 6. Results of MGDrivE minimum simulation time release scenarios.

Resistance rate	Release ratio	Weekly releases	% Reduction	% Resistant
0.001	1:1	13	76.75	35.97
0.001	3:1	13	89.94	33.06
0.001	5:1	13	94.10	32.95
0.001	1:1	26	88.20	37.76
0.001	3:1	26	96.89	33.80
0.001	5:1	26	97.76	36.97
0.001	1:1	39	93.77	38.18
0.001	3:1	39	93.51	44.04
0.001	5:1	39	91.26	45.12
0.0001	1:1	13	88.97	44.57
0.0001	3:1	13	96.19	35.26
0.0001	5:1	13	97.51	23.38
0.0001	1:1	26	94.79	31.94
0.0001	3:1	26	99.04	43.35
0.0001	5:1	26	99.43	36.04
0.0001	1:1	39	97.90	40.70
0.0001	3:1	39	98.65	42.70
0.0001	5:1	39	98.30	43.06
0.00005	1:1	13	90.75	48.99
0.00005	3:1	13	96.24	41.04
0.00005	5:1	13	97.97	29.44
0.00005	1:1	26	96.03	29.04
0.00005	3:1	26	98.82	32.38
0.00005	5:1	26	99.42	43.50
0.00005	1:1	39	98.04	31.41
0.00005	3:1	39	98.89	46.04
0.00005	5:1	39	98.29	48.63
0.00001	1:1	13	95.51	55.12
0.00001	3:1	13	97.66	25.63
0.00001	5:1	13	98.10	12.64
0.00001	1:1	26	97.83	17.68
0.00001	3:1	26	99.57	17.05
0.00001	5:1	26	99.74	43.20
0.00001	1:1	39	99.56	52.20
0.00001	3:1	39	99.71	53.93
0.00001	5:1	39	99.63	58.33

4.3 Change in risk to human health endpoints

4.3.1 Maximum simulation time

The Bayesian network assembled for gene drive deployment in Ponce calculated risk of dengue and Zika transmission based on spatial inputs and MGDrivE simulation outcomes. Risk estimates described here reflect the pupal development rates and dengue transmission thresholds described in Focks et al. (2000) for herd immunity levels of 0 and for temperatures of 26°C (average temperature of Ponce is 26.2°C). The risk of dengue transmission at maximum mosquito population estimate prior to simulated deployment of gene drive mosquitoes for the entire study area was estimated to be a 60.6% probability of below transmission threshold, a 7.69% probability at transmission threshold, and a 31.7% probability of being above the transmission threshold for the entire study region (Table 7). Overall risk of Zika transmission at maximum mosquito population estimate for all risk regions was estimated at 78.6% probability of being high risk (Table 8).

Risk estimates for dengue transmission (Appendix 1) at maximum mosquito population estimate for RR1 and RR2 had the highest probability of being below transmission threshold based on pupae/person calculations (RR1 = 79.1%, RR2 = 59.5% probability of being below transmission threshold). RR3 was the only risk region that was at risk of dengue transmission at maximum mosquito population estimate; 43.3% probability of being below threshold, 7.69% probability of being at the threshold, and 49.1% probability of being above the dengue transmission threshold. When examining the outcomes of the release of gene drive *Ae. aegypti* to reduce dengue transmission without any other interventions, the risk distributions were identical for some scenarios. Resistance rates of 0.001 had a distribution similar to that of the maximum mosquito population estimate simulations (probability of being below transmission threshold: RR1 = 79.1%, RR2 = 58.1%, RR3 = 40.0%). There was no change in probability distributions until the release scenario modeling a resistance generation rate of 0.0001, release ratio of 5:1, and 13

Table 7. Dengue transmission risk estimates for maximum simulation time scenarios. CC signifies maximum mosquito population estimate. At transmission threshold signifies that dengue transmission is possible based on pupae densities. Above transmission threshold signifies that dengue transmission is likely based on pupae densities.

Rate of accurate HDR	Release ratio	Weekly releases	Below transmission threshold (%)	At transmission threshold (%)	Above transmission threshold (%)
CC	CC	CC	60.63	7.69	31.70
0.999	3340	13	59.07	7.69	33.23
0.999	10021	13	59.07	7.69	33.23
0.999	16702	13	59.07	7.69	33.23
0.999	3340	26	59.07	7.69	33.23
0.999	10021	26	59.07	7.69	33.23
0.999	16702	26	59.07	7.69	33.23
0.999	3340	39	59.07	7.69	33.23
0.999	10021	39	59.07	7.69	33.23
0.999	16702	39	59.07	7.69	33.23
0.9999	3340	13	59.07	7.69	33.23
0.9999	10021	13	59.37	7.69	32.93
0.9999	16702	13	60.37	7.69	31.97
0.9999	3340	26	60.37	7.69	31.97
0.9999	10021	26	63.90	7.69	28.40
0.9999	16702	26	62.30	7.69	30.00
0.9999	3340	39	60.03	7.69	32.27
0.9999	10021	39	63.90	7.69	28.40
0.9999	16702	39	67.13	7.69	25.17
0.99995	3340	13	61.00	7.69	31.30
0.99995	10021	13	63.27	7.69	29.03
0.99995	16702	13	65.53	7.69	26.80
0.99995	3340	26	63.90	7.69	28.40
0.99995	10021	26	67.13	7.69	25.17
0.99995	16702	26	65.53	7.69	26.80
0.99995	3340	39	63.27	7.69	29.03
0.99995	10021	39	66.17	7.69	26.13
0.99995	16702	39	65.53	7.69	26.80
0.99999	3340	13	68.43	7.69	23.87
0.99999	10021	13	69.07	7.69	23.23
0.99999	16702	13	70.37	7.69	21.93
0.99999	3340	26	70.03	7.69	22.27
0.99999	10021	26	71.67	7.69	20.63
0.99999	16702	26	72.93	7.69	19.37
0.99999	3340	39	72.63	7.69	19.67
0.99999	10021	39	73.27	7.69	19.03
0.99999	16702	39	74.23	7.69	18.07

Table 8. Zika transmission risk estimates for maximum simulation time MGDrivE release scenarios. CC signifies maximum mosquito population estimate simulations with no simulated gene drive releases.

Rate of accurate HDR	Release ratio	Weekly releases	Zero transmission (%)	Low transmission (%)	Medium transmission (%)	High transmission (%)
CC	CC	CC	7.14	7.14	7.14	78.60
0.999	3340	13	12.40	11.96	17.00	58.67
0.999	10021	13	12.40	11.96	17.00	58.67
0.999	16702	13	12.40	11.96	17.00	58.67
0.999	3340	26	12.40	11.96	17.00	58.67
0.999	10021	26	12.40	11.96	17.00	58.67
0.999	16702	26	12.40	11.96	17.00	58.67
0.999	3340	39	12.40	11.96	17.00	58.67
0.999	10021	39	12.40	11.96	17.00	58.67
0.999	16702	39	12.40	11.96	17.00	58.67
0.9999	3340	13	12.40	11.96	17.00	58.67
0.9999	10021	13	13.57	11.90	16.84	57.70
0.9999	16702	13	17.17	11.76	16.34	54.73
0.9999	3340	26	17.17	11.76	16.34	54.73
0.9999	10021	26	30.30	11.30	14.47	43.93
0.9999	16702	26	24.33	11.53	15.30	48.83
0.9999	3340	39	15.97	11.83	16.50	55.70
0.9999	10021	39	30.30	11.30	14.47	43.93
0.9999	16702	39	42.27	10.90	12.77	34.10
0.99995	3340	13	19.53	11.70	15.97	52.77
0.99995	10021	13	27.93	11.40	14.80	45.90
0.99995	16702	13	36.30	11.10	13.60	39.03
0.99995	3340	26	30.30	11.30	14.47	43.93
0.99995	10021	26	42.27	10.90	12.77	34.10
0.99995	16702	26	36.30	11.10	13.60	39.03
0.99995	3340	39	27.93	11.40	14.80	45.90
0.99995	10021	39	38.63	11.03	13.27	37.03
0.99995	16702	39	36.30	11.10	13.60	39.03
0.99999	3340	13	47.03	10.73	12.07	30.17
0.99999	10021	13	49.43	10.63	11.74	28.20
0.99999	16702	13	54.20	10.46	11.07	24.23
0.99999	3340	26	53.00	10.50	11.24	25.27
0.99999	10021	26	59.00	10.30	10.40	20.33
0.99999	16702	26	63.73	10.13	9.70	16.37
0.99999	3340	39	62.57	10.16	9.87	17.40
0.99999	10021	39	64.97	10.10	9.54	15.40
0.99999	16702	39	68.53	9.96	9.03	12.47

weekly releases. All simulations with a 1:1 release ratio had similar probability distributions to the maximum mosquito population estimate levels.

RR1 showed no change in probability distribution for all scenarios at the maximum simulation time; probability of being below dengue transmission threshold was constant at 79.1%. RR2 varied from 58.1% at the lowest (0.001 resistance generation, 1:1 release ratio, 13 weekly releases) to 74.1% at its highest likelihood of being below transmission threshold (0.00001 resistance generation rate, 5:1 release ratio, 39 weekly releases). RR3 varied in probability of being below dengue transmission threshold from 40.0% (0.001 resistance generation, 1:1 release ratio, 13 weekly releases) to 69.5% (0.00001 resistance generation rate, 5:1 release ratio, 39 weekly releases).

The risk calculations for Zika transmission in each region for the maximum simulation time (Appendix 2) were constant for all simulations with release parameters of 0.001 resistance generation and for a resistance generation rate of 0.0001, release ratio of 1:1, and 13 weekly releases (High risk per region: RR1 = 39.6%, RR2 = 62.9%, RR3 = 73.5%). The remaining release parameters increased the likelihood of zero transmission as the resistance generation rate decreased or either the release ratio or duration of releases increased. The RR1 Zika transmission release ratio varied from 11.5% (0.001 resistance rate, 1:1 release ratio, & 13 weekly releases) to 71.2% (0.00001 resistance generation rate, 5:1 release ratio, & 39 weekly releases). The zero transmission Zika risk calculations for RR2 varied from 15.6% (0.001 resistance generation, 1:1 release ratio, 13 weekly releases) to 64.4% probability (0.00001 resistance generation rate, 5:1 release ratio, & 39 weekly releases). The likelihood of zero transmission in RR3 increased from 10.1% to 70.0% (0.00001 resistance generation rate, 5:1 release ratio, & 39 weekly releases) (Appendix 2).

4.3.2 Minimum simulation time

All of the simulations examined for the timepoint when mosquito populations were at their lowest gave constant dengue transmission risk distributions for each region (Table 9, Appendix 3). All risk estimates used the transmission rates based off of 0% herd immunity and 26°C (Focks et al. 2000). The risk estimate for RR1 being below the dengue transmission threshold was 81.9%. Risk in RR2 was 80.0%. The RR3 risk estimate for being below the dengue transmission threshold was 76.4%.

Analyzing risk of Zika transmission with the MGDrivE simulation output for the time period with the lowest average mosquito populations (Appendix 4) showed an increase in the probability of zero transmission as the resistance generation rate was decreased (Table 10). Resistance generation rates of 0.001, a release ratio of 1:1, and 13 weekly releases (the least invasive release scenario) resulted in zero Zika transmission risk of 25.6%, low risk of 20.6%, medium risk of 19.0%, and high risk of 34.8% in RR1. Zika transmission risk estimates for RR2 showed zero transmission probability of 23.4%, low risk of 16.8%, medium risk of 42.3%, and 17.5% probability for high risk. Estimates for Zika transmission in RR3 ranged from 18.7% probability of zero transmission, 28.8% low transmission, 33.3% medium transmission, and 19.3% probability of high Zika transmission. Increasing the release ratio to 3:1 and weekly releases to 26 increased the highest likelihood of zero transmission for this resistance generation level: zero Zika transmission probability of 59.8% for RR1 (low = 20.6%, med. = 8.7%, high = 10.8%), zero Zika transmission probability of 38.0% for RR2 (low = 33.8%, med. = 14.0%, high = 14.3%), and zero Zika transmission of 26.4% for RR3 (low = 44.7%, med. = 11.3%, high = 17.7%).

The resistance generation rate of 0.0001 resulted in risk estimates for zero Zika transmission of 38.7% for RR1 (low = 34.5%, med. = 10.9%, high = 15.8%), 25.7% for RR2 (low = 42.5%, med. = 17.5, high = 14.4%), and 20.8% for RR3 (low = 33.5%, med. = 28.0%, high = 17.7%) with the least invasive release scenario. Increasing the release ratio to 5:1 and weekly releases to 26

Table 9. Dengue transmission risk estimate for time point to minimum population size MGDrivE release scenarios for entire study area. CC signifies maximum mosquito population estimate simulations with no simulated gene drive releases.

Rate of accurate HDR	Release ratio	Weekly releases	Below transmission threshold (%)	At transmission threshold (%)	Above transmission threshold (%)
CC	CC	CC	60.63	7.69	31.70
0.999	3340	13	79.43	7.69	12.90
0.999	10021	13	79.50	7.69	12.80
0.999	16702	13	79.50	7.69	12.80
0.999	3340	26	79.50	7.69	12.80
0.999	10021	26	79.50	7.69	12.80
0.999	16702	26	79.50	7.69	12.80
0.999	3340	39	79.50	7.69	12.80
0.999	10021	39	79.50	7.69	12.80
0.999	16702	39	79.50	7.69	12.80
0.9999	3340	13	79.50	7.69	12.80
0.9999	10021	13	79.50	7.69	12.80
0.9999	16702	13	79.50	7.69	12.80
0.9999	3340	26	79.50	7.69	12.80
0.9999	10021	26	79.50	7.69	12.80
0.9999	16702	26	79.50	7.69	12.80
0.9999	3340	39	79.50	7.69	12.80
0.9999	10021	39	79.50	7.69	12.80
0.9999	16702	39	79.50	7.69	12.80
0.99995	3340	13	79.50	7.69	12.80
0.99995	10021	13	79.50	7.69	12.80
0.99995	16702	13	79.50	7.69	12.80
0.99995	3340	26	79.50	7.69	12.80
0.99995	10021	26	79.50	7.69	12.80
0.99995	16702	26	79.50	7.69	12.80
0.99995	3340	39	79.50	7.69	12.80
0.99995	10021	39	79.50	7.69	12.80
0.99995	16702	39	79.50	7.69	12.80
0.99999	3340	13	79.47	7.69	12.83
0.99999	10021	13	79.50	7.69	12.80
0.99999	16702	13	79.50	7.69	12.80
0.99999	3340	26	79.47	7.69	12.83
0.99999	10021	26	79.50	7.69	12.80
0.99999	16702	26	79.50	7.69	12.80
0.99999	3340	39	79.50	7.69	12.80
0.99999	10021	39	79.50	7.69	12.80
0.99999	16702	39	79.50	7.69	12.80

Table 10. Zika transmission risk estimates for time point to minimum population size MGDrivE release scenarios. CC signifies maximum mosquito population estimate simulations with no simulated gene drive releases. High transmission for release scenarios was stable at 14.27%, signifying that this was at the lowest possible estimation for all scenarios modeled.

Rate of accurate HDR	Release ratio	Weekly releases	Zero transmission (%)	Low transmission (%)	Medium transmission (%)	High transmission (%)
CC	CC	CC	7.14	7.14	7.14	78.60
0.999	3340	13	22.57	22.07	31.53	23.87
0.999	10021	13	27.17	40.07	18.29	14.53
0.999	16702	13	35.50	36.67	13.39	14.50
0.999	3340	26	26.33	38.70	19.53	15.50
0.999	10021	26	41.40	33.03	11.34	14.27
0.999	16702	26	42.97	31.87	10.88	14.27
0.999	3340	39	33.47	38.73	13.51	14.27
0.999	10021	39	34.43	38.53	12.74	14.27
0.999	16702	39	28.97	39.40	16.90	14.70
0.9999	3340	13	28.40	36.83	18.80	15.97
0.9999	10021	13	38.60	33.70	13.22	14.47
0.9999	16702	13	40.80	32.23	12.04	14.27
0.9999	3340	26	37.20	35.43	12.86	14.50
0.9999	10021	26	46.07	28.73	10.94	14.27
0.9999	16702	26	48.70	26.17	10.91	14.27
0.9999	3340	39	41.93	32.73	11.04	14.27
0.9999	10021	39	43.33	31.50	10.91	14.27
0.9999	16702	39	42.10	32.30	11.34	14.27
0.99995	3340	13	32.03	34.73	17.77	15.53
0.99995	10021	13	40.87	32.30	12.36	14.50
0.99995	16702	13	46.73	28.07	10.98	14.27
0.99995	3340	26	40.10	32.47	12.77	14.67
0.99995	10021	26	46.70	28.00	11.04	14.27
0.99995	16702	26	50.07	24.80	10.88	14.27
0.99995	3340	39	43.30	31.40	11.04	14.27
0.99995	10021	39	43.90	30.93	10.88	14.27
0.99995	16702	39	42.40	32.33	10.98	14.27
0.99999	3340	13	40.80	30.27	13.59	15.37
0.99999	10021	13	44.27	29.00	12.26	14.50
0.99999	16702	13	48.47	25.60	11.64	14.27
0.99999	3340	26	43.90	30.13	11.68	14.27
0.99999	10021	26	53.23	21.47	11.08	14.27
0.99999	16702	26	56.07	18.77	10.91	14.27
0.99999	3340	39	55.03	19.67	11.08	14.27
0.99999	10021	39	58.57	16.33	10.88	14.27
0.99999	16702	39	58.93	15.93	10.88	14.27

gave the highest likelihood of zero transmission of 65.4% for RR1 (low = 15.1%, med. = 8.7%, high = 10.8%), 42.5% for RR2 (low = 29.5%, med. = 13.7%, high = 14.3%), and 38.2% for RR3 (low = 33.9%, med. = 10.3%, high = 17.7%) (Appendix 4).

Zero Zika transmission risk estimates varied from 45.1% for RR1 (low = 30.6%, med. = 10.2%, high = 14.2%), 28.9% for RR2 (low = 37.5%, med. = 18.9%, high = 14.7%), and 22.1% for RR3 (low = 36.1%, med. = 24.2%, high = 17.7%) for the simulations modeling the resistance generation rate of 0.00005, 1:1 release ratio, and 13 weekly releases. The highest probabilities of zero transmission for this resistance rate were increased to 65.0% for RR1 (low = 15.5%, med. = 8.7%, high = 10.8%), 45.1% for RR2 (low = 26.9%, med. = 13.6%, high = 14.3%), and 40.1% for RR3 (low = 32.0%, med. = 10.3%, high = 17.7%) with a release ratio of 5:1 and 26 weekly releases (Appendix 4).

The final resistance generation rate of 0.00001 gave risk estimates for zero Zika transmission of 58.2% for RR1 (low = 19.0%, med. = 9.7%, high = 13.1%), 36.7% for RR2 (low = 32.88%, med. = 16.0%, high = 14.5%), and 27.5% for RR3 (low = 39.0%, med. = 15.1%, high = 18.5%) when modeled with a 1:1 release ratio and 13 weekly releases. These risk estimates were increased to 64.6% for RR1 (low = 15.9%, med. = 8.7%, high = 10.8%), 54.9% for RR2 (low = 17.2%, med. = 13.6%, high = 14.3%), and 57.3% for RR3 (low = 14.7%, med. = 10.3%, high = 17.7%) when the release ratio was increased to 5:1 with 39 weekly releases (Appendix 4).

4.4 Uncertainties and sensitivity analysis

Model sensitivity analysis shows the amount of influence that changes in one node can assert on the other nodes within the Bayesian network (Pollino et al. 2007; Marcot et al. 2012; Norsys 2014). Model sensitivity is calculated as reduction in entropy or mutual information. Entropy means the amount of information that is held within a variable. Shared entropy between to variables is the amount of mutual information that can be gathered by observing only one of the

variables, showing their mutual dependence on one another. Sensitivity analysis determines to what degree each input influences the endpoint. Larger values of mutual information show a greater influence on an endpoint.

Results of the sensitivity analysis in Netica™ focusing on entropy reduction showed that for the scenarios examined for the maximum simulation time the rate of accurate homology-directed repair (HDR) and conversely the rate of resistance generation was the most influential for Zika transmission, proportion of surviving mosquitoes that are resistant to the drive, and resulting population after the release of the gene drive. Sensitivity analysis showed that temperature was most influential on dengue transmission, followed by accurate HDR, release ratio, and duration of releases. For each node analyzed the overall pattern of importance was the same. Frequency of releases followed rate of resistance generation. The release ratio was the least important of the input nodes analyzed for each of the maximum simulation endpoints.

Sensitivity analysis for the maximum simulation time Zika transmission node showed that the rate of HDR had an entropy reduction of 5.77%. The frequency of releases node had an entropy reduction of 0.295%, and the release ratio node had an entropy reduction of 0.217%. The dengue transmission node showed entropy reductions of 1.39% for temperature, 0.657% for HDR, 0.033% for release frequency, and 0.0243% for release ratio. The node representing the proportion of surviving mosquitoes that were resistant showed that HDR had a 4.84% entropy reduction, the frequency of releases 0.803%, and release ratio 0.376% reduction. When examining the intermediate node of resulting population after gene drive release I saw an entropy reduction of 14.7% for HDR, 0.0109% for release frequency, and 0.008% for release ratio (Figure 9).

When running the sensitivity analysis on the Bayesian network that showed the results at the time point associated with the lowest possible population reduction, there was no consistent pattern of which input variables provided the most entropy reduction. The time to minimum

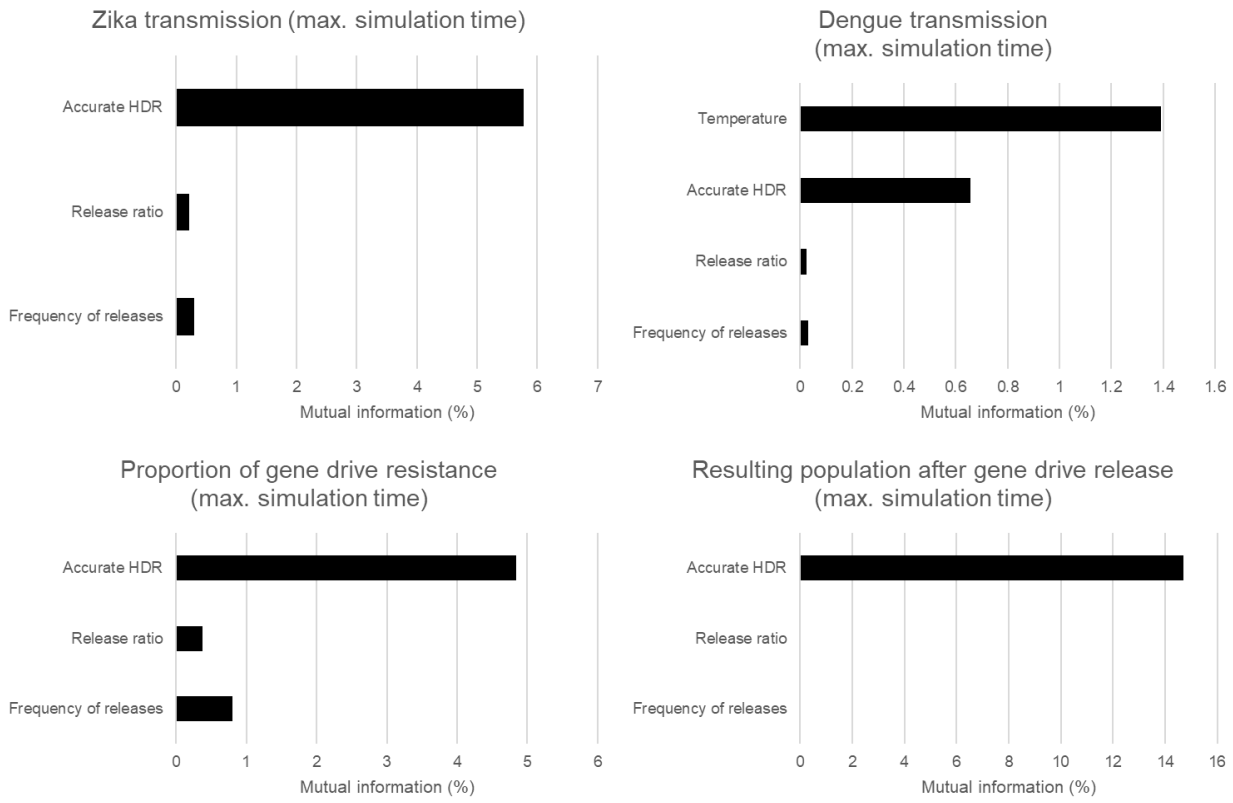

Figure 9. Results of entropy reduction analysis for maximum simulation time Bayesian network.

population Zika transmission node showed an entropy reduction of 0.282% for rate of accurate HDR, 0.0744% for release ratio, and 0.0675% for duration of releases. Sensitivity analysis of the dengue transmission node showed an entropy reduction of 5.35% for temperature, 0.000707% for accurate HDR, 0.000173% for release ratio, and 0.000151% for duration of releases.

The node representing the rate of resistance within the surviving mosquito populations showed an entropy reduction of 16.9% for HDR, 1.52% for release frequency, and 0.012% for release ratio. The intermediate node showing the distribution of the resulting population number after the release of the gene drive showed entropy reductions of 10.4% for HDR, 2.09% for release ratio, and 2.04% for release frequency (Figure 10). Rates of accurate HDR was the most influential input node in relation to the endpoints and intermediate node examined with

sensitivity analysis, except for dengue transmission, where temperature was most influential. Release ratio was more influential that the frequency of releases for dengue and Zika transmission, as well as the resulting population size after the gene drive release. Frequency of releases was more influential that release ratio when examining the mutual information between input nodes and the node estimating the probability of the proportion of the surviving mosquito population developing resistance to the drive.

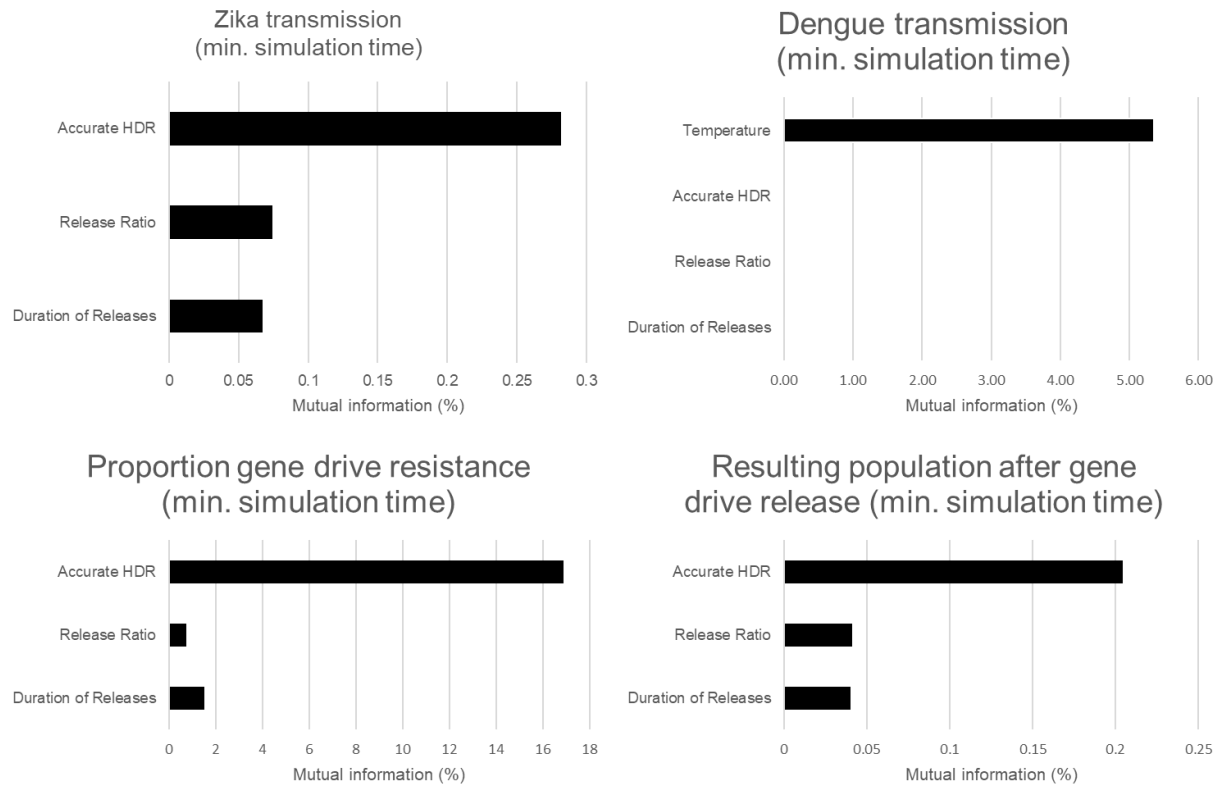

Figure 10. Results of entropy reduction analysis for the Bayesian network showing time point associated with the greatest average population reduction.

5. Discussion

5.1 The risk assessment approach meets NASEM 2016 recommendations

The National Academies of Sciences, Engineering, and Medicine recommended the use of the BN-RRM for the risk assessment of gene drives within chapter six of their report, *Gene Drives*

on the Horizon (NASEM 2016). Several criteria were listed that future risk assessments must include to effectively model the risk associated with the development and deployment of gene drive technologies. These are: 1) acknowledging the fact that the field of synthetic biology is still in its infancy and uncertainties that lie behind every advancement need to be identified, 2) quantitatively and probabilistically describing the direct and indirect effects of the development and deployment of gene drive technologies, 3) discussing alternative management approaches in the risk assessment and quantifying the trade-offs between management scenarios, and 4) the model requires the input of all relevant stakeholders as well as management in order to effectively describe the cultural importance of endpoints within the specific study area.

Criterion 1. Within the context of gene drive research, the amount of data regarding the propagation of gene drives within a target population is lacking. Lab releases have been few in number and relegated to a small subset of species. Research conducted to model the use of population suppression gene drives within my study are based off of work done on the malaria vector *Anopheles gambiae* (Hammond et al. 2016). There are most likely differences in how gene drives will propagate within different species of arbovirus vectors, as well as how fast the rates of resistance to the gene drive will occur. The Bayesian networks developed in this study incorporate uncertainties within the probability distributions of each node. Nodes that have lacking or insufficient data have been parameterized by even distributions to signify that there is insufficient data to determine calculations of risk. These even distributions signal that there are uncertainties that exist. If the findings of sensitivity analyses show that these areas are contributing to uncertainty these should be areas where future research is directed. Because the rate of accurate HDR and subsequently, the rates of resistance allele generation, proved to be very influential in the model results, extremely accurate estimations of potential drive systems would need to be calculated to inform future gene drive risk assessments. In terms of decision making, this means that management efforts need to prioritize research so that a complete

understanding of gene drive technologies can be developed. As suitable data are obtained, nodes can be further parameterized to better capture the full distributions of likely outcomes and other pathways can be included so that specific management goals can be addressed by the model. Any potential management actions taken before this is accomplished would be taken without regard for a true risk assessment, and any subsequent management actions would need to establish the fact that the knowledge available is likely to be incomplete.

Criterion 2. The Bayesian networks I developed for the risk assessment of gene drive technologies highlight important research needs and areas of uncertainties that need to be addressed. To further inform the probability distributions of specific outcomes for each proposed release of engineered vector populations additional lab conducted releases of specific transgenic gene drive mosquitoes will be required. Field releases will greatly inform these uncertainties but should only be done once the impacts of field releases can be fully understood and their trade-offs quantified. Within the Bayesian networks I developed, additional pathways can be easily integrated that describe other endpoints and effects that stakeholders may deem necessary.

These additional pathways may include the impacts to agricultural and ecological endpoints, as well as the likelihood of horizontal gene transfer of the drive or parts of the engineered genetic element into non-target species. Specific data that would inform these additional pathways describing unintended consequences depends on the potential endpoints being considered. Agricultural endpoints could be associated with the ecosystem services that the *Ae. aegypti* mosquito performs within Ponce. Studies examining their role as pollinators should extensively describe what specific functions the mosquito fulfils and in what circumstances. Ecological endpoints that could be impacted include food-web interactions. All species that eat the mosquito could be impacted by their eradication. The specific percentage of predator diet that the mosquito makes up should be defined. Any rates of recolonization by other mosquito

species would inform the impact on the food web. Although this was not accomplished in this study, my research has made it clear that a strong understanding of the role of the target species in the environment must be understood to grasp the full range of unintended outcomes.

In the scenarios I modeled, incomplete eradication will most likely result in the replacement of the wild type population with one that has developed resistance to the gene drive. The existing management plan would then have to contend with a vector population that has already developed widespread resistance to contemporary pyrethroid pesticides (Agramonte et al. 2017), as well as resistance to the particular gene drive deployed.

Criterion 3. Discussing alternative management strategies and quantifying the 'what-if' scenarios (Pearl & Mackenzie 2018) make this approach ideal for use within an adaptive management framework (Landis et al. 2017). Any causal pathways that can be quantitatively described can be included into the Bayesian network. The Bayesian networks developed here can be further developed to include alternative management strategies by incorporating additional nodes that describe additional relevant causal pathways. Several pyrethroid insecticides were included for study within my Bayesian networks. The synergistic or antagonistic effects of other pesticides can be included given that there is sufficient toxicity data to develop dose-response curves to inform the conditional probability tables. Other pest management strategies can also be included, given that there are enough data about the management options, or through expert elicitation (Marcot et al. 2006). For example, an extensive effort to reduce standing water used as mosquito breeding habitat could be included within the nodes that model the maximum mosquito population estimate of each patch. If a dengue or Zika vaccine were to be developed and used, the infection rate within the human population could be modeled given the new probability of transmission rates.

Criterion 4. The BN-RRM approach includes stakeholder input in an iterative manner that makes sure that one group does not dominate the input into the risk assessment. In this study direct

input from people within Puerto Rico and other stakeholder groups was not included. Attempts were made to reach out to the Puerto Rico Vector Control Unit and other arbovirus researchers in Puerto Rico. However, Puerto Rico is dealing with economic strife, disaster relief, and political uprisings (Bonilla 2020) that may have led to some having less time to respond to inquiries. The public and other stakeholders must be included when conducting a site-specific risk assessment for a tentative project looking at an actual release of gene drive engineered mosquitoes. By expanding the sphere of influence into the risk assessment process, those that will be directly impacted have access to determining the cultural values that will be included to determine the endpoints of the Bayesian network. This is especially helpful within the context of ecological risk assessment when defining what characteristics of the endpoints need to be measured and what terms like 'recovered' actually mean.

5.2 Application to the field of synthetic biology

The field of synthetic biology research is expanding and moving closer to field releases of transgenic organisms. To date there have been no quantitative risk assessments conducted for their use. By developing the Bayesian networks within this work, I have shown that the approach as recommended by the National Academies (NASEM 2016) can be fully integrated with highly site-specific spatial components to directly estimate the impacts of gene drive releases aimed at curtailing disease transmission. The BN-RRM can most likely be used to model synthetic biology release risk for any other purpose in any location, given sufficient data. The degree of specificity of future risk assessments will depend on the degree that data are available for the study area, the characteristics of the target species, and rates of propagation of the gene drive throughout the target population. However, even if the data sets are lacking, the Bayesian network approach highlights areas that require more information to be gathered and other areas that contribute to the uncertainties of gene drive deployment. In this sense, the Bayesian network approach can be a highly effective tool for management to help direct future

research into areas that will inform the overall risk assessment of the use of gene drive technologies.

For any future risk assessment of gene drive and synthetic biology, specific information will be required to make any judgements about the intended and unintended outcomes due to their use. In the system that I modeled, incorporating population suppression gene drives as a management tool to reduce disease in Puerto Rico, the greatest data needs are related to the target species. Extensive trapping data would greatly inform the modeling by giving much more accurate estimates of existing *Ae. aegypti* populations. More information gained from lab releases of the drive in subsets of the vector population would greatly improve the knowledge of how the drive spreads through the target population, as well as estimates of the rates of resistance that will most likely develop as a result. The amount of time that the drive would remain effective could be established for gene drives that would result in a high likelihood of non-eradication. To even begin to model the unintended consequences of the release in Ponce, the life history parameters and rates of recolonization of closely related vector and non-vector species would need to be understood. Rates of hybridization between the *Ae. aegypti* and any other potential mates would have to be established to determine the probability of gene transfer between target and non-target species. A substantial survey of the role that the *Ae. aegypti* servers within the ecosystem and any ecosystem services that the mosquito supplies could influence any estimates regarding the risk associated with the complete eradication of the species and the impacts that would have on local food-webs.

5.3 Gene drive risk assessment within adaptive management

Adaptive management is an iterative process where resource managers update their strategies as concurrent strategies provide additional insights to their efficacy (Landis et al. 2017). Adaptive management uses knowledge gained through current monitoring programs to constantly update and improve available options for resource managers. The adaptive

management process can be thought of as a loop. The process is driven by public engagement and governance. Existing problems are defined that reflect the culturally significant endpoints of the study area, goals are explicitly stated, management actions are explored, and data that are relevant to determining risk to the endpoints through the monitoring programs are examined. At this point the impacts of current management practices are re-evaluated with updated information and the process starts over with the goal of constantly improving the outcomes (Landis et al. 2017). Adaptive management will be a substantial regulatory tool to constantly evaluate the interactions of a gene drive with the environment. Gene drives have not been leveraged as a management option before and any information that is gathered that quantitatively describes the outcomes can be used to make better judgements about subsequent management actions.

The BN-RRM has been shown to appropriately integrate within the adaptive management framework (Landis et al. 2017). Any management option that includes the deployment of gene drives will require the input of stakeholders and local governance. Existing and alternative management strategies must be examined to determine if the use of gene drives is warranted, given the vast uncertainties involved with their use, including rates of horizontal gene transfer and the lack of a reversibility once the gene drive has been deployed.

Gene drives and synthetic biology present a true ‘wicked problem’ (Rittel and Webber 1973). Almost all of the criteria listed by Rittel and Webber (1973) can be associated with use of a gene drive. In almost any sense, the dilemmas posed by the use of a gene drive are planning problems; a gene drive would only be used to solve some sort of greater issue, in my case study the issue would be the reduction in debilitating mosquito borne disease. Because of the uncertainties involved with gene drive use, the full breadth of consequential outcomes is almost inconceivable to define (Rittel and Webber 1973). The unknown consequences of the deployment of a gene drive are so great because the research is still in its infancy and because

the world in which a gene drive would be introduced is inherently an open system with no way to completely model every interaction that could possibly take place. The best we can do is model potential outcomes, but this is in itself another wicked problem. The range of stakeholders that are involved with potential uses of gene drives also lends itself to the issue. Each distinct group of people that can be affected by a gene drive will have differing socio-economic backgrounds with differing attitudes about the appropriateness of a solution that gene drives could potentially solve (Rittel and Webber 1973). Furthermore, the impacts of a gene drive may prove to be completely irreversible, and any potential solutions to reverse a drive will present yet another wicked problem. The starkest characteristic of gene drives as a wicked problem is that each and every potential use will present its own set of unique circumstances with no clear resolutions. The characteristics of a target species will differ between location, the aspects of a certain gene drive architecture will be different depending on the genetic makeup of the target species, and the social values of the stakeholders won't be consistent even within one proposed use and will vary over time. Because every proposed usage of a gene drive will be unique, the lessons from one circumstance may very well lend themselves to another proposed use but no two will be identical and commonalities should not be necessarily assumed (Rittel and Webber 1973).

The synthetic biology risk assessment framework recommended by the National Academies (NASEM 2016) and explored within this study shows that risk can be calculated for various management options related to the release of gene drive engineered *Ae. aegypti* as vectors to control dengue and Zika virus. This framework can easily be adapted to alternate locations and the inclusion of alternative management strategies can be incorporated given that the data sets are sufficient. Although no two gene drive proposals will be identical, certain similarities between them can be characterized so that the potential outcomes can be estimated. The BN-RRM can be adapted to specific landscapes if the geospatial data are available at that site. The advancements and characteristics of gene drive architectures can be incorporated into the

modeling process so that the ability of the drive to spread within the target population can be reasonably understood and parameterized.

In the scenario I explored, the BN-RRM can provide disease prevention specialists with a powerful tool to evaluate proposed and current management practices with regard to the potential release of gene drives. If an adaptive management process were undertaken, the BN-RRM can be easily updated with any additional information as it becomes available. The risk estimates would reflect the most current information available and the trade-offs incurred by their use, or other management strategies, can be easily visualized.

Within this study the integration of alternative management strategies was shown through the inclusion of a pesticide application pathway. This causal pathway illustrates how current management approaches can be used together with gene drive technologies to drive mosquito populations to levels below disease transmission thresholds. Additionally, this shows how incomplete eradication of the target species can still be a useful outcome to managers that have alternative and additional strategies in place to control the remaining mosquito populations. If avoiding resistance to the gene drive is an ideal endpoint, the Bayesian network can calculate what release scenarios could be used to avoid the likely development of resistance. If these scenarios resulted in non-eradication events, they could still be a beneficial outcome within an overarching adaptive management scheme that includes current mosquito control measures. Pesticide application schemes can then be modeled using site-specific information to determine the likelihood of further reducing populations based on dose-response modeling parameterized into the pesticide application pathway of the Bayesian network.

6. Conclusions and Next Steps

6.1 Resistance generation

Lab tested gene drives, such as those demonstrated by Hammond et al. (2016), lack a sufficiently low resistance generation rate for them to be effective as a stand-alone tool for complete eradication of mosquito species when modeled over large spatial scales. Landscape scale releases, as demonstrated by the results of this study, will only result in rebounds of resistant or wild-type mosquitoes. For complete eradication as a result of the gene drive, the resistance generation rate of the target population must be below some critical level as determined by the size of the starting population of mosquitoes (Champer et al. 2017; Marshall et al. 2017; Noble et al. 2017; Unckless et al. 2017). If all habitat patches within the release area are connected, there will be a high likelihood that resistance developing in one patch, due to the high selective pressure against the fitness reduction of the drive, will spread to the remaining patches. As demonstrated by this study, successful population suppression will be dependent upon the number of gene drive engineered mosquitoes released, the duration of releases, and the characteristics of the landscape, the population dynamics of the target species, and more importantly, the resistance allele generation rate.

Resistance generation due to non-homologous end-joining must be circumvented for gene drive technologies to be viable for population suppression. As gene drive and synthetic biology technology progresses, techniques such as multiplexing (Marshall et al. 2017) could sufficiently reduce the rate of resistance generation to the gene drive within the target species to rates that could lead to complete population suppression. Currently, the rates of resistance generation to demonstrated gene drives are insufficient. What my modeling has shown is that the gene drive releases of any scenario do reduce the target population up to a point. Depending on the parameters surrounding the releases of the gene drive engineered organisms there are varying degrees of success in terms of reducing overall population numbers.

Currently available gene drives may not be suitable for complete eradication of certain target populations of mosquitoes. However, they can be seen as a management option for effectively reducing populations to levels below thresholds required for disease transmission. Results of my modeling show that gene drives may be suitable as one tool within an adaptive management framework that can leverage other existing management options to potentially eradicate a population of *Ae. aegypti*. However, the implications of complete eradication of the Ae. aegypti from one area are not well understood. Current research into the ecosystem services provided by the Ae. aegypti are not well documented. The complete impacts of elimination from a specific community can be described with direct monitoring after a release of the engineered mosquitoes. By incorporating alternative management strategies within the BN-RRM, the trade-offs of various management strategies can be easily quantified. If the goal is to effectively reduce vector populations to some predetermined level and then aggressively leverage contemporary mosquito management strategies to further reduce the populations, the desired outcome may be achievable. The estimation of risk for the direct and indirect effects can be obtainable through the risk assessment framework recommended by the National Academies of Sciences, Engineering, and Medicine (NASEM 2016). To incorporate the pathways to describe the indirect effects associated with gene drive engineered *Ae. aegypti* and a method of disease control, much more data regarding the ecosystem services that the mosquito provides in Ponce needs to be gathered. To inform food-web interactions that could be impacted, the role of the mosquito in the diet of local predators needs to be quantified. Rates of hybridization with non-target species and gene transfer are unknown for the non-vector mosquito species and would have to be determined to inform the risk of horizontal gene transfer. The agricultural impacts resulting from eradication of the mosquito in Ponce are currently unquantified and would need to be researched in order to develop risk estimates for this endpoint. Undoubtedly there will be unintended consequences and future risk assessments can incorporate these pathways into the modeling framework to highlight research needs that would inform management decisions.

The BN-RRM has shown its utility as part of an epidemiological risk assessment framework. Causal disease pathways, where a mechanistic understanding of how disease transmission occurs (Chen and Pollino 2007; Marcot 2017; Landis et al. 2020), were included to quantitatively describe the risk of transmission rates for diseases vectored by the *Ae. aegypti* mosquito within Ponce, Puerto Rico. Although the gene drive modeled after the construct created by Hammond et al. (2016) was unsuccessful in causing eradication in simulations, it was successful in causing drastic declines in simulated mosquito populations. Resource managers could use the BN-RRM to explore decisions based on the trade-offs between overall population reduction and the likelihood of resistance to the gene drive developing in the remaining population. Multiple synergistic management options can also be included to determine what the best possible outcome or set of outcomes would be depending on the scenarios depicted.

This work builds off of the approach recommended by the National Academies of Science, Engineering, and Medicine (NASEM 2016), and goes a step further by incorporating site specific details into a case-study related to the use of gene drives as vectors to control disease. The Bayesian networks provide estimations of risk for the direct and indirect effects of gene drive deployment in a way that can be highly useful to the decision-making process. The work that I have done highlights some of the data requirements for future quantitative risk assessments of gene drive technologies. These data requirements inform the site-specific release scenarios that are able to be integrated into the BN-RRM approach. Furthermore, this approach illustrates the flexibility of the BN-RRM to any potential future release of gene drives by being able to incorporate site-specific and organismal specific characteristics into the model that can easily be adapted for different research programs.

The trade-offs of alternative management actions are represented in a meaningful way that fits well within an adaptive management decision making framework (Landis et al. 2017). The BN-RRM shows how specific decisions can be evaluated regarding gene drive deployment. The

framework also shows how information relating to gene drive technology can be assembled in a way that directly supports the decision-making process. The BN-RRM lets researchers and decision makers understand the trade-offs associated with management actions and improves the overall effectiveness and utility of the novel technology posed by gene drive and synthetic biology products.

6.2 Next steps

1) In order to further develop the estimation of risk associated with the use of gene drives, the full range of possible outcomes needs to be examined. In my study ecological and agricultural pathways have not been included due to lack of research into those areas. These could be included and left unparameterized with even distributions to signify knowledge gaps. The complete range of risk needs to be known to fully comprehend what the trade-offs associated with the use of gene drives are. The BN-RRM can be used as the risk assessment tool for synthetic biology moving forward that will clearly highlight areas of uncertainty. This study shows that the aspects of the spread of the gene drive through a population can be absorbed into the BN-RRM framework with additional stressors (pesticides) modeled. The landscape characteristics that will affect the spread of the gene drive will differ from location to location and as the system that is being modeled becomes larger. If geospatial data exists for locations of gene drive use, then these landscape conditions can be interpreted as inputs into the BN-RRM. The role of temperature on the vector density of *Ae. aegypti* and rates of Zika transmission is an area that needs to be addressed. Although Zika transmission thresholds were developed by Barrera et al. (2017), these were in relation to trapping data and rates of seroprevalence of disease and the influence of temperature was not accounted for. Sensitivity analysis showed that temperature is highly influential (Figures 9 and 10) in the dengue transmission pathway (Focks et al. 2000) and most likely has a similar role in Zika transmission (Barrera et al. 2017). Areas of uncertainty represent research gaps that are required to comprehend the full range of

risk. It is suggested that gene drive research be monitored using an adaptive management approach (NASEM 2016). The BN-RRM can be used to constantly update risk calculations based on any new data that becomes available.

2) The role of *Ae. aegypti* within Puerto Rico would need to be better understood to fully understand the ecological impacts of their eradication. There are 36 known species of mosquito that inhabit parts of Puerto Rico (Fox 1953). The principal vector for dengue and Zika virus is the *Ae. aegypti* mosquito. The *Aedes* spp. do act as pollinators for some orchid species (Thein 1969; Gorham 1976; Lahondère et al. 2020), but these plant species do not occur within Puerto Rico and are more commonly found in Northern latitudes. Eradication of the *Ae. aegypti* may pose little threat to pollination success of plant species within Puerto Rico, though there is a lack of research into this area, which must be addressed if field deployments of gene drive engineered mosquitoes were to be proposed as a management strategy. A full understanding of the ecosystem services that the *Ae. aegypti* performs is a necessary requirement for determining risk of gene drive use and the unintended consequences as they relate to the ecological interactions in the study area. Furthermore, within this study Ponce was treated as a closed system and emigration and dispersal of the gene drive was not considered to other bordering municipalities. To have a complete range of risk calculation the likelihood of the gene drive spreading beyond political borders within well connected habitat needs to be addressed. Dispersal beyond legal boundaries can be incorporated by expanding the spatial scope to the rest of Puerto Rico, given that sufficient spatial data exists to do that, if not than this highlights yet another area to be researched in terms of having a complete understanding of risk.

3) Predation upon *Ae. aegypti* within Puerto Rico is not a significant controlling factor on regulating their populations (Barrera et al. 2006). Although there are several species that have been documented as taking *Ae. aegypti* as food items, none of these predators specialize on eating the *Ae. aegypti* mosquito solely (Arrivillaga & Barrera 2003). Introduction of biocontrol

agents into specific *Ae. aegypti* habitats can reduce populations, but because the mosquito can use highly cryptic containers as habitats, the predator species are not likely to be present in many areas (Krol et al. 2019). Understanding the role of *Ae. aegypti* as a food source would be required to parameterize nodes and pathways that describe ecological impacts of their eradication. Other mosquito species could expand their range and fill the role of food items for predators, but this another area where research must be conducted if field deployments were to be proposed. If the *Ae. aegypti* were to be fully eradicated from a location, the rates of recolonization of other mosquito species in the area would need to be quantified in order to better describe the agricultural and ecological impacts resulting from their extirpation, especially if a closely related species is also a disease vector.

4) The BN-RRM can be adapted to include both ecological and human health endpoints. The National Academies of Sciences, Engineering, and Medicine's recommended approach to the regulation and risk assessment of gene drives (NASEM 2016) is achievable using realistic landscape characteristics and parameters that describe how the gene drive spreads throughout and reduces the target population, impacting human health endpoints. As the research into gene drives progresses, new data can be incorporated into the risk assessment process in an iterative fashion that can be constantly updated to suit manager's needs. Through this recommended approach (NASEM 2016), the effects resulting from the use of gene drives as methods to control disease can be quantified in a probabilistic manner. However, there are large data gaps regarding the role that the *Ae. aegypti* mosquito servers within the ecosystem and these areas need to be addressed in a probabilistic way before a complete risk assessment of gene drive engineered mosquitoes can be fully completed.

6. REFERENCES

Adelman ZN, & Tu Z. 2016. Control of mosquito-borne infectious diseases: sex and gene drive. Trends in parasitology. 32:219-229.

Altrock PM, Traulsen A, & Reed FA. 2011. Stability properties of underdominance in finite subdivided populations. PLoS Comput Biol, 7:e1002260 [accessed 7 June 2020].

Anderson SM & Landis WG. 2012. A pilot application of regional scale risk assessment to the forestry management of the Upper Grande Ronde Watershed, Oregon. Human and Ecological Risk Assessment 18:705-732.

ArboNET. Centers for Disease Control and Prevention. 2019. ArboNET disease maps. Available online at https://wwwn.cdc.gov/arbonet/maps/ADB_Diseases_Map/index.html (accessed 24 February 2019).

Ardlie KG. 1998. Putting the brake on drive: Meiotic drive of t haplotypes in natural populations of mice. Trends Genet. 14:189-193.

Arrivillaga J, & Barrera R. 2004. Food as a limiting factor for Aedes aegypti in water-storage containers. Journal of vector Ecology, 29:11-20.

Ayre KK, & Landis WG. 2012. A Bayesian approach to landscape ecological risk assessment applied to the Upper Grande Ronde Watershed, Oregon. Human and Ecological Risk Assessment 18:946-970.

Ayre KK, Caldwell CA, Stinson J, & Landis WG. 2014. Analysis of regional scale risk of whirling disease in populations of Colorado and Rio Grande cutthroat trout using a Bayesian belief network model. Society for Risk Analysis. 34:1589-1605.

Barrera R, Acevedo V, Felix GE, Hemme RR, Vazquez J, Munoz JL, & Amador M. 2017. Impact of autocidal gravid ovitraps on chikungunya virus incidence in Aedes aegypti (Diptera: *Culicidae*) in areas with and without traps. Journal of medical entomology, 54:387-395.

Barrera R, Amador M, & Clark GG. 2006. Ecological factors influencing Aedes aegypti (Diptera: Culicidae) productivity in artificial containers in Salinas, Puerto Rico. Journal of medical entomology, 43:484-492.

Barrera R, Felix G, Acevedo V, Amador M, Rodriguez D, Rivera L, Gonzalez O, Nazario N, Ortiz M, Muñoz-Jordan, JL, & Waterman SH. 2019. Impacts of Hurricanes Irma and Maria on Aedes aegypti Populations, Aquatic Habitats, and Mosquito Infections with Dengue, Chikungunya, and Zika Viruses in Puerto Rico. The American journal of tropical medicine and hygiene, 100:1413-1420.

Barrett AD. 2018. Current status of Zika vaccine development: Zika vaccines advance into clinical evaluation. npj Vaccines, 3:1-4.

Beck HE, Zimmermann NE, McVicar TR, Vergopolan N, Berg A, & Wood EF. 2018. Data descriptor: present and future Köppen-Geiger climate classification maps at 1-km resolution. Scientific data. Available online at https://www.nature.com/articles/sdata2018214 (accessed 24 February 2019).

Beeman RW, Friesen KS, & Denell RE. 1992. Maternal-effect selfish genes in flour beetles. Science, 256:89-92.

Bernoulli D. 1760. Essai d'une nouvelle analyse de la mortalité causée par la petite vérole, et des avantages de l'inoculation pour la prévenir. Histoire de l'Acad., Roy. Sci.(Paris) avec Mem, 1-45.

Bonilla, Y. (2020). The swarm of disaster. Political Geography, 102182. (accessed 16 July 2020).

Buchman A, Gamez S, Li M, Antoshechkin I, Li HH, Wang HW, Chen CH, Klein MJ, Duchemin JB, Paradkar PN & Akbari OS. 2019. Engineered resistance to Zika virus in transgenic

Aedes aegypti expressing a polycistronic cluster of synthetic small RNAs. Proceedings of the National Academy of Sciences, 116:3656-3661.

Buchman A, Marshall JM, Ostrovski D, Yang T, & Akbari OS. 2018. Synthetically engineered Medea gene drive system in the worldwide crop pest *Drosophila suzukii*. Proceedings of the National Academy of Sciences, 115:4725-4730.

Cain J, Batchelor C & Waughray D. 1999. Belief networks: a framework for the participatory development of natural resource management strategies. Environment, Development and Sustainability 1:123-133.

Carvalho DO, McKemey AR, Garziera L, Lacroix R, Donnelly CA, Alphey L, Malavasi A, & Capurro ML. 2015. Suppression of a field population of Aedes aegypti in Brazil by sustained release of transgenic male mosquitoes. PLoS neglected tropical diseases, 9(7), e0003864.

CDC. Center for Disease Control and Prevention. 2016a. Success in mosquito control: an integrated approach. Available from https://www.epa.gov/mosquitocontrol/success-mosquito-control-integrated-approach (accessed 13 March 2019).

CDC. Center for Disease Control and Prevention. 2017. Centers for Disease Control and Prevention. Zika virus. Insecticide resistance testing in Puerto Rico. Retrieved: February, 24, 2019, from https://www.cdc.gov/zika/vector/testing-puertorico.html.

CDC. Centers for Disease Control and Prevention. 2016b. CDC awards $13 million to establish the first Vector Control unit in Puerto Rico. Retrieved: February 18, 2020, from https://www.cdc.gov/media/releases/2016/s0927-vector-control-puerto-rico.html.

Champer J, Reeves R, Oh SY, Liu C, Liu J, Clark AG, & Messer PW. 2017. Novel CRISPR/Cas9 gene drive constructs reveal insights into mechanisms of resistance allele

formation and drive efficiency in genetically diverse populations. PLoS genetics, 13:e1006796.

Chen SH, & Pollino CA. 2012. Good practice in Bayesian network modelling. Environmental Modelling & Software. 37:134-145.

Colnar AM & Landis WG. 2007. Conceptual model development for invasive species and a regional risk assessment case study: the European green crab, *Carcinus maenas*, at Cherry Point, Washington, USA. Human and Ecological Risk Assessment 13:120-155.

Cox J, Grillet ME, Ramos OM, Amador M, Barrera R. 2007. Habitat segregation of dengue vectors along an urban environmental gradient. Am J Trop Med Hyg. 76:820–826

Curtis CF. 1968. Possible use of translocations to fix desirable genes in insect pest populations. Nature, 218:368-369.

Deines AM, Chen VC, & Landis WG. 2005. Modeling the Risks of Nonindigenous Species Introductions Using a Patch-Dynamics Approach Incorporating Contaminant Effects as a Disturbance. Risk Analysis: An International Journal. 25:1637-1651.

Del Rio-Galvan SL, Flores AE, Barrera R, Lopez-Monroy B, Felix G, Amador M, & Ponce-Garcia G. 2016. Susceptibility to temephos and spinosad in *Aedes aegypti* (Diptera: Culicidae) from Puerto Rico. Journal of medical entomology. 53:1211-1217.

EPA. Experimental Use Permit Issued for 93167-EUP-2 to Allow for Releases of OX5034 Aedes aegypti in Florida and Texas Insurance Letter. 2020. Available from https://www.regulations.gov/document?D=EPA-HQ-OPP-2019-0274-0353 (accessed 25 June 2020).

EPA. United States Environmental Protection Agency. 2001. Insect Growth Regulators: S-Hydroprene (128966), S-Kinoprene (107502), Methoprene (105401), S-Methoprene

(105402) Fact Sheet. Available from https://www3.epa.gov/pesticides/chem_search/reg_actions/registration/fs_G-107_06-Dec-01.pdf (accessed 17 March 2019).

EPA. United States Environmental Protection Agency. 2016a. Controlling mosquitos at the larval stage. Available from https://www.epa.gov/mosquitocontrol/controlling-mosquitoes-larval-stage (accessed 17 March 2019).

EPA. United States Environmental Protection Agency. 2016b. Success in mosquito control: an integrated approach. Available from https://www.epa.gov/mosquitocontrol/success-mosquito-control-integrated-approach#larval%20stage (accessed 17 March 2019).

EPA. United States Environmental Protection Agency. 2016c. Malathion. Available from https://www.epa.gov/mosquitocontrol/malathion (accessed 17 March 2019).

EPA. United States Environmental Protection Agency. 2018. Naled for mosquito control. Available from https://www.epa.gov/mosquitocontrol/naled-mosquito-control (accessed 17 March 2019).

EPA. United States Environmental Protection Agency. 2019. Permethrin, resmethrin, d-phenothrin (Sumithrin ®): synthetic pyrethroids for mosquito control. Available from https://www.epa.gov/mosquitocontrol/permethrin-resmethrin-d-phenothrin-sumithrinr-synthetic-pyrethroids-mosquito-control (accessed 17 March 2019).

Focks DA, Brenner RJ, Hayes J, & Daniels E. 2000. Transmission thresholds for dengue in terms of Aedes aegypti pupae per person with discussion of their utility in source reduction efforts. The American journal of tropical medicine and hygiene, 62:11-18.

Foster GG, Whitten MJ, Prout T, & Gill R. 1972. Chromosome rearrangements for the control of insect pests. Science, 176:875-880.

Fox I. 1953. Notes on Puerto Rican mosquitoes including a new species of Culex. The Journal of parasitology, 39:178-182.

Fraser Jr, MJ. 2012. Insect transgenesis: current applications and future prospects. Annual review of entomology, 57:267-289.

Gantz VM, Jasinskiene N, Tatarenkova O, Fazekas A, Macias VM, Bier E, & James AA. 2015. Highly efficient Cas9-mediated gene drive for population modification of the malaria vector mosquito Anopheles stephensi. Proceedings of the National Academy of Sciences, 112:E6736-E6743.

Gorham JR. 1976. Orchid pollination by Aedes mosquitoes in Alaska. American Midland Naturalist, 208-210.

Gorham JR. 1976. Orchid pollination by Aedes mosquitoes in Alaska. *American Midland Naturalist*, 208-210.

Graham SE, Chariton AA, & Landis WG. 2019. Using Bayesian networks to predict risk to estuary water quality and patterns of benthic environmental DNA in Queensland. Integrated environmental assessment and management, 15:93-111.

Gratz SJ, Ukken FP, Rubinstein CD, Thiede G, Donohue LK, Cummings AM, & O'Connor-Giles KM. 2014. Highly specific and efficient CRISPR/Cas9-catalyzed homology-directed repair in Drosophila. Genetics. 196:961-971.

Hamer WH. 1906. The Milroy Lectures on Epidemic Diseases in England: The Evidence of Variability and of Persistency of Type; Delivered Before the Royal College of Physicians of London, March 1st, 6th, and 8th, 1906. Bedford Press.

Hammond A, Galizi R, Kyrou K, Simoni A, Siniscalchi C, Katsanos D, Gribble M, Baker D, Marois E, Russell S, & Burt, A. 2016. A CRISPR-Cas9 gene drive system targeting

female reproduction in the malaria mosquito vector Anopheles gambiae. Nature biotechnology, 34:78-83.

Harrington LC, Scott TW, Lerdthusnee K, Coleman RC, Costero A, Clark GG, Jones JJ, Kitthawee S, Kittayapong P, Sithiprasasna R, & Edman, JD. 2005. Dispersal of the dengue vector Aedes aegypti within and between rural communities. The American journal of tropical medicine and hygiene, 72:209-220.

Hayes EH & Landis WG. 2004. Regional ecological risk assessment of a near shore marine environment: Cherry Point, WA. Human and Ecological Risk Assessment 10:299-325.

Hickey WA, & Craig Jr, GB. 1966. Genetic distortion of sex ratio in a mosquito, *Aedes aegypti*. Genetics. 53:1177.

Hines EH & Landis WG. 2013. Regional risk assessment of the Puyallup River watershed and the evaluation of low impact development in meeting management goals. Integrated Environmental Assessment and Management 10:269-278.

Hiraizumi, Y, & Crow JF. 1960. Heterozygous effects on viability, fertility, rate of development, and longevity of Drosophila chromosomes that are lethal when homozygous. Genetics, 45:1071.

Humanitarian OpenStreetMap. 2020. HOTOSM Puerto Rico Buildings (OpenStreetMap Export). Available online at https://data.humdata.org/dataset/hotosm_pri_buildings (accessed 12 December 2019).

Jasin, M. 1996. Genetic manipulation of genomes with rare-cutting endonucleases. Trends in Genetics, 12:224-228.

Kermack WO, & McKendrick AG. 1927. A contribution to the mathematical theory of epidemics. Proceedings of the royal society of London. Series A, Containing papers of a mathematical and physical character, 115:700-721.

Knott GJ & Doudna JA. 2018. CRISPR-Cas guides the future of genetic engineering. Science, 361(6405), 866-869.

Krol L, Gorsich EE, Hunting ER, Govender D, van Bodegom PM, & Schrama M. 2019. Eutrophication governs predator-prey interactions and temperature effects in Aedes aegypti populations. Parasites & vectors, 12:179.

Kyrou K, Hammond AM, Galizi R, Kranjc N, Burt A, Beaghton AK, Nolan T & Crisanti A. 2018. A CRISPR–Cas9 gene drive targeting doublesex causes complete population suppression in caged Anopheles gambiae mosquitoes. Nature biotechnology. 36:1062-1066.

Lahondère C, Vinauger C, Okubo RP, Wolff GH, Chan JK, Akbari OS, & Riffell JA. 2020. The olfactory basis of orchid pollination by mosquitoes. Proceedings of the National Academy of Sciences, 117:708-716.

Landis WG, Chu VR, Graham SE, Harris MJ, Markiewicz AJ, Mitchell CJ, von Stackelberg KE, & Stark JD. 2020. Integration of Chlorpyrifos Acetylcholinesterase Inhibition, Water Temperature, and Dissolved Oxygen Concentration into a Regional Scale Multiple Stressor Risk Assessment Estimating Risk to Chinook Salmon. Integrated Environmental Assessment and Management, 16:28-42.

Landis WG, Markiewicz AJ, Ayre KK, Johns AF, Harris MJ, Stinson JM, & Summers HM. 2017. A general risk-based adaptive management scheme incorporating the Bayesian Network Relative Risk Model with the South River, Virginia, as case study. Integrated environmental assessment and management, 13:115-126.

Landis, W. G., and Wiegers J. 1997. Design considerations and a suggested approach for regional and comparative ecological risk assessment. Human and Ecological Risk Assessment. 3:287-297.

Landis, W. G., and Wiegers J. 2005. Introduction to the regional risk assessment using the relative risk model. In M. C. Newman (editor), Environmental and Ecological Risk Assessment, CRC Press, Boca Raton, FL. Pp. 11-35.

Landis, W. G., J. L. Durda, M. L. Brooks, and P. M. Chapman. 2013. Ecological risk assessment in the context of global climate change. Environmental Toxicology and Chemistry 32:79-92.

Leftwich PT, Edgington MP, Harvey-Samuel T, Carabajal Paladino LZ, Norman VC, & Alphey L. 2018. Recent advances in threshold-dependent gene drives for mosquitoes. Biochemical Society Transactions, 46:1203-1212.

Macias VM, Ohm JR, & Rasgon JL. 2017. Gene drive for mosquito control: where did it come from and where are we headed? International Journal of Environmental Research and Public Health, 14:1006.

Marcot BG, Steventon JD, Sutherland GD, & McCann RK. 2006. Guidelines for developing and updating Bayesian belief networks applied to ecological modeling and conservation. Canadian Journal of Forest Research 36:3063-3074.

Marcot BG. 2012. Metrics for evaluating performance and uncertainty of Bayesian network models. Ecological modelling. 230:50-62.

Marshall JM, Buchman A, & Akbari OS. 2017. Overcoming evolved resistance to population-suppressing homing-based gene drives. Scientific reports, 7(1), 1-12.

McClintock, B. 1951. Chromosome organization and genic expression. Cold Spring Harb. Symp. Quant. Biol. 16:13-47.

McClintock, B. 1956. Intranuclear systems controlling gene action and mutation. Brookhaven Symp. Biol. 8:58–74.

McDermott SR, and Noor MA. 2010. The role of meiotic drive in hybrid male sterility. Philos. Trans. R. Soc. Lond B. Biol. Sci. 365:1265-1272.

NASEM. National Academies of Sciences, Engineering, and Medicine. 2016. Gene Drives on the Horizon: Advancing Science, Navigating Uncertainty, and Aligning Research with Public Values. The National Academies Press, Washington, DC. Available online at https://doi.org/10.17226/23405 (accessed 17 October 2018).

Nelson MJ. 1986. *Aedes aegypti*: Biology and Ecology. Pan American Health Organization. Washington, D.C.

NLCD. Multi-Resolution Land Characteristics Consortium (U.S.). 2016. National land cover dataset. [Research Triangle Park, NC] :[Multi-Resolution Land Characteristics Consortium].

Noble C, Olejarz J, Esvelt KM, Church GM, & Nowak MA. 2017. Evolutionary dynamics of CRISPR gene drives. Science advances, 3(4), e1601964.

Norsys. 2014. Netica. Vancouver (CA): Norsys Software Corporation.

Noyd DH, & Sharp TM. 2015. Recent advances in dengue: relevance to Puerto Rico. Puerto Rico health sciences journal, 34:65.

Pearl J, & Mackenzie D. 2018. The book of why: the new science of cause and effect. Basic Books.

Pollino CA, Woodberry O, Nicholson A, Korb K, & Hart BT. 2007. Parameterisation and evaluation of a Bayesian network for use in an ecological risk assessment. Environmental Modelling & Software. 22:1140-1152.

Prowse TA, Cassey P, Ross JV, Pfitzner C, Wittmann TA, & Thomas P. 2017. Dodging silver bullets: good CRISPR gene-drive design is critical for eradicating exotic vertebrates. Proceedings of the Royal Society B: Biological Sciences, 284(1860), 20170799 (accessed 5 January 2020).

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rhoades MM, & Dempsey E. 1985. Structural heterogeneity of chromosome 10 in races of maize and teosinte. Pp. 1-18 in Plant Genetics, M. Freeling, ed. New York: Alan R. Liss.

Rittel HW, & Webber MM. 1973. Dilemmas in a general theory of planning. Policy sciences. 4:155-169.

Sánchez CHM, Wu SL, Bennett JB, & Marshall JM. 2020. MGDrivE: A modular simulation framework for the spread of gene drives through spatially explicit mosquito populations. Methods in Ecology and Evolution, 11:229-239.

Scott MJ, Gould F, Lorenzen M, Grubbs N, Edwards O, & O'Brochta D. 2018. Agricultural production: assessment of the potential use of Cas9-mediated gene drive systems for agricultural pest control. Journal of Responsible Innovation. 5(sup1):S98-S120.

Scott MJ, Gould F, Lorenzen M, Grubbs N, Edwards O, & O'Brochta D 2018. Agricultural production: assessment of the potential use of Cas9-mediated gene drive systems for agricultural pest control. Journal of Responsible Innovation, 5:S98-S120.

SERCC. The Southeast Regional Climate Center. 2007. Historical Climate Summaries for Puerto Rico and the U.S. Virgin Islands. Southeast Regional Climate Center, University of North Carolina at Chapel Hill. Chapel Hill, North Carolina. Available from: https://sercc.com/climateinfo/historical/historical_pr.html (accessed 1 July 2020).

Silver LM. 1993. The peculiar journey of a selfish chromosome: Mouse t haplotypes and meiotic drive. Trends Genet. 9:250-254.

Soderlund DM, Clark JM, Sheets LP, Mullin LS, Piccirillo VJ, Sargent D, Stevens JT, & Weiner ML. 2002. Mechanisms of pyrethroid neurotoxicity: implications for cumulative risk assessment. Toxicology. 171:3-59.

Sweeny TL, & Barr AR. 1978. Sex ratio distortion caused by meiotic drive in a mosquito, *Culex pipiens* L. Genetics. 88:427-446.

Thien LB. 1969. Mosquito pollination of *Habenaria obtusata* (*Orchidaceae*). *American Journal of Botany*, *56*:232-237.

U.S. Census Bureau. 2019. QuickFacts Ponce Municipio, Puerto Rico. Available online at https://www.census.gov/quickfacts/poncemunicipiopuertorico (accessed 26 February 2019)

Webber BL, Raghu S, & Edwards OR. 2015. Opinion: Is CRISPR-based gene drive a biocontrol silver bullet or global conservation threat? Proceedings of the National Academy of Sciences, 112(34), 10565-10567.

Weterings E, & Chen DJ. 2008. The endless tale of non-homologous end-joining. Cell research. 18:114-124.

Wicker T, Sabot F, Hua-Van A, Bennetzen JL, Capy P, Chalhoub B, Flavell A, Leroy P, Morgante M, Panaud O, Paux E, SanMiguel P, and Schulman A.H. 2007. A unified classification system for eukaryotic transposable elements. Nat. Rev. Genet. 8:973-982.

Wiegers JK, Feder HM, Mortensen LS, Shaw DG, Wilson VJ, & Landis WG. 1998. A regional multiple-stressor rank-based ecological risk assessment for the fjord of Port Valdez, Alaska. Human and Ecological Risk Assessment, 4:1125-1173.

Windbichler N, Menichelli M, Papathanos PA, Thyme SB, Li H, Ulge UY, Hovde BT, Baker D, Monnat RJ, Burt A & Crisanti A. 2011. A synthetic homing endonuclease-based gene drive system in the human malaria mosquito. Nature, 473(7346), 212-215.WHO. World Health Organization. 2017. Vector-borne diseases. Available from https://www.who.int/news-room/fact-sheets/detail/vector-borne-diseases (accessed 19 March 2019).

World Health Organization [WHO]: World Malaria Report 2019. 2019. Available online: https://www.who.int/publications-detail/world-malaria-report-2019 (accessed on 2 January 2020).

Appendix 1. Dengue transmission risk estimates for maximum simulation time.

Risk Region	Rate of accurate HDR	Release ratio	Weekly releases	Below transmission threshold (%)	At transmission threshold (%)	Above transmission threshold (%)
1	CC	CC	CC	79.1	7.69	13.2
2	CC	CC	CC	59.5	7.69	32.8
3	CC	CC	CC	43.3	7.69	49.1
1	0.999	3340	13	79.1	7.69	13.2
2	0.999	3340	13	58.1	7.69	34.2
3	0.999	3340	13	40	7.69	52.3
1	0.999	10021	13	79.1	7.69	13.2
2	0.999	10021	13	58.1	7.69	34.2
3	0.999	10021	13	40	7.69	52.3
1	0.999	16702	13	79.1	7.69	13.2
2	0.999	16702	13	58.1	7.69	34.2
3	0.999	16702	13	40	7.69	52.3
1	0.999	3340	26	79.1	7.69	13.2
2	0.999	3340	26	58.1	7.69	34.2
3	0.999	3340	26	40	7.69	52.3
1	0.999	10021	26	79.1	7.69	13.2
2	0.999	10021	26	58.1	7.69	34.2
3	0.999	10021	26	40	7.69	52.3
1	0.999	16702	26	79.1	7.69	13.2
2	0.999	16702	26	58.1	7.69	34.2
3	0.999	16702	26	40	7.69	52.3
1	0.999	3340	39	79.1	7.69	13.2
2	0.999	3340	39	58.1	7.69	34.2
3	0.999	3340	39	40	7.69	52.3
1	0.999	10021	39	79.1	7.69	13.2
2	0.999	10021	39	58.1	7.69	34.2
3	0.999	10021	39	40	7.69	52.3
1	0.999	16702	39	79.1	7.69	13.2
2	0.999	16702	39	58.1	7.69	34.2
3	0.999	16702	39	40	7.69	52.3
1	0.9999	3340	13	79.1	7.69	13.2
2	0.9999	3340	13	58.1	7.69	34.2
3	0.9999	3340	13	40	7.69	52.3
1	0.9999	10021	13	79.1	7.69	13.2
2	0.9999	10021	13	58.4	7.69	33.9
3	0.9999	10021	13	40.6	7.69	51.7
1	0.9999	16702	13	79.1	7.69	13.2
2	0.9999	16702	13	59.5	7.69	32.9
3	0.9999	16702	13	42.5	7.69	49.8

Risk Region	Rate of accurate HDR	Release ratio	Weekly releases	Below transmission threshold (%)	At transmission threshold (%)	Above transmission threshold (%)
1	0.9999	3340	26	79.1	7.69	13.2
2	0.9999	3340	26	59.5	7.69	32.9
3	0.9999	3340	26	42.5	7.69	49.8
1	0.9999	10021	26	79.1	7.69	13.2
2	0.9999	10021	26	63.2	7.69	29.1
3	0.9999	10021	26	49.4	7.69	42.9
1	0.9999	16702	26	79.1	7.69	13.2
2	0.9999	16702	26	61.5	7.69	30.8
3	0.9999	16702	26	46.3	7.69	46
1	0.9999	3340	39	79.1	7.69	13.2
2	0.9999	3340	39	59.1	7.69	33.2
3	0.9999	3340	39	41.9	7.69	50.4
1	0.9999	10021	39	79.1	7.69	13.2
2	0.9999	10021	39	63.2	7.69	29.1
3	0.9999	10021	39	49.4	7.69	42.9
1	0.9999	16702	39	79.1	7.69	13.2
2	0.9999	16702	39	66.6	7.69	25.7
3	0.9999	16702	39	55.7	7.69	36.6
1	0.99995	3340	13	79.1	7.69	13.2
2	0.99995	3340	13	60.1	7.69	32.2
3	0.99995	3340	13	43.8	7.69	48.5
1	0.99995	10021	13	79.1	7.69	13.2
2	0.99995	10021	13	62.5	7.69	29.8
3	0.99995	10021	13	48.2	7.69	44.1
1	0.99995	16702	13	79.1	7.69	13.2
2	0.99995	16702	13	64.9	7.69	27.4
3	0.99995	16702	13	52.6	7.69	39.8
1	0.99995	3340	26	79.1	7.69	13.2
2	0.99995	3340	26	63.2	7.69	29.1
3	0.99995	3340	26	49.4	7.69	42.9
1	0.99995	10021	26	79.1	7.69	13.2
2	0.99995	10021	26	66.6	7.69	25.7
3	0.99995	10021	26	55.7	7.69	36.6
1	0.99995	16702	26	79.1	7.69	13.2
2	0.99995	16702	26	64.9	7.69	27.4
3	0.99995	16702	26	52.6	7.69	39.8
1	0.99995	3340	39	79.1	7.69	13.2
2	0.99995	3340	39	62.5	7.69	29.8
3	0.99995	3340	39	48.2	7.69	44.1

Risk Region	Rate of accurate HDR	Release ratio	Weekly releases	Below transmission threshold (%)	At transmission threshold (%)	Above transmission threshold (%)
1	0.99995	10021	39	79.1	7.69	13.2
2	0.99995	10021	39	65.6	7.69	26.7
3	0.99995	10021	39	53.8	7.69	38.5
1	0.99995	16702	39	79.1	7.69	13.2
2	0.99995	16702	39	64.9	7.69	27.4
3	0.99995	16702	39	52.6	7.69	39.8
1	0.99999	3340	13	79.1	7.69	13.2
2	0.99999	3340	13	68	7.69	24.3
3	0.99999	3340	13	58.2	7.69	34.1
1	0.99999	10021	13	79.1	7.69	13.2
2	0.99999	10021	13	68.6	7.69	23.7
3	0.99999	10021	13	59.5	7.69	32.8
1	0.99999	16702	13	79.1	7.69	13.2
2	0.99999	16702	13	70	7.69	22.3
3	0.99999	16702	13	62	7.69	30.3
1	0.99999	3340	26	79.1	7.69	13.2
2	0.99999	3340	26	69.7	7.69	22.6
3	0.99999	3340	26	61.3	7.69	31
1	0.99999	10021	26	79.1	7.69	13.2
2	0.99999	10021	26	71.4	7.69	20.9
3	0.99999	10021	26	64.5	7.69	27.8
1	0.99999	16702	26	79.1	7.69	13.2
2	0.99999	16702	26	72.7	7.69	19.6
3	0.99999	16702	26	67	7.69	25.3
1	0.99999	3340	39	79.1	7.69	13.2
2	0.99999	3340	39	72.4	7.69	19.9
3	0.99999	3340	39	66.4	7.69	25.9
1	0.99999	10021	39	79.1	7.69	13.2
2	0.99999	10021	39	73.1	7.69	19.2
3	0.99999	10021	39	67.6	7.69	24.7
1	0.99999	16702	39	79.1	7.69	13.2
2	0.99999	16702	39	74.1	7.69	18.2
3	0.99999	16702	39	69.5	7.69	22.8

Appendix 2. Zika transmission risk estimates for maximum simulation time.

Risk Region	Rate of accurate HDR	Release ratio	Weekly releases	Zero transmission (%)	Low transmission (%)	Medium transmission (%)	High transmission (%)
1	CC	CC	CC	7.14	7.14	7.14	78.6
2	CC	CC	CC	7.14	7.14	7.14	78.6
3	CC	CC	CC	7.14	7.14	7.14	78.6
1	0.999	3340	13	11.5	15.9	33.1	39.6
2	0.999	3340	13	15.6	11.7	9.75	62.9
3	0.999	3340	13	10.1	8.29	8.16	73.5
1	0.999	10021	13	11.5	15.9	33.1	39.6
2	0.999	10021	13	15.6	11.7	9.75	62.9
3	0.999	10021	13	10.1	8.29	8.16	73.5
1	0.999	16702	13	11.5	15.9	33.1	39.6
2	0.999	16702	13	15.6	11.7	9.75	62.9
3	0.999	16702	13	10.1	8.29	8.16	73.5
1	0.999	3340	26	11.5	15.9	33.1	39.6
2	0.999	3340	26	15.6	11.7	9.75	62.9
3	0.999	3340	26	10.1	8.29	8.16	73.5
1	0.999	10021	26	11.5	15.9	33.1	39.6
2	0.999	10021	26	15.6	11.7	9.75	62.9
3	0.999	10021	26	10.1	8.29	8.16	73.5
1	0.999	16702	26	11.5	15.9	33.1	39.6
2	0.999	16702	26	15.6	11.7	9.75	62.9
3	0.999	16702	26	10.1	8.29	8.16	73.5
1	0.999	3340	39	11.5	15.9	33.1	39.6
2	0.999	3340	39	15.6	11.7	9.75	62.9
3	0.999	3340	39	10.1	8.29	8.16	73.5
1	0.999	10021	39	11.5	15.9	33.1	39.6
2	0.999	10021	39	15.6	11.7	9.75	62.9
3	0.999	10021	39	10.1	8.29	8.16	73.5
1	0.999	16702	39	11.5	15.9	33.1	39.6
2	0.999	16702	39	15.6	11.7	9.75	62.9
3	0.999	16702	39	10.1	8.29	8.16	73.5
1	0.9999	3340	13	11.5	15.9	33.1	39.6
2	0.9999	3340	13	15.6	11.7	9.75	62.9
3	0.9999	3340	13	10.1	8.29	8.16	73.5
1	0.9999	10021	13	12.7	15.7	32.6	39
2	0.9999	10021	13	16.6	11.7	9.75	61.9
3	0.9999	10021	13	11.4	8.29	8.16	72.2
1	0.9999	16702	13	16.5	15.3	31.1	37
2	0.9999	16702	13	19.8	11.7	9.75	58.8
3	0.9999	16702	13	15.2	8.29	8.16	68.4

Risk Region	Rate of accurate HDR	Release ratio	Weekly releases	Zero transmission (%)	Low transmission (%)	Medium transmission (%)	High transmission (%)
1	0.9999	3340	26	16.5	15.3	31.1	37
2	0.9999	3340	26	19.8	11.7	9.75	58.8
3	0.9999	3340	26	15.2	8.29	8.16	68.4
1	0.9999	10021	26	30.5	13.9	25.5	30.1
2	0.9999	10021	26	31.2	11.7	9.75	47.4
3	0.9999	10021	26	29.2	8.29	8.16	54.3
1	0.9999	16702	26	24.2	14.6	28	33.2
2	0.9999	16702	26	26	11.7	9.75	52.6
3	0.9999	16702	26	22.8	8.29	8.16	60.7
1	0.9999	3340	39	15.3	15.5	31.6	37.7
2	0.9999	3340	39	18.7	11.7	9.75	59.8
3	0.9999	3340	39	13.9	8.29	8.16	69.6
1	0.9999	10021	39	30.5	13.9	25.5	30.1
2	0.9999	10021	39	31.2	11.7	9.75	47.4
3	0.9999	10021	39	29.2	8.29	8.16	54.3
1	0.9999	16702	39	43.2	12.7	20.4	23.7
2	0.9999	16702	39	41.6	11.7	9.75	37
3	0.9999	16702	39	42	8.29	8.16	41.6
1	0.99995	3340	13	19.1	15.1	30	35.8
2	0.99995	3340	13	21.8	11.7	9.75	56.7
3	0.99995	3340	13	17.7	8.29	8.16	65.8
1	0.99995	10021	13	28	14.2	26.5	31.3
2	0.99995	10021	13	29.1	11.7	9.75	49.5
3	0.99995	10021	13	26.7	8.29	8.16	56.9
1	0.99995	16702	13	36.9	13.3	22.9	26.9
2	0.99995	16702	13	36.4	11.7	9.75	42.2
3	0.99995	16702	13	35.6	8.29	8.16	48
1	0.99995	3340	26	30.5	13.9	25.5	30.1
2	0.99995	3340	26	31.2	11.7	9.75	47.4
3	0.99995	3340	26	29.2	8.29	8.16	54.3
1	0.99995	10021	26	43.2	12.7	20.4	23.7
2	0.99995	10021	26	41.6	11.7	9.75	37
3	0.99995	10021	26	42	8.29	8.16	41.6
1	0.99995	16702	26	36.9	13.3	22.9	26.9
2	0.99995	16702	26	36.4	11.7	9.75	42.2
3	0.99995	16702	26	35.6	8.29	8.16	48
1	0.99995	3340	39	28	14.2	26.5	31.3
2	0.99995	3340	39	29.1	11.7	9.75	49.5
3	0.99995	3340	39	26.7	8.29	8.16	56.9

Risk Region	Rate of accurate HDR	Release ratio	Weekly releases	Zero transmission (%)	Low transmission (%)	Medium transmission (%)	High transmission (%)
1	0.99995	10021	39	39.4	13.1	21.9	25.6
2	0.99995	10021	39	38.4	11.7	9.75	40.1
3	0.99995	10021	39	38.1	8.29	8.16	45.4
1	0.99995	16702	39	36.9	13.3	22.9	26.9
2	0.99995	16702	39	36.4	11.7	9.75	42.2
3	0.99995	16702	39	35.6	8.29	8.16	48
1	0.99999	3340	13	48.3	12.2	18.3	21.2
2	0.99999	3340	13	45.7	11.7	9.75	32.8
3	0.99999	3340	13	47.1	8.29	8.16	36.5
1	0.99999	10021	13	50.9	11.9	17.3	19.9
2	0.99999	10021	13	47.8	11.7	9.75	30.8
3	0.99999	10021	13	49.6	8.29	8.16	33.9
1	0.99999	16702	13	56	11.4	15.3	17.3
2	0.99999	16702	13	51.9	11.7	9.75	26.6
3	0.99999	16702	13	54.7	8.29	8.16	28.8
1	0.99999	3340	26	54.7	11.5	15.8	18
2	0.99999	3340	26	50.9	11.7	9.75	27.7
3	0.99999	3340	26	53.4	8.29	8.16	30.1
1	0.99999	10021	26	61.1	10.9	13.3	14.8
2	0.99999	10021	26	56.1	11.7	9.75	22.5
3	0.99999	10021	26	59.8	8.29	8.16	23.7
1	0.99999	16702	26	66.1	10.4	11.2	12.2
2	0.99999	16702	26	60.2	11.7	9.75	18.3
3	0.99999	16702	26	64.9	8.29	8.16	18.6
1	0.99999	3340	39	64.9	10.5	11.7	12.9
2	0.99999	3340	39	59.2	11.7	9.75	19.4
3	0.99999	3340	39	63.6	8.29	8.16	19.9
1	0.99999	10021	39	67.4	10.3	10.7	11.6
2	0.99999	10021	39	61.3	11.7	9.75	17.3
3	0.99999	10021	39	66.2	8.29	8.16	17.3
1	0.99999	16702	39	71.2	9.88	9.19	9.7
2	0.99999	16702	39	64.4	11.7	9.75	14.2
3	0.99999	16702	39	70	8.29	8.16	13.5

Appendix 3. Dengue transmission risk estimates for minimum population timepoint.

Risk Region	Rate of accurate HDR	Release ratio	Weekly releases	Below transmission threshold (%)	At transmission threshold (%)	Above transmission threshold (%)
1	CC	CC	CC	79.1	7.69	13.2
2	CC	CC	CC	59.5	7.69	32.8
3	CC	CC	CC	43.3	7.69	49.1
1	0.999	3340	13	81.9	7.69	10.5
2	0.999	3340	13	80	7.69	12.3
3	0.999	3340	13	76.4	7.69	15.9
1	0.999	10021	13	81.9	7.69	10.4
2	0.999	10021	13	80	7.69	12.3
3	0.999	10021	13	76.6	7.69	15.7
1	0.999	16702	13	81.9	7.69	10.4
2	0.999	16702	13	80	7.69	12.3
3	0.999	16702	13	76.6	7.69	15.7
1	0.999	3340	26	81.9	7.69	10.4
2	0.999	3340	26	80	7.69	12.3
3	0.999	3340	26	76.6	7.69	15.7
1	0.999	10021	26	81.9	7.69	10.4
2	0.999	10021	26	80	7.69	12.3
3	0.999	10021	26	76.6	7.69	15.7
1	0.999	16702	26	81.9	7.69	10.4
2	0.999	16702	26	80	7.69	12.3
3	0.999	16702	26	76.6	7.69	15.7
1	0.999	3340	39	81.9	7.69	10.4
2	0.999	3340	39	80	7.69	12.3
3	0.999	3340	39	76.6	7.69	15.7
1	0.999	10021	39	81.9	7.69	10.4
2	0.999	10021	39	80	7.69	12.3
3	0.999	10021	39	76.6	7.69	15.7
1	0.999	16702	39	81.9	7.69	10.4
2	0.999	16702	39	80	7.69	12.3
3	0.999	16702	39	76.6	7.69	15.7
1	0.9999	3340	13	81.9	7.69	10.4
2	0.9999	3340	13	80	7.69	12.3
3	0.9999	3340	13	76.6	7.69	15.7
1	0.9999	10021	13	81.9	7.69	10.4
2	0.9999	10021	13	80	7.69	12.3
3	0.9999	10021	13	76.6	7.69	15.7

Risk Region	Rate of accurate HDR	Release ratio	Weekly releases	Below transmission threshold (%)	At transmission threshold (%)	Above transmission threshold (%)
1	0.9999	16702	13	81.9	7.69	10.4
2	0.9999	16702	13	80	7.69	12.3
3	0.9999	16702	13	76.6	7.69	15.7
1	0.9999	3340	26	81.9	7.69	10.4
2	0.9999	3340	26	80	7.69	12.3
3	0.9999	3340	26	76.6	7.69	15.7
1	0.9999	10021	26	81.9	7.69	10.4
2	0.9999	10021	26	80	7.69	12.3
3	0.9999	10021	26	76.6	7.69	15.7
1	0.9999	16702	26	81.9	7.69	10.4
2	0.9999	16702	26	80	7.69	12.3
3	0.9999	16702	26	76.6	7.69	15.7
1	0.9999	3340	39	81.9	7.69	10.4
2	0.9999	3340	39	80	7.69	12.3
3	0.9999	3340	39	76.6	7.69	15.7
1	0.9999	10021	39	81.9	7.69	10.4
2	0.9999	10021	39	80	7.69	12.3
3	0.9999	10021	39	76.6	7.69	15.7
1	0.9999	16702	39	81.9	7.69	10.4
2	0.9999	16702	39	80	7.69	12.3
3	0.9999	16702	39	76.6	7.69	15.7
1	0.99995	3340	13	81.9	7.69	10.4
2	0.99995	3340	13	80	7.69	12.3
3	0.99995	3340	13	76.6	7.69	15.7
1	0.99995	10021	13	81.9	7.69	10.4
2	0.99995	10021	13	80	7.69	12.3
3	0.99995	10021	13	76.6	7.69	15.7
1	0.99995	16702	13	81.9	7.69	10.4
2	0.99995	16702	13	80	7.69	12.3
3	0.99995	16702	13	76.6	7.69	15.7
1	0.99995	3340	26	81.9	7.69	10.4
2	0.99995	3340	26	80	7.69	12.3
3	0.99995	3340	26	76.6	7.69	15.7
1	0.99995	10021	26	81.9	7.69	10.4
2	0.99995	10021	26	80	7.69	12.3
3	0.99995	10021	26	76.6	7.69	15.7
1	0.99995	16702	26	81.9	7.69	10.4
2	0.99995	16702	26	80	7.69	12.3
3	0.99995	16702	26	76.6	7.69	15.7

Risk Region	Rate of accurate HDR	Release ratio	Weekly releases	Below transmission threshold (%)	At transmission threshold (%)	Above transmission threshold (%)
1	0.99995	3340	39	81.9	7.69	10.4
2	0.99995	3340	39	80	7.69	12.3
3	0.99995	3340	39	76.6	7.69	15.7
1	0.99995	10021	39	81.9	7.69	10.4
2	0.99995	10021	39	80	7.69	12.3
3	0.99995	10021	39	76.6	7.69	15.7
1	0.99995	16702	39	81.9	7.69	10.4
2	0.99995	16702	39	80	7.69	12.3
3	0.99995	16702	39	76.6	7.69	15.7
1	0.99999	3340	13	81.9	7.69	10.4
2	0.99999	3340	13	80	7.69	12.3
3	0.99999	3340	13	76.5	7.69	15.8
1	0.99999	10021	13	81.9	7.69	10.4
2	0.99999	10021	13	80	7.69	12.3
3	0.99999	10021	13	76.6	7.69	15.7
1	0.99999	16702	13	81.9	7.69	10.4
2	0.99999	16702	13	80	7.69	12.3
3	0.99999	16702	13	76.6	7.69	15.7
1	0.99999	3340	26	81.9	7.69	10.4
2	0.99999	3340	26	80	7.69	12.3
3	0.99999	3340	26	76.5	7.69	15.8
1	0.99999	10021	26	81.9	7.69	10.4
2	0.99999	10021	26	80	7.69	12.3
3	0.99999	10021	26	76.6	7.69	15.7
1	0.99999	16702	26	81.9	7.69	10.4
2	0.99999	16702	26	80	7.69	12.3
3	0.99999	16702	26	76.6	7.69	15.7
1	0.99999	3340	39	81.9	7.69	10.4
2	0.99999	3340	39	80	7.69	12.3
3	0.99999	3340	39	76.6	7.69	15.7
1	0.99999	10021	39	81.9	7.69	10.4
2	0.99999	10021	39	80	7.69	12.3
3	0.99999	10021	39	76.6	7.69	15.7
1	0.99999	16702	39	81.9	7.69	10.4
2	0.99999	16702	39	80	7.69	12.3
3	0.99999	16702	39	76.6	7.69	15.7

Appendix 4. Zika transmission risk estimates for minimum population timepoint.

Risk Region	Rate of accurate HDR	Release ratio	Weekly releases	Zero transmission (%)	Low transmission (%)	Medium transmission (%)	High transmission (%)
1	CC	CC	CC	7.14	7.14	7.14	78.6
2	CC	CC	CC	7.14	7.14	7.14	78.6
3	CC	CC	CC	7.14	7.14	7.14	78.6
1	0.999	3340	13	25.6	20.6	19	34.8
2	0.999	3340	13	23.4	16.8	42.3	17.5
3	0.999	3340	13	18.7	28.8	33.3	19.3
1	0.999	10021	13	36.8	42.9	8.97	11.4
2	0.999	10021	13	23.7	43.4	18.4	14.5
3	0.999	10021	13	21	33.9	27.5	17.7
1	0.999	16702	13	49.1	30.6	8.97	11.4
2	0.999	16702	13	32.5	37.6	15.6	14.4
3	0.999	16702	13	24.9	41.8	15.6	17.7
1	0.999	3340	26	34.8	40.9	10.2	14.2
2	0.999	3340	26	23.7	42.3	19.4	14.6
3	0.999	3340	26	20.5	32.9	29	17.7
1	0.999	10021	26	59.8	20.6	8.73	10.8
2	0.999	10021	26	38	33.8	14	14.3
3	0.999	10021	26	26.4	44.7	11.3	17.7
1	0.999	16702	26	62.6	17.8	8.73	10.8
2	0.999	16702	26	39.6	32.5	13.6	14.3
3	0.999	16702	26	26.7	45.3	10.3	17.7
1	0.999	3340	39	46.7	33.7	8.73	10.8
2	0.999	3340	39	28.9	41	15.8	14.3
3	0.999	3340	39	24.8	41.5	16	17.7
1	0.999	10021	39	47.1	33.3	8.73	10.8
2	0.999	10021	39	30.8	39.5	15.4	14.3
3	0.999	10021	39	25.4	42.8	14.1	17.7
1	0.999	16702	39	40.7	38.1	9.2	12
2	0.999	16702	39	24	43.7	17.8	14.4
3	0.999	16702	39	22.2	36.4	23.7	17.7

Risk Region	Rate of accurate HDR	Release ratio	Weekly releases	Zero transmission (%)	Low transmission (%)	Medium transmission (%)	High transmission (%)
1	0.9999	3340	13	38.7	34.5	10.9	15.8
2	0.9999	3340	13	25.7	42.5	17.5	14.4
3	0.9999	3340	13	20.8	33.5	28	17.7
1	0.9999	10021	13	56.6	23	8.97	11.4
2	0.9999	10021	13	34.4	36.6	14.7	14.3
3	0.9999	10021	13	24.8	41.5	16	17.7
1	0.9999	16702	13	61	19.4	8.73	10.8
2	0.9999	16702	13	37	34.5	14.2	14.3
3	0.9999	16702	13	24.4	42.8	13.2	17.7
1	0.9999	3340	26	53.4	26.2	8.97	11.4
2	0.9999	3340	26	32.8	37.3	15.5	14.4
3	0.9999	3340	26	25.4	42.8	14.1	17.7
1	0.9999	10021	26	64.2	16.3	8.73	10.8
2	0.9999	10021	26	41.6	30.3	13.8	14.3
3	0.9999	10021	26	32.4	39.6	10.3	17.7
1	0.9999	16702	26	65.4	15.1	8.73	10.8
2	0.9999	16702	26	42.5	29.5	13.7	14.3
3	0.9999	16702	26	38.2	33.9	10.3	17.7
1	0.9999	3340	39	61.8	18.6	8.73	10.8
2	0.9999	3340	39	37.3	34.3	14.1	14.3
3	0.9999	3340	39	26.7	45.3	10.3	17.7
1	0.9999	10021	39	63.4	17.1	8.73	10.8
2	0.9999	10021	39	39.9	32.1	13.7	14.3
3	0.9999	10021	39	26.7	45.3	10.3	17.7
1	0.9999	16702	39	62.2	18.2	8.73	10.8
2	0.9999	16702	39	37.7	34	14	14.3
3	0.9999	16702	39	26.4	44.7	11.3	17.7
1	0.99995	3340	13	45.1	30.6	10.2	14.2
2	0.99995	3340	13	28.9	37.5	18.9	14.7
3	0.99995	3340	13	22.1	36.1	24.2	17.7
1	0.99995	10021	13	58.6	21	8.97	11.4
2	0.99995	10021	13	36.4	34.4	14.9	14.4
3	0.99995	10021	13	27.6	41.5	13.2	17.7

Risk Region	Rate of accurate HDR	Release ratio	Weekly releases	Zero transmission (%)	Low transmission (%)	Medium transmission (%)	High transmission (%)
1	0.99995	16702	13	61	19.4	8.73	10.8
2	0.99995	16702	13	42.9	29	13.9	14.3
3	0.99995	16702	13	36.3	35.8	10.3	17.7
1	0.99995	3340	26	57.8	21	9.2	12
2	0.99995	3340	26	36	35.2	14.5	14.3
3	0.99995	3340	26	26.5	41.2	14.6	17.7
1	0.99995	10021	26	63.4	17.1	8.73	10.8
2	0.99995	10021	26	43.2	28.9	13.6	14.3
3	0.99995	10021	26	33.5	38	10.8	17.7
1	0.99995	16702	26	65	15.5	8.73	10.8
2	0.99995	16702	26	45.1	26.9	13.6	14.3
3	0.99995	16702	26	40.1	32	10.3	17.7
1	0.99995	3340	39	61.4	19	8.73	10.8
2	0.99995	3340	39	38.6	33	14.1	14.3
3	0.99995	3340	39	29.9	42.2	10.3	17.7
1	0.99995	10021	39	63.8	16.7	8.73	10.8
2	0.99995	10021	39	41.2	30.8	13.6	14.3
3	0.99995	10021	39	26.7	45.3	10.3	17.7
1	0.99995	16702	39	62.2	18.2	8.73	10.8
2	0.99995	16702	39	38.3	33.5	13.9	14.3
3	0.99995	16702	39	26.7	45.3	10.3	17.7
1	0.99999	3340	13	58.2	19	9.68	13.1
2	0.99999	3340	13	36.7	32.8	16	14.5
3	0.99999	3340	13	27.5	39	15.1	18.5
1	0.99999	10021	13	61.4	18.2	8.97	11.4
2	0.99999	10021	13	39.9	31.1	14.6	14.4
3	0.99999	10021	13	31.5	37.7	13.2	17.7
1	0.99999	16702	13	60.2	20.2	8.73	10.8
2	0.99999	16702	13	43.2	28.5	14	14.3
3	0.99999	16702	13	42	28.1	12.2	17.7
1	0.99999	3340	26	60.6	19.8	8.73	10.8
2	0.99999	3340	26	40.6	31	14.1	14.3
3	0.99999	3340	26	30.5	39.6	12.2	17.7

Risk Region	Rate of accurate HDR	Release ratio	Weekly releases	Zero transmission (%)	Low transmission (%)	Medium transmission (%)	High transmission (%)
1	0.99999	10021	26	65	15.5	8.73	10.8
2	0.99999	10021	26	49	23	13.7	14.3
3	0.99999	10021	26	45.7	25.9	10.8	17.7
1	0.99999	16702	26	65	15.5	8.73	10.8
2	0.99999	16702	26	52.3	19.7	13.7	14.3
3	0.99999	16702	26	50.9	21.1	10.3	17.7
1	0.99999	3340	39	65	15.5	8.73	10.8
2	0.99999	3340	39	50.6	21.4	13.7	14.3
3	0.99999	3340	39	49.5	22.1	10.8	17.7
1	0.99999	10021	39	65.4	15.1	8.73	10.8
2	0.99999	10021	39	53.6	18.5	13.6	14.3
3	0.99999	10021	39	56.7	15.4	10.3	17.7
1	0.99999	16702	39	64.6	15.9	8.73	10.8
2	0.99999	16702	39	54.9	17.2	13.6	14.3
3	0.99999	16702	39	57.3	14.7	10.3	17.7

Appendix 5. Conceptual model and Bayesian network node derivation and quantification

Variable (node)	Discretization methodology & justification	States
SOURCE		
Region	Risk regions delineated by average mosquito dispersal distance using GIS analysis. Average mosquito flight distance found to be 100m (Harrington et al. 2005; Cox et al. 2007).	RR1
		RR2
		RR3
Landclass	Land class was determined through GIS analysis using National Land Cover Database shapefile. The land class of the subregions were defined by the dominant land cover for each patch.	High-density housing (HDH)
		Low-density housing (LDH)
		Forest (F)
		Non-forest vegetation (NFV)
		Wetlands (W)
Buildings	Number of buildings was determined in GIS using a shapefile available from the Humanitarian OpenStreetMap Team. Shapefile showed polygons of all existing buildings within the study area. Polygons were converted to points.	Low = 2379 buildings
		Medium = 4334 buildings
		High = 11069 buildings
Release ratio	Release ratios reflect the average wild type mosquito population per patch: (total mosquito population / # of patches). Ratios chosen in agreement with previous modeling studies (Robert et al. 2014; Pham et al. 2019; and Sánchez et al. 2020)	1:1 (3340 individuals)
		3:1 (10021 individuals)
		5:1 (16702 individuals)
Accurate HDR	Accurate homology directed repair (HDR). Rate of accurate HDR reflect one level (0.999) shown in previous lab studies (Hammond et al. 2016), and increasing levels theorized to be capable through the use of multiplexing (Marshall et al. 2017). 1 - (accurate HDR) = resistance generation rate.	0.999
		0.9999
		0.99995
		0.99999
Duration of Releases	Frequency of releases in weeks, e.g. 13 = one release every week for 13 weeks. Weeks chosen in agreement with previous modeling studies (Robert et al. 2014; Pham et al. 2019; and Sánchez et al. 2020)	13 weeks
		26 weeks
		39 weeks
STRESSOR		
Wild population		8687 to 16413 (individuals)

Wild population (continued)	Existing mosquito population estimates were obtained by combining the *Ae. aegypti* abundance estimates of the Barrera et al. (2019) study with our GIS analysis of the existing landscape characteristics of Ponce. A 50/50 sex ratio is assumed in the sex of offspring. Original population estimates are for female mosquitoes only, multiplying this by 2 gives an approximation of both male and female mosquitoes. States chosen represent ranges of mosquitoes per region based on Barrera et al. (2017) linear model equations. WildPop (landclass, Buildings) = HDH = 3.4*2*(# of Buildings), LDH = 9.8*2*(# of Buildings), F = 2.6*2*(# of Buildings), W = 2.1*2*(# of Buildings), NFV = 2.7*2*(# of Buildings).	16413 to 20476 (individuals)
		20476 to 30216 (individuals)
		30216 to 66565 (individuals)
		66565 to 84166 (individuals)
		84166 to 217000 (individuals)
Pesticide application	Pesticides chosen for examination are from Agramonte et al. 2017. This study tested various levels of several pyrethroid insecticides on strains of *Ae. aegypti* found in Puerto Rico. Authors examined resistance levels, dose-response relationships show pesticide efficacy. Analysis done in R.	Permethrin
		Etofenprox
		Deltamethrin
		DDT
		Transfluthrin
Mols pesticide (applied)	Pesticides chosen for examination are from Agramonte et al. (2017) Levels of dose developed from natural Jenk's optimization for pesticide levels in Agramonte et al. (2017).	1e-20 to 1.32e-11 (mols)
		1.32e-11 to 3.58e-11 (mols)
		3.58e-11 to 7.73e-11 (mols)
		7.73e-11 to 1.55e-10 (mols)
		1.55e-10 to 2.34e-10 (mols)
		2.34e-10 to 3.09e-10 (mols)
HABITAT		
Human population	Human population estimates derived from GIS analysis of 2010 U.S. Census data.	14170 (persons)
		38629 (persons)
		113430 (persons)
Herd immunity	Levels of herd immunity derived from Focks et al. (2000).	0% (seroprevalence)
		0.33% (seroprevalence)
		0.67% (seroprevalence)
Temperature		22°C

Temperature (continued)	Temperatures that are correlated to dengue thresholds and pupal production are derived from Focks et al. (2000).	24°C
		26°C
		28°C
		30°C
		32°C
Dengue threshold (pupae/1000 people)	Specific dengue transmission thresholds from Focks et al. (2000) table 6. Values are specific thresholds for each combination of herd immunity and temperature.	70 (pupae/1000 people)
		100 (pupae/1000 people)
		130 (pupae/1000 people)
		190 (pupae/1000 people)
		260 (pupae/1000 people)
		380 (pupae/1000 people)
		530 (pupae/1000 people)
		750 (pupae/1000 people)
		1420 (pupae/1000 people)
		1690 (pupae/1000 people)
		2030 (pupae/1000 people)
		2920 (pupae/1000 people)
		4260 (pupae/1000 people)
		4470 (pupae/1000 people)
		9220 (pupae/1000 people)
		9570 (pupae/1000 people)
		14100 (pupae/1000 people)
		30550 (pupae/1000 people)
EFFECTS		
Resulting population after gene drive release	CPT filled out with case learning file from output of MGDrivE simulations. Data was either aggregated to show the population levels for the time period (day) in which the average number of mosquito per patch was at its lowest over the entire simulation, or the resulting population estimates and genotype at the end of the simulation period (6.75 years).	0 (individuals)
		0 to 100 (individuals)
		100 to 2741 (individuals)
		2741 to 7626 (individuals)
		7626 to 13729 (individuals)
		13729 to 65701 (individuals)
Proportion of gene drive resistant		0 to 0.05

Proportion of gene drive resistant (continued)	Resistance to the drive occurs primarily through non-homologous end joining (NHEJ; Unckless et al. 2017). Rates of resistance generation in MGDrivE are determined by 1 - (rate of accurate HDR).	0.05 to 0.25
		0.25 to 0.75
		0.75 to 1
Proportion survival	Pesticides chosen for examination are from Agramonte et al. (2017). This study tested various levels of several pyrethroid insecticides on strains of *Ae. aegypti* found in Puerto Rico. Authors examined resistance levels, dose-response relationships show pesticide efficacy. Analysis done in R. Full range of survival proportion in increments of 0.1 to estimate full range of effect. Survival (Pesticide, Mols) = Permethrin = 1/(1+exp(1.2428*(log(Mols)-log(4.2368e-12)))), Etofenprox = 1/(1+exp(1.4305*(log(Mols)-log(1.7591e-11)))), Deltamethrin = 1/(1+exp((6.4233e-01)*(log(Mols)-log(1.0283e-13)))), DDT = 1/(1+exp((4.8912e-01)*(log(Mols)-log(6.0348e-12)))), Transfluthrin = 1/(1+exp((5.7621e-01)*(log(Mols)-log(3.0401e-12))))	0 to 0.1
		0.1 to 0.2
		0.2 to 0.3
		0.3 to 0.4
		0.4 to 0.5
		0.5 to 0.6
		0.6 to 0.7
		0.7 to 0.8
		0.8 to 0.9
		0.9 to 1
Resulting population after pesticide application	Outcomes of dose-response modeling for pesticides examined in Agramonte et al. (2017) applied to the population estimates resulting from the gene drive. States are 0 to 1 to represent extinction, 0 to 100 to represent near extinction, and then Jenk's natural breaks optimization.	0 to 1 (individuals)
		1 to 100 (individuals)
		100 to 2741 (individuals)
		2741 to 7626 (individuals)
		7626 to 13729 (individuals)
		13729 to 64030.7 (individuals)
Pupae	Equations for pupal production from standing crop of female mosquitoes are derived from Focks et al. (2000). A 50/50 sex ratio was assumed, total population/2 gives approximation of female mosquitoes.	0 to 1 (pupae)
		1 to 114 (pupae)
		114 to 3124 (pupae)
		3124 to 8693 (pupae)

Pupae (continued)	Pupae (Temperature, TotalAfterPesticide) = 22 = ((TotalAfterPesticide)/2)*1.14, 24 = ((TotalAfterPesticide)/2)*0.94, 26 = ((TotalAfterPesticide)/2)*0.75, 28 = ((TotalAfterPesticide)/2)*0.57, 30 = ((TotalAfterPesticide)/2)*0.41, 32 = ((TotalAfterPesticide)/2)*0.26	8693 to 33117.5 (pupae)
Pupae per 1000 people	Pupae per 1000 people estimates are derived from dividing the pupae estimates per the human population per risk region and multiplying by 1000.	0 to 100 (pupae/1000 people)
		100 to 200 (pupae/1000 people)
		200 to 300 (pupae/1000 people)
		300 to 2318.82 (pupae/1000 people)
IMPACTS		
Zika transmission	Transmission thresholds derived from Barrera et al. (2017). Zero = 0 to 0.2 females/trap/day, Low = 0.2 to 0.8 females/trap/day, Medium = 0.8 to 1.5 females/trap/day, High = 1.8 to >6.8 females/trap/day.	Zero
		Low
		Medium
		High
Dengue reduction	Dengue transmission thresholds derived from Focks et al. (2000). Dengue transmission states represent the likelihood of being below the transmission threshold, at the transmission threshold, or above the transmission threshold in terms of pupae/1000 people.	Below threshold
		At threshold
		Above threshold