ROSÂNGELA PEREGRINA SANCHES

**Análise do número de reprodutibilidade basal na fase inicial de
doenças causadas por vetores**

Tese apresentada à Faculdade de Medicina da
Universidade de São Paulo para a obtenção
do título de Doutor em Ciências

Programa de Patologia
Orientador: Prof. Dr. Eduardo Massad

(Versão corrigida. Resolução CoPGr 6018, de 13 de Outubro de 2011.
A versão original está disponível na Biblioteca FMUSP)

SÃO PAULO
2015

Dados Internacionais de Catalogação na Publicação (CIP)

Preparada pela Biblioteca da
Faculdade de Medicina da Universidade de São Paulo

©reprodução autorizada pelo autor

Sanches, Rosângela Peregrina

Análise do número de reprodutibilidade basal na fase inicial de doenças
causadas por vetores / Rosângela Peregrina Sanches. -- São Paulo, 2015.

Tese(doutorado)--Faculdade de Medicina da Universidade de São Paulo.
Programa de Patologia.

Orientador: Eduardo Massad.

Descritores: 1.Número básico de reprodução 2.Dengue 3.Modelos teóricos
4.Modelos matemáticos 5.Funções verossimilhança 6.Estudo comparativo

USP/FM/DBD-399/15

Dedicatória

*Dedico esta tese em primeiro lugar a Deus, e também aos meus pais
Antonio e Elza, e ao meu namorado Gilson.*

Agradecimentos

Ao professor Eduardo Massad pela orientação, por tudo que me ensinou e por toda a ajuda durante a realização desse projeto.

Aos meus pais Elza e Antonio, por sempre me apoiarem, me incentivarem a seguir meus sonhos, e me darem mesmo que de longe todo o seu carinho e amor.

Ao meu namorado Gilson pelo apoio, compreensão, carinho, paciência, amor.

Aos meus companheiros de laboratório, Franciane, Teresa, Raphael, Édina, Edson, Silvia, Hélio, por todos os momentos que compartilhamos, pelas discussões produtivas, e também por todas as risadas e momentos de descontração que tivemos.

Aos professores Marcelo Buratini, Neli Ortega e Marcos Amaku, que participaram da minha banca de qualificação, pelas sugestões oferecidas a este trabalho, e também aos professores David Greenhalgh pela ajuda com as contas e Jianhong Wu pelo estágio que me foi oferecido na York University.

Aos funcionários da DIM: Iracene, Vilma e Paulo César, por toda ajuda durante o doutorado.

Aos todos os meus amigos que mesmo à distância sempre me deram palavras de apoio e incentivo.

A Coordenação de Aperfeiçoamento Pessoal de Nível Superior (CAPES) pelo suporte financeiro.

Acima de tudo agradeço à Deus por sempre iluminar o meu caminho.

*“Foi o tempo que dedicaste à tua rosa
que a fez tão importante.”
(Antoine de Saint-Exupéry)*

Esta tese está de acordo com as seguintes normas, em vigor no momento desta publicação:

Referências: Adaptado de *International Committee of Medical Journals Editors (Vancouver)*.

Universidade de São Paulo. Faculdade de Medicina. Serviço de Biblioteca e Documentação. *Guia de apresentação de dissertações, teses e monografias da FMUSP*. Elaborado por Anneliese Carneiro da Cunha, Maria Julia de A. L. Freddi, Maria F. Crestana, Marinalva de Souza Aragão, Suely Campos Cardoso, Valéria Vilhena. 3a ed. São Paulo: Divisão de Biblioteca e Documentação; 2011.

Abreviaturas dos títulos dos periódicos de acordo com *List of Journals Indexed in Index Medicus*.

Sumário

Página

Lista de Figuras

Lista de Tabelas

Resumo

Abstract

1	INTRODUÇÃO	1
1.1	Dengue	1
1.1.1	Histórico da Dengue	8
1.1.2	A Dengue no mundo	10
1.2	Modelagem matemática em epidemiologia	13
1.3	Número de Reprodutibilidade Basal	16
1.3.1	A história do R_0 na visão de Heesterbeek.	22
1.3.2	O Número de Reprodutibilidade Basal e a Dengue	29
1.3.3	Breve introdução ao cálculo de R_0 neste projeto	31
1.4	Objetivo	32
2	MÉTODOS	33
2.1	Método para o cálculo do R_0 proposto por Ross e Macdonald.	33
2.2	Método para o cálculo do R_0 dado por Nishiura.	36
2.2.1	Método para o cálculo de R_0 dado por White e Pagano.	37
2.3	Cálculos adicionais	38
2.4	Determinação da fase inicial da doença	43
3	RESULTADOS	44
3.1	Cálculo do valor de R_0 utilizando o método de Ross-Macdonald.	44

3.2	Cálculo do valor de R_0 utilizando os métodos de Nishiura e White e Pagano.	52
3.2.1	Método de Nishiura.	52
3.2.2	Método de White e Pagano.	59
4	DISCUSSÃO	66
5	CONCLUSÃO	74
6	ANEXOS	76
	REFERÊNCIAS	83
	APÊNDICES	

Lista de Figuras

Figura		Página
1	Ciclo da Dengue. Fonte: http://www.brasil.gov.br/saude/2010/03/ciclo_da_dengue	2
2	Área de risco de dengue no mundo em 2011. Fonte: Murray et al. (2013).	11
3	Diagrama do modelo compartimental referente ao sistema (1).	19
4	Diagrama do modelo compartimental referente ao sistema (6).	21
5	Em pontos azuis os dados reais, e na linha em vermelho o ajuste dado pela equação (18) no período de 2009-2010.	35
6	Na linha em azul indivíduos infectados dados pela equação (20), na linha em vermelho o ajuste obtido utilizando o Modelo de Ross-Macdonald (Macdonald, 1952) no período de 2009-2010.	36
7	Diagrama do modelo compartimental referente ao Modelo de Ross-Macdonald (Macdonald, 1952), o sistema (40).	48
8	Figura (a) em pontos azuis os dados reais, na linha em vermelho o ajuste dado pela equação (18) no período de 2009-2010, na Figura (b) a linha azul apresenta os indivíduos infectados dados pela equação (20), na linha em vermelho o ajuste obtido utilizando o Modelo de Ross-Macdonald (Macdonald, 1952) (sistema de equações (40)) no período de 2009-2010.	51
9	Figura (a) em pontos azuis os dados reais, na linha em vermelho o ajuste dado pela equação (18) no período de 2010-2011, na Figura (b) a linha azul apresenta os indivíduos infectados dados pela equação (20), na linha em vermelho o ajuste obtido utilizando o Modelo de Ross-Macdonald (Macdonald, 1952) (sistema de equações (40)) no período de 2010-2011.	52

Lista de Tabelas

Tabela	Página
1	Valores biológicos utilizados na simulação. 50
2	Valores de R_0 para o período de 2009-2010. 65
3	Valores de R_0 para o período de 2010-2011. 65
4	Dados de dengue na cidade de Ribeirão Preto no período de 2009- 2010(CVE, 2012) 76
5	Dados de dengue na cidade de Ribeirão Preto no período de 2010- 2011(CVE, 2012) 79

Resumo

Sanches RP. *Análise do número de reprodutibilidade basal na fase inicial de doenças causadas por vetores* [Tese]. São Paulo: Faculdade de Medicina, Universidade de São Paulo; 2015.

O número de reprodutibilidade basal, R_0 , é definido como o número esperado de casos secundários de uma doença produzidos por um indivíduo infectado em uma população suscetível durante seu período de infecciosidade. Tem-se que, para $R_0 < 1$ a doença não consegue se manter na população, e para $R_0 > 1$ a doença irá se estabelecer. O cálculo do valor de R_0 pode ser feito de diversas maneiras, como por exemplo: a partir da análise de estabilidade de um modelo compartimental, através da matriz de próxima geração, da fase final de uma epidemia, entre outros. Neste trabalho foi estudado o cálculo de R_0 a partir da fase inicial de crescimento de um surto, em que ao fazer este cálculo não é suposto crescimento exponencial da doença, o que é proposto implicitamente na maior parte dos estudos. Foram estudadas as técnicas propostas por Nishiura, Ross-Macdonald e White e Pagano. O objetivo deste estudo foi comparar essas técnicas e avaliar como cada técnica calcula o valor do número de reprodutibilidade basal, aplicando-as a doenças causadas por vetores, neste caso em particular foram utilizados dados de dengue. Foram utilizados dados da cidade de Ribeir-

rão Preto nos períodos de 2009-2010 e 2010-2011, em ambos os casos a cidade apresentou um surto epidêmico. Os resultados apresentados pelos três métodos são numericamente diferentes. Pode-se concluir que todos os métodos acertam na previsão de que a dengue irá se propagar na cidade estudada, o que é verdade para os casos estudados, e que apesar de serem numericamente diferentes a análise semanal dos dados mostra que os valores calculados apresentam um mesmo padrão ao longo do tempo.

Descritores: número básico de reprodução; dengue; modelos teóricos; modelos matemáticos; funções verossimilhança; estudo comparativo.

Abstract

Sanches RP. *Analysis of the basic reproduction number from the initial growth phase of the outbreak in diseases caused by vectors* [Thesis]. São Paulo: Faculdade de Medicina, Universidade de São Paulo; 2015.

The basic reproduction number, R_0 , is defined as the expected number of secondary cases of a disease produced by a single infection in a susceptible population along its infectiousness period. If $R_0 < 1$ the disease cannot establish in the population, and if $R_0 > 1$ we expect the disease spread in the population. The value of R_0 can be estimated in several ways, for example, with the stability analysis of a compartmental model, through the matrix of next generation, using the final phase of an epidemic, etc. In this work we studied methods for estimating R_0 from the initial growth phase of the outbreak, without assuming exponential growth of cases, which is suggested in most studies. We used the methods proposed by Nishiura, Ross-Macdonald and White and Pagano. The objective of this work was to compare these techniques and to evaluate how these technique estimate the value of the basic reproduction number, applying them to diseases caused by vectors. In this particular case we used data of dengue. We used data from the city of Ribeirão Preto in the periods of 2009-2010 and 2010-2011, in both cases the city had an outbreak. The results obtained by the three methods are

numerically different. We can conclude that all methods are correct in the sense that dengue will spread in the city studied, what is true for the cases studied, although they are numerically different. Weekly analysis of the data show that the estimated values have a same pattern over time.

Descriptors: basic reproduction number; dengue; models, theoretical; mathematical models; likelihood functions; comparative study.

1 INTRODUÇÃO

Nesta seção são apresentados os tópicos estudados nesta pesquisa, são eles: dengue, modelagem matemática, número de reprodutibilidade basal, e objetivos deste trabalho.

1.1 Dengue

Ao se estudar a história da disseminação de doenças infecciosas, uma característica importante a ser observada é o movimento de massas das populações humanas, sendo um dos principais exemplos o da colonização das Américas. A população indígena que habitava as Américas não possuía qualquer imunidade às doenças trazidas por Europeus e Africanos, levando a uma rápida infestação de várias doenças, incluindo a dengue (Holmes et al., 1990).

A dengue é uma arbovirose que possui quatro sorotipos, a saber, DEN-1, DEN-2, DEN-3, e DEN-4, os quais pertencem ao gênero Flavivírus, da família Flaviviridae. É uma doença urbana em que seu vírus é mantido em um ciclo de vida que envolve humanos e o mosquito *Aedes* (Chowell e Clark, 1995). Os sorotipos a que são atribuídos uma maior mortalidade dessa doença são: DEN-2 e DEN-3 (Guzman e Istúriz, 2010).

Essa doença é transmitida ao homem através da picada da fêmea infectada do mosquito do gênero *Aedes*, mais especificamente, *Aedes aegypti* e *Aedes albopictus* (Gubler, 1998; da Cunha e Nogueira, 2005; Guzman e Istúriz, 2010). O ciclo da doença acontece da seguinte forma: a fêmea do mosquito *Aedes* alimenta-se do sangue de uma pessoa infectada durante seu período infeccioso; e após o período de incubação extrínseco, cerca de 7 a 12 dias, o mosquito passa a transmitir o vírus da dengue (Gubler, 1998; Guzman e Istúriz, 2010). O período de incubação do vírus nos humanos é em geral de 4 a 7 dias, no entanto pode variar

de 2 a 15 dias (da Cunha e Nogueira, 2005; Gubler, 1998; Guzman e Istúriz, 2010).

Na Figura (1) é mostrado o ciclo dessa doença.

Figura 1: Ciclo da Dengue. Fonte: http://www.brasil.gov.br/saude/2010/03/ciclo_da_dengue.

Pessoas infectadas apresentam um quadro clínico que pode variar desde uma infecção assintomática, até o caso mais severo onde a pessoa infectada pode sofrer hemorragia. Casos mais graves dessa doença podem causar a morte do indivíduo (Teixeira e Barreto, 2009). No entanto, os sintomas mais comuns são: febre alta, cansaço, prostração, inapetência, e dores musculares e nas articulações (Guzman e Istúriz, 2010). Os sintomas surgem após o período de incubação intrínseco (Holmes et al., 1990). Muitas vezes essa doença é confundida com outras doenças virais ou bacterianas (Teixeira e Barreto, 2009). Ao ser infectado por um sorotipo acredita-se que o indivíduo adquire imunidade duradoura a este sorotipo, e uma breve imunidade aos outros sorotipos (Simmons et al., 2012).

Os casos mais graves de dengue são conhecidos como a Dengue Hemorrágica e a Síndrome de Choque da Dengue que é a forma mais grave da doença. No caso da Dengue Hemorrágica, como o próprio nome diz, a pessoa infectada

sofre hemorragias que podem ser fatais. Na Síndrome de Choque da Dengue a pessoa infectada apresenta hipotensão, que é a queda da pressão arterial, e um choque profundo devido à perda de plasma e falha no sistema circulatório (Holmes et al., 1990). A Dengue Hemorrágica mata uma a cada cinco pessoas infectadas se estas não forem hospitalizadas e tiverem o suporte necessário (Nature , 2007).

Existem algumas hipóteses sobre o que poderia levar uma pessoa a desenvolver a Dengue Hemorrágica e a Síndrome de Choque da Dengue. A hipótese mais comum é a de que haveria um aumento da replicação viral devido a presença de anticorpos da dengue em níveis subneutralizados. Uma pessoa ao ser infectada pelo vírus da dengue produz anticorpos contra o sorotipo pelo qual foi infectada. Existe um período de tempo (geralmente de 12 semanas) em que estes anticorpos imunizam esta pessoa contra todos sorotipos da dengue, no entanto eles decaem rapidamente e se a pessoa é infectada por outro sorotipo antes dos anticorpos decaírem abaixo de um segundo nível então pode haver uma replicação viral (Holmes et al., 1990).

No entanto, nem todos os casos estão associados a uma segunda infecção, o que leva à hipótese de que algum sorotipo tenha uma probabilidade maior de causar as formas mais graves da doença. Existe também a teoria de que algumas pessoas da população tem uma maior probabilidade genética de desenvolver as formas mais graves da doença (Holmes et al., 1990).

Apesar de todas as medidas tomadas em relação à dengue, essa doença ainda é um problema de saúde pública. Segundo dados publicados na *Nature* em 2007, a dengue era, em número de casos, a segunda doença mais transmitida no mundo, em se tratando de doenças causadas por vetores, perdendo apenas para Malária (Nature , 2007; Barreto e Teixeira, 2008). O crescimento atual dessa doença é alarmante, e a preocupação tanto da sociedade quanto das autoridades de saúde cresce cada dia mais (Barreto e Teixeira, 2008). Segundo

informações da Organização Mundial de Saúde, atualmente a dengue é a doença causada por vetores que se dissemina espacialmente mais rapidamente no mundo (WHO, 2015).

Vários fatores são tidos como influenciadores da dinâmica da dengue, dentre eles são destacados:

- temperatura e chuva que influenciam o desenvolvimento da população de vetores;
- aumento da população mundial;
- viagens para locais onde existe risco de dengue;
- falta de programas de vigilância epidemiológica;
- cepa do vírus, que pode influenciar na magnitude e duração da viremia nas pessoas;
- densidade, competência e comportamento do vetor em uma população de humanos;
- suscetibilidade da população.

Acredita-se que fatores como temperatura e umidade contribuem para a sobrevivência do mosquito adulto durante o período de incubação extrínseca do vírus, e observou-se também que o aumento de temperatura faz com que o período de incubação extrínseca diminua. Os três últimos itens são apontados como fatores importantes para o início e conservação do vírus em uma população (McBride e Bielefeldt-Ohmann, 2000).

Sabe-se que a dengue é uma doença endêmica em países de clima tropical e subtropical, devido às condições climáticas favoráveis ao seu desenvolvimento (Guzman e Istúriz, 2010). Porém, essa doença também tem chegado a

regiões que não possuem estes tipos de clima, como países da América do Norte e Europa. Nestes casos a disseminação da doença se dá através de viagens e migrações (Gubler, 1998).

Uma forma de erradicar essa doença é controlar os mosquitos que a transmitem. Uma alternativa efetiva é reduzir as fontes de desenvolvimento larval, por exemplo, fazendo limpeza de quintais, retirando recipientes que acumulam água, ou seja, retirando qualquer objeto que possa servir de criadouro para o mosquito (Gubler, 1998). Já no caso de viagens para áreas tropicais, não existe um método específico, mas alguns cuidados podem ser tomados como ficar em locais com ar condicionado, espalhar inseticidas que matam mosquitos adultos, usar repelente, uso do mosquiteiro, etc. (Gubler e Clark, 1995).

Segundo o Ministério da Saúde a melhor medida a ser tomada em relação a prevenção da dengue, é a redução da densidade de mosquitos, agindo principalmente nos focos de desenvolvimento das larvas, como recipientes que acumulem água, desde uma tampinha de garrafa até uma caixa d'água destampada. O Ministério também recomenda o uso de roupas que reduzam a exposição da pele, mosquiteiros, inseticidas e repelentes.

Nenhuma vacina licenciada da dengue está disponível para a população, e uma vacina realmente efetiva e acessível seria um avanço importante para o controle dessa doença. O Instituto Butantan vem desenvolvendo uma vacina tetravalente contra a dengue desde 2010. Esta pesquisa está sendo feito em parceria com o *National Institutes of Health* (NIH), dos Estados Unidos. Essa pesquisa vem se mostrando promissora, e os experimentos realizados mostram que a resposta imune é satisfatória contra os quatro sorotipos. Uma nova fase de testes já foi aprovada, e contará com a participação de 17 mil voluntários (Toledo, 2015; FAPESP, 2015)

Foi estimado que atualmente mais de 3,6 bilhões de pessoas vivem

em áreas onde a dengue tem grande possibilidade de ser transmitida. Segundo a Organização Mundial de Saúde, em 2012, a dengue foi a doença causada por vetores mais importante no mundo, isto devido a sua disseminação global, atingindo inclusive áreas não afetadas por essa doença anteriormente (Murray et al., 2013). Segundo Bhatt et al. (2013) existem mais de 390 milhões de infecções por dengue por ano, sendo que 96 milhões apresentam algum nível de gravidade.

Um estudo publicado em 2013 relata que a região do mundo com um maior número de casos é o Sudeste da Ásia, onde cerca de 75% da população mundial exposta à dengue vive nesta região, e o número de casos de dengue severa é cerca de 18 vezes maior que nas Américas. Em 2010 foram registrados 1,6 milhões de casos de dengue nas Américas, sendo 49.000 de dengue severa, e nesta região as epidemias são cíclicas como no Sudeste da Ásia. Na América Latina esses ciclos estão se tornando cada vez mais frequentes e seu número de casos vem aumentando. A Europa tem poucos casos de dengue quando comparados as Américas e Sudeste da Ásia. Mas recentemente, houve um surto na Ilha da Madeira, em Portugal, que resultou em 2.164 casos. Pouco se sabe sobre a dengue na África, mas dados sugerem que 22 países africanos (40,74% do continente) registraram casos esporádicos ou surtos de dengue entre os anos de 1960 a 2010 (Murray et al., 2013).

Neste ano de 2015 no Brasil, segundo dados do Ministério da Saúde até a 30ª semana epidemiológica foram notificados 1.350.406 casos prováveis de dengue no país, sendo que 64,4% dos casos foram registrados na região Sudeste, totalizando 869.346 casos. Neste período foram registradas 614 mortes por dengue, o que representa um aumento de 57% quando comparado a esse mesmo período de tempo, em 2014, e a região Sudeste concentrou 71% desses casos (Ministério da Saúde, 2015).

Existem poucos estudos relacionados aos custos indiretos da den-

gue, tais como: ausência na escola e trabalho, as despesas das famílias, perda de produtividade.

No trabalho de Martelli et al. (2015), foram calculados tanto os custos diretos da dengue, que são os gastos com serviços médicos como, hospitalização, compra de medicamentos, custos ambulatoriais, quanto os custos indiretos da dengue, citados anteriormente. Esta pesquisa foi desenvolvida no Brasil, entre Setembro de 2012 e Agosto de 2013. Os estudos foram feitos nas cidades de Goiânia, Belo Horizonte, Rio de Janeiro, Teresina, Recife, e Belém, o que engloba as regiões: Centro-Oeste, Sudeste, Nordeste, e Norte. Foram calculados os custos diretos e indiretos para casos hospitalares e ambulatorias nos setores público e privado.

Neste pesquisa foram entrevistados pacientes com suspeita clínica de dengue, estes casos foram definidos de acordo com o guia de orientações para dengue no Brasil. Todas as cidades estudadas incluem crianças e adultos. Foram utilizados dados de 2.035 pacientes sintomáticos de dengue. Na entrevista os pacientes reponderam a questões relativas as características da dengue, histórico de contágio de dengue anterior ao atual, data do início dos sintomas/sinais de dengue, uso de serviços médicos, ausência no trabalho ou na escola/faculdade, e perda de produtividade (Martelli et al., 2015).

Nos custos indiretos foram calculados a perda na produtividade devido a doença e morte, em que seu valor levou em consideração o número de dias perdidos de trabalho pelos pacientes ou cuidadores durante o curso da doença. Também foram inclusos os dias perdidos na escola/faculdade pelos pacientes, em que seu valor foi calculado levando em consideração o nível de escolaridade e foi baseado no custo anual por estudante de US\$ 6.612. O número de dias perdidos no trabalho ou escola/faculdade foram obtidos através das entrevistas (Martelli et al., 2015).

Martelli et al. (2015) obtiveram que os custos diretos da dengue para: casos ambulatoriais no setor público variaram de US\$ 31 no Rio de Janeiro a US\$ 89 em Belo Horizonte; casos ambulatoriais no setor privado variaram de US\$ 77 em Recife a US\$ 168 em Goiânia; casos hospitalares no setor público variaram de US\$ 238 em Belém a US\$ 479 em Belo Horizonte; e casos hospitalares no setor privado variaram de US\$ 318 em Belém a US\$ 1.577 no Recife. Eles obtiveram que os custos indiretos da dengue para: casos ambulatoriais no setor público variaram de US\$ 28 em Belém a US\$ 102 em Goiânia, casos ambulatoriais no setor privado variaram de US\$ 107 em Belo Horizonte a US\$ 367 em Teresina; casos hospitalares no setor público variaram de US\$ 52 em Recife a US\$ 171 em Goiânia; e casos hospitalares no setor privado variaram de US\$ 56 no Rio de Janeiro a US\$ 900 em Recife.

1.1.1 Histórico da Dengue

As primeiras epidemias de dengue, relatadas na literatura médica ocorreram em Jakarta (ilha de Java), e em Alexandria e Cairo (Egito) em 1779, e na Filadélfia (EUA) em 1780 (da Cunha e Nogueira, 2005; Gubler, 1998). No entanto, existem relatos de doenças clinicamente compatíveis anteriores aos citados acima. O mais antigo encontra-se na Enciclopédia Chinesa de sintomas e tratamentos de doenças, publicada primeiramente na Dinastia Chin (265 a 420 D.C.) e editada em 610 D.C. na Dinastia Tang, que traz relatos de uma doença chamada pelos chineses de “veneno da água”, e recebeu este nome pois os chineses acreditavam que esta doença se devia a uma relação entre insetos voadores associados à água (Gubler, 1998). Existem relatos também compatíveis com a dengue no Caribe em 1635 e 1699 (Guzman e Istúriz, 2010). A dengue, ou uma infecção similar, teve uma ampla distribuição geográfica antes do século XVIII, e esta foi a primeira pandemia conhecida de dengue ou infecção similar (Gubler, 1998).

Foram registradas epidemias simultâneas em 1779 e 1780 na Ásia, África e América do Norte. No entanto, o primeiro caso registrado de dengue foi relatado em 1789 por Benjamim Rush, e recebeu o nome de “febre quebra-ossos”. Entre os anos de 1946 e 1963, não houve epidemias de dengue, e isto foi atribuído a um sucesso parcial na erradicação de seu vetor, o mosquito *Aedes aegypti*, devido a programas focados em conter a febre amarela (Guzman e Istúriz, 2010).

No passado chegou-se a acreditar que a dengue poderia ser causada pelos mais diferentes organismos como bactérias e protozoários, assim como sua transmissão já foi considerada ser através de miasmas, da via respiratória e por fim através de mosquitos. A comprovação de que a transmissão dessa doença se dá através do mosquito *Aedes* foi feita em 1906 por Bancroft, e a comprovação de que o agente causador dessa doença era um vírus aconteceu durante a Segunda Guerra Mundial, quando pesquisadores japoneses e do exército dos Estados Unidos conseguiram isolar os sorotipos DEN-1 e DEN-2 (da Cunha e Nogueira, 2005).

A primeira epidemia de Dengue Hemorrágica conhecida aconteceu entre os anos de 1953 e 1954 em Manila, nas Filipinas, quando este tipo de dengue foi a causa de muitas hospitalizações. Outras epidemias aconteceram em Bancoque, na Tailândia, em 1958, e nos anos 60 na Malásia, em Singapura e no Vietnã (Gubler, 1998; Gubler e Clark, 1995). O crescimento da economia e a urbanização do Sudeste da Ásia fizeram com que epidemias de dengue clássica e dengue hemorrágica se espalhassem por esta região durante os anos 70 (Gubler e Clark, 1995).

As mudanças tanto ecológicas como demográficas devido à Segunda Guerra Mundial contribuíram para o aumento do número de casos de dengue por facilitarem sua transmissão na região da Ásia e Pacífico, e a expansão econômica e urbanização contribuíram para a circulação dos vetores e a propagação do vírus (Guzman e Istúriz, 2010).

Existem evidências da ocorrência de epidemias de dengue no Brasil desde 1846, nas cidades de São Paulo e Rio de Janeiro, existem também surtos relacionados a dengue nos anos de 1851 e 1853, também na cidade de São Paulo. No entanto, as primeiras referências na literatura médica aconteceram em 1916, em São Paulo, e em 1923, em Niterói (Teixeira et al., 1999; CVE, 2015). Já a circulação do vírus da dengue foi comprovada no Brasil, somente em 1982, quando foram isolados os sorotipos DEN-1 e DEN-4, em Boa Vista, Roraima. Em 1986, foram notificados novos casos de dengue no Brasil, no Rio de Janeiro, causado pelo sorotipo DEN-1, o qual causou várias epidemias em diversas regiões do país (da Cunha e Nogueira, 2005; Teixeira et al., 1999).

O primeiro caso de dengue hemorrágica documentado no Brasil, causado pelo sorotipo DEN-2, aconteceu no Rio de Janeiro, e então começaram surgir formas mais graves da dengue hemorrágica, não só no Rio de Janeiro, como em outras regiões do país (da Cunha e Nogueira, 2005).

1.1.2 A Dengue no mundo

A dengue é uma doença endêmica em regiões tropicais e subtropicais do planeta, ocorrendo em mais de 100 países, e é também a arbovirose que se espalha mais rapidamente por essas regiões. Cerca de 2,5 bilhões de pessoas residem em áreas com risco de transmissão de dengue, e a Organização Mundial de Saúde estima que milhares de pessoas morrem por ano devido a essa doença, em sua maioria crianças (Guzman e Istúriz, 2010).

Nos últimos 50 anos a incidência de dengue aumentou em 30 vezes, aumentando sua expansão geográfica, e atingindo mais países (WHO, 2009). Na Figura (2) as áreas em laranja representam as regiões do mundo com risco de dengue em 2011 (Murray et al., 2013).

Até o fim dos anos 80 as regiões que apresentavam um número maior

Figura 2: Área de risco de dengue no mundo em 2011. Fonte: Murray et al. (2013).

de casos de dengue eram o Sudeste da Ásia, e o Oeste do Pacífico. Contudo, recentemente essa situação mudou, e a incidência de dengue tem aumentado nas regiões da América Central e América do Sul, que continham cerca de 70% dos casos de dengue no mundo (Teixeira e Barreto, 2009).

Os casos de dengue nas Américas tem crescido drasticamente. Segundo a Organização Pan-Americana de Saúde, o número total de casos de dengue nos anos 80 foi de 1.033.417, nos anos 90 cresceu para 2.725.405, em entre os anos de 2000-2007 houveram 4.759.007 casos de dengue. O número de casos de dengue hemorrágica também cresceu. Nos anos 80 foram notificados 13.398 casos, nos anos 90, 58.419 casos, e no período de 2000-2007 foram notificados 111.724 casos. Durante os 27 anos dessa pesquisa, 54,5% dos casos de dengue eram do Brasil, mas a Venezuela foi o país que mais notificou casos de dengue hemorrágica, 35,1% do número total de casos (Martín et al., 2010). Com o passar dos anos o Brasil se tornou o país que apresenta o maior número de casos dessa doença nas Américas, sendo responsável por cerca de 70% dos casos de toda a América (Rodriguez-Barraquer et al., 2008). Segundo a Organização Pan-Americana de Saúde, foram

notificados em 2014, nas Américas, 1.176.529 casos de dengue, sendo que 798 casos causaram a morte do indivíduo (PAHO, 2015).

Nos Estados Unidos, a dengue surge a partir de casos importados por viajantes e imigrantes e, neste caso, as infecções secundárias são muito raras (Guzman e Istúriz, 2010).

A dengue começou a ser um problema de saúde pública no Sudeste da Ásia durante a Segunda Guerra Mundial. Mas apesar da guerra ter servido como ponte para a chegada da dengue nesta região, o que deu condições para que essa doença se estabelecesse no Sudeste da Ásia foi a urbanização deste local após a guerra. Manila, a capital de Filipinas, teve sua primeira epidemia de dengue hemorrágica em 1953/1954, com um segundo surto em 1956. Bancoque, capital da Tailândia, notificou uma epidemia em 1958, no entanto, poucos casos foram identificados na Tailândia nos anos 50, e desde então os ciclos de epidemia de dengue, cada vez mais frequentes e duradouros vêm ocorrendo no Sudeste da Ásia (Ooi e Gubler, 2008).

O primeiro surto de dengue na China ocorreu em 1978, na província de Guangdong, e desde então surtos de dengue vêm ocorrendo nas províncias de Hainan, Guangxi, Fujian, e Zhejiang. A dengue tem se espalhado por várias regiões da China, e entre os anos de 1978 até 2008 foram registrados 655.324 casos e 610 mortes (Wu et al., 2010).

A história da dengue na África é pouco documentada. Entretanto, estudos sugerem que a dengue está presente na África desde 1926 quando causou uma epidemia em Durban, na África do Sul. Os quatro sorotipos de dengue circulam na África desde 1980 e o número de casos vem crescendo desde então devido a uma má vigilância epidemiológica nesta região do mundo. Isto acontece porque é suposto que a dengue não é um problema de saúde pública neste continente, e as formas graves dessa doença são relatadas raramente. A maioria dos casos

de dengue são infecções subclínicas e não são diagnosticadas, sendo geralmente tratadas como malária ou outros tipos de febres endêmicas (Sang, 2006).

Na Europa a maioria dos casos de dengue são trazidos por viajantes que voltam de áreas endêmicas. Os casos de dengue aumentaram de 64 em 1999, para 224 em 2002, e tem se mantido entre 100-170 casos anualmente (Guzman e Istúriz, 2010).

1.2 Modelagem matemática em epidemiologia

A modelagem matemática tem como objetivo descrever fenômenos observados em diversas áreas como epidemiologia, engenharia, ecologia, etc. Os modelos matemáticos surgem a partir da necessidade de reproduzir padrões observados no dia-a-dia, a fim de compreendê-los melhor para então poder interferir em seu processo. Os modelos são a tradução da linguagem do mundo real em linguagem matemática.

“O objetivo da epidemiologia matemática é integrar biologia e dados circunstanciais dentro de um todo coerente.” George MacDonald (Cummings e Lesser, 2013).

Os modelos matemáticos se tornaram ferramentas importantes na análise da disseminação e controle de doenças infecciosas, o processo de formulação de um modelo ajuda a esclarecer os processos e parâmetros envolvidos na dinâmica de uma doença (Hethcote, 2000).

Os modelos podem identificar comportamentos que não são claros em situações reais, como por exemplo, os limiares que ditam o surgimento de uma epidemia e que são encontrados na maioria dos modelos epidemiológicos, sendo portanto um de seus principais resultados (Brauer e Castillo-Chavez, 2012). De acordo com a teoria de limiares formulada por Kermack e McKendric, a introdução de um número pequeno de indivíduos infectados em uma população suscetível

provocará uma epidemia se a densidade de suscetíveis estiver acima de um certo valor crítico (Anderson e May, 1991).

Entender quais são as características de transmissão de doenças infecciosas em uma população pode ajudar na erradicação de doenças, ou pode pelo menos propiciar um decréscimo em sua transmissão.

Os modelos matemáticos são usados para testar, comparar, avaliar, planejar, implementar e otimizar a prevenção, e programas de controle de doenças. A modelagem matemática na área epidemiológica, pode contribuir para a formulação e análise de inquéritos epidemiológicos, os quais sugerem quais dados devem ser recolhidos, ajudam a identificar as tendências, fazem previsões gerais sobre o comportamento da doença, e estimam a incerteza nas previsões (Hethcote, 2000).

Ao se formular uma teoria matemática são construídos modelos analíticos que geralmente são conjuntos de equações diferenciais, e em boa parte das vezes esses modelos não são simples o bastante para sejam resolvidos sem a ajuda de simulações numéricas, as quais envolvem uso de computadores. Uma vantagem em se utilizar essa ferramenta é que podem ser testadas as mais variadas hipóteses, o que muitas vezes não pode ser feito na vida real devido às questões éticas, e também pelo tempo que seria necessário para realizar tal experimento (Massad, 1996).

Existem dois tipos de modelos matemáticos: os determinísticos, e os estocásticos. Modelos determinísticos, são aqueles em que a resposta do modelo depende unicamente dos valores dados aos parâmetros, e suas condições iniciais. Já nos modelos estocásticos, a resposta dada pelo modelo depende da aleatoriedade associada a este sistema, e não somente dos parâmetros e suas condições iniciais.

Os estudos nessa área surgiram em 1760 com Daniel Bernoulli,

quando ele usou um modelo matemático para avaliar a eficácia da imunização contra a varíola, a fim de influenciar políticas de saúde pública (Anderson e May, 1991; Brauer e Castillo-Chavez, 2012). Um passo importante foi dado em 1840 por Willian Farr, em que ele ajustou dados de mortes por varíola a uma curva normal (Anderson e May, 1991). Já P. D. En'ko deu as primeiras contribuições para a epidemiologia na matemática moderna. Entre os anos de 1873-1894, ele mostrou um bom entendimento sobre limiares de doenças infantis (Brauer e Castillo-Chavez, 2012; Heesterbeek, 2002).

Outros trabalhos importantes foram feitos por Harmer, Ross e Moshkovskii, que formularam teorias específicas sobre a transmissão de doenças utilizando modelos matemáticos simples. Hudson, Kermack, McKendrick e Soper também deram contribuições importantes, as quais servem de base para estrutura teórica em estudos sobre epidemiologia matemática, juntamente com J. Brownlee que deu contribuições estatísticas importantes (Anderson e May, 1991; Brauer e Castillo-Chavez, 2012).

Dentre essas contribuições duas podem ser destacadas: a primeira é de Harmer com a “lei de ação das massas”, a qual diz que a taxa de disseminação de uma doença é assumida ser proporcional ao produto entre a densidade de pessoas suscetíveis e a densidade de pessoas infectadas; a segunda é de Ronald Ross com seu trabalho pioneiro sobre a dinâmica de malária (Anderson e May, 1991).

Outra contribuição muito importante é de George Macdonald, em seu estudo sobre a malária de 1952, em que definiu o Número de Reprodutibilidade Basal, R_0 , um conceito muito importante em epidemiologia matemática. Segundo Heesterbeek (1992), o R_0 é a maior contribuição que a matemática trouxe para a epidemiologia.

A modelagem matemática é uma ferramenta muito importante. No entanto, ela só é realmente útil se existir uma cooperação entre epidemiologistas e

pesquisadores da área de exatas. Felizmente, existem exemplos dessa cooperação como o que aconteceu em São Paulo em uma campanha contra a rubéola, neste caso um modelo matemático desenvolvido por Massad et al. (1994), foi aliado a campanha de vacinação para o controle da rubéola (Massad, 1996).

1.3 Número de Reprodutibilidade Basal

Uma das principais preocupações em termos de saúde pública é em relação ao número de doenças que estão surgindo e também com aquelas que aparentemente haviam sido controladas, mas que voltaram a surgir como: dengue, febre amarela, malária, entre outras (Lopez et al., 2002).

Ao estudar doenças infecciosas, um dos interesses centrais é avaliar sua habilidade de contágio em uma população, ou seja, o potencial de transmissão de uma doença. Para tanto é utilizado o número de reprodutibilidade basal, R_0 , que é de extrema importância no estudo de epidemias, em especial para avaliar a eficiência de estratégias de controle (Heesterbeek, 2002).

Segundo Smith (1964) apud Nishiura (2010) *, se uma proporção da população maior do que $1-(1/R_0)$ estiver permanentemente protegida, a doença não consegue se propagar, e será erradicada ao longo do tempo.

Heesterbeek escreveu em 2002, que nos últimos 10-15 anos o R_0 havia sido um dos ingredientes mais presentes em artigos de modelagem matemática que estudavam a disseminação de doenças (Heesterbeek, 2002).

O número de reprodutibilidade basal, R_0 , é definido como o número esperado de casos secundários de uma doença produzidos em uma população completamente suscetível, causados por um indivíduo infectado durante seu período infeccioso (Diekmann et al., 1980; van den Driessche e Watmough, 2002; Heesterbeek, 1992). Para doenças transmitidas por vetores, o R_0 pode ser definido como

*Smith CEG. Factors in the transmission of virus infections from animal to man. *Scientific Basics of Medicine Annual Review*. 1964: 125-50.

o número de casos secundários, causados por um indivíduo inicialmente infectado por um vetor (Massad et al., 2010; Lopez et al., 2002).

Se $R_0 < 1$, então em média cada indivíduo infectado produz menos do que um indivíduo infectado durante seu período infeccioso, e portanto, a doença não consegue se estabelecer na população. Mas, se $R_0 > 1$, então em média cada indivíduo infectado produz mais do que um indivíduo infectado, e portanto, a doença pode vir a causar uma epidemia (Diekmann et al., 1980; van den Driessche e Watmough, 2002; Heesterbeek, 1992).

É importante saber que é possível existirem pessoas infectadas mesmo com $R_0 < 1$, neste caso tem-se que o número de infectados começa a crescer mas em pouco tempo esse número diminui até a doença desaparecer da população (Heesterbeek, 1992).

Existem inúmeras maneiras de se calcular o valor de R_0 , como por exemplo: através da análise de estabilidade de um sistema de equações diferenciais; o método da matriz de próxima geração; a partir do final de uma epidemia; utilizando a função de sobrevivência; a partir de taxa de crescimento inicial; etc (Smith?, 2008).*

A seguir são apresentados de forma simplificada dois exemplos para o cálculo de R_0 .

- **Análise de estabilidade de um sistema de equações diferenciais**

Essa análise é feita com o objetivo de averiguar a estabilidade dos pontos de equilíbrio de um sistema de equações, e estes são dados por: equilíbrio trivial, onde todas as populações estão extintas; equilíbrio livre da doença em que, como o próprio nome diz a população está livre da doença; e o equilíbrio endêmico, no qual coexistem indivíduos suscetíveis e infectados.

*O ponto de interrogação após o sobrenome Smith, está correto. Robert Smith? diz ter um nome popular e que portanto fica totalmente invisível no site de buscas Google, e além disso ele tem como homônimo, o vocalista de uma banda chamada *The Cure*. Portanto ele decidiu colocar um ponto de interrogação ao fim de seu sobrenome, para diferenciar seu nome dos outros.

Os pontos de equilíbrio são obtidos igualando cada uma das derivadas do sistema de equações diferenciais a zero. Após calcular os pontos de equilíbrio é necessário calcular a matriz Jacobiana, a qual é uma matriz com as derivadas parciais do sistema de equações, onde a primeira coluna é dada pelas derivadas parciais de todas as equações em relação a primeira variável, a segunda coluna é dada pelas derivadas parciais de todas as equações em relação a segunda variável e assim por diante (Boyce e DiPrima, 2002; Smith?, 2008).

A estabilidade do sistema de equações, é determinada através do sinal dos autovalores do polinômio característico, dados por $P(\lambda) = \det(A - \lambda I)$, onde A é a matriz Jacobiana calculada no ponto de equilíbrio, I a matriz identidade e λ os autovalores associados a matriz A . Se todos os autovalores são negativos, o equilíbrio é estável, se pelo menos um dos autovalores é positivo o equilíbrio é instável (Boyce e DiPrima, 2002).

Para exemplificar as informações dadas acima, considere o exemplo dado por Smith? (2008), onde é dado um sistema de equações diferenciais simples, especificamente um Modelo SIS (Suscetível-Infectado-Suscetível), em que uma pessoa suscetível se torna infectada através do contato com uma pessoa infectada, e depois de recuperar da doença volta a ser suscetível.

$$\begin{cases} \frac{dS(t)}{dt} = bI(t) - aS(t)I(t), \\ \frac{dI(t)}{dt} = aS(t)I(t) - bI(t), \end{cases} \quad (1)$$

em que S é a população de suscetíveis, I a população de infectados, b é a taxa de recuperação e a é a taxa de contato. Este modelo está representado na Figura 3.

A Matriz Jacobiana deste sistema é:

$$F = \begin{bmatrix} -aI(t) & b - aS(t) \\ aI(t) & aS(t) - b \end{bmatrix} \quad (2)$$

Figura 3: Diagrama do modelo compartimental referente ao sistema (1).

Considerando o equilíbrio livre da doença, tem-se que o ponto de equilíbrio é dado por $(N,0)$, porque neste caso toda a população é suscetível e ninguém é infectado. Avaliando a Matriz Jacobiana neste ponto de equilíbrio, são obtidos os seguintes autovalores:

$$\lambda_1 = 0, \lambda_2 = aN - b. \quad (3)$$

Portanto, a estabilidade do sistema depende do autovalor λ_2 , já que λ_1 é nulo.

Se $aN - b < 0$ o ponto de equilíbrio livre da doença é estável (uma vez que ambos os autovalores não são positivos), e se $aN - b > 0$ existe um autovalor positivo então este equilíbrio é instável. Desenvolvendo esta inequação tem-se,

$$\begin{aligned} aN - b &< 0, \\ aN &< b, \\ \frac{aN}{b} &< 1. \end{aligned} \quad (4)$$

Então, é possível definir,

$$R_0 = \frac{aN}{b}. \quad (5)$$

Tem-se que se $R_0 > 1$ o equilíbrio é instável e a doença será en-

dêmica, e se $R_0 < 1$ o equilíbrio é estável e portanto a doença irá se extinguir (Smith?, 2008).

- **Matriz de próxima geração**

Este método é utilizado para obter uma expressão R_0 para um modelo compartimental, em que o R_0 , é o autovalor dominante da matriz de próxima geração. De forma resumida, tem-se que x_i é o número ou proporção de indivíduos no i -ésimo compartimento do modelo, $F_i(x)$ é a taxa de aparecimento de novas infecções, e $V_i(x)$, é diferença entre a taxa de saída do compartimento i e taxa de entrada no compartimento i . A diferença $F_i(x) - V_i(x)$ dá a taxa de mudança de x_i (van den Driessche e Watmough, 2002; Smith?, 2008).

A multiplicação de FV^{-1} determina a taxa pela qual indivíduos infectados produzem novos indivíduos infectados (van den Driessche e Watmough, 2002; Smith?, 2008).

Assumindo F_i e V_i satisfazem as condições dadas por van den Driessche e J Watmough (2002) é possível obter a matriz de próxima geração dada por FV^{-1} , em que F é matriz das derivadas parciais de F_i , e V^{-1} é a matriz inversa das derivadas parciais de V_i .

Para exemplificar as definições dadas, considere o sistema para doenças causadas por vetores, dado por van den Driessche e J Watmough (2002), em que humanos suscetíveis se tornam infectados através do contato com mosquitos infectados e mosquitos suscetíveis se tornam infectados através do contato com humanos infectados. Este é um Modelo SIS para humanos e SI para vetores,

$$\begin{cases} \frac{dI(t)}{dt} = \beta_s S(t)V(t) - (b + \gamma)I(t), \\ \frac{dV(t)}{dt} = \beta_m M(t)I(t) - cV(t), \\ \frac{dS(t)}{dt} = b - bS(t) + \gamma I(t) - \beta_s S(t)V(t), \\ \frac{dM(t)}{dt} = c - cM(t) - \beta_m M(t)I(t), \end{cases} \quad (6)$$

em que I é a população de humanos infectados, V é a população de vetores infectados, S é a população de humanos suscetíveis e M é a população de vetores suscetíveis, β_s é a taxa pela qual humanos se tornam infectados, β_m é taxa pela qual vetores se tornam infectados, b é a taxa de mortalidade dos humanos, c é taxa de mortalidade dos vetores, e γ é a taxa de recuperação dos humanos, neste modelo humanos voltam a ser suscetíveis. Sua população total é constante, em que $S + I = 1$ e $M + V = 1$. Este modelo está representado na Figura 4.

Figura 4: Diagrama do modelo compartimental referente ao sistema (6).

Portanto,

$$F = \begin{bmatrix} 0 & \beta_s \\ \beta_m & 0 \end{bmatrix}, \quad V = \begin{bmatrix} b + \gamma & 0 \\ 0 & c \end{bmatrix}, \quad (7)$$

a inversa de V é dada por,

$$V^{-1} = \begin{bmatrix} \frac{1}{b+\gamma} & 0 \\ 0 & \frac{1}{c} \end{bmatrix}, \quad (8)$$

e FV^{-1} ,

$$FV^{-1} = \begin{bmatrix} 0 & \frac{\beta_s}{c} \\ \frac{\beta_m}{b+\gamma} & 0 \end{bmatrix}. \quad (9)$$

Obtendo o autovalor dominante,

$$R_0 = \sqrt{\frac{\beta_s \beta_m}{c(b + \gamma)}} \quad (10)$$

1.3.1 A história do R_0 na visão de Heesterbeek.

Nesta subseção é dado um resumo de um estudo muito interessante onde a história do R_0 é contada. Heesterbeek (2002), escreveu um artigo, em que dá um desenvolvimento gradual da história do R_0 em demografia, ecologia e epidemiologia, seu resumo é dado a seguir.

Em demografia e ecologia, o R_0 tem uma interpretação similar, em que é dado como o número esperado de descendentes fêmeas, que atingiram a idade fértil durante sua vida. As primeiras definições para o conceito de R_0 , em demografia, surgiram em 1925 com Dublin e Lotka. Já o conceito que é utilizado hoje, em epidemiologia foi desenvolvido somente em 1980.

O conceito de limiar teve sua origem em demografia e ecologia, como interpretação do número de descendentes por geração, e em epidemiologia, em termos de densidade crítica de uma população. Mesmo tendo basicamente a mesma definição, levou algum tempo até que os pesquisadores em epidemiologia percebessem que o conceito de potencial de reprodutibilidade é poderoso e claro para doenças infecciosas. Uma possível razão para isto está ligada aos dados, enquanto que em demografia e ecologia o desenvolvimento do conceito de R_0 estava ligado ao conhecimento prático, em epidemiologia o interesse dos pesquisadores estava voltado em apresentar uma teoria matematicamente coerente.

Segundo Heesterbeek (2002), as primeiras ideias sobre R_0 em demografia, surgiram em 1886 com Richard Bösch, que calculou o total de nascimentos para a cidade de Berlin em 1879, o que chamou de “reprodução total de uma população”, no entanto sua ideia não foi a base para o desenvolvimento do R_0 em

demografia. Kuczynski, um estudante de Borch, desenvolveu seus cálculos para Berlin nos anos de 1891-1895 e desenvolveu uma teoria sobre esse conceito em 1932, entretanto Kuczynski fez este trabalho conhecendo o trabalho de Lotka e Dublin, e não em resposta ao trabalho inicial de Borch (Kuczynski, (1932) apud Heesterbeek (2002))*.

Lotka desenvolveu ao longo do tempo uma linha de raciocínio sobre o conceito de R_0 que, em seu artigo com F.R. Sharpe em 1911, foi escrita pela primeira vez a relação (Lotka e Sharpe, (1911) apud Heesterbeek (2002))**,

$$r > 0 \Leftrightarrow \int_0^{\infty} b(a)p(a)da > 1, \quad (11)$$

em que r é a taxa de reprodução por indivíduo, $b(a)$ é a taxa de concepções por indivíduo com idade a , e p é a probabilidade de sobrevivência até a idade a .

Ronald Ross em seu trabalho sobre malária, sugeriu que haveria um limiar chamado de “densidade crítica de mosquitos”, de modo que, para conter a malária, não seria preciso matar todos os mosquitos, e sim diminuir a razão entre mosquitos e homens abaixo de um limiar. Ross publicou uma série de artigos sobre a transmissão da malária antes de construir uma teoria geral de fenômenos epidemiológicos. Em uma série de três artigos (Ross (1916); Ross e Hudson (1917a,b)) ele desenvolveu uma aproximação a qual ele se refere como “a priori pathometry”, sendo que dois desses artigos foram desenvolvidos com Hilda Hudson. Eles desenvolveram uma equação geral, que descreve a dinâmica populacional de um agente infeccioso. No entanto, eles não chegaram a desenvolver uma teoria sobre o R_0 (Heesterbeek, 2002).

Existem duas principais razões para que o desenvolvimento da teoria do R_0 , no campo da epidemiologia tenha levado tanto tempo para se desenvolver. A primeira, foi ter como ponto inicial a densidade crítica ao invés da ideia limiar

*Kuczynski RR. *Fertility and Reproduction*. New York: Falcon Press; 1932.

**Sharpe FR, Lotka AJ. A problem in age distribution. *Philosophical Magazine*.1911; 21:435-8.

de reprodutibilidade, e a segunda é o fato de que Ross desenvolveu sua ideia especificamente para malária, que é uma doença que em sua dinâmica envolve humanos e mosquitos.

Lotka estudou os artigos de Ross e Hudson, e publicou um artigo em resposta a estes intitulado “*Contribution to quantitative epidemiology*” (Lotka A J, apud Heesterbeek (2002) *), onde ele faz algumas pequenas modificações na equação publicada por Ross e Hudson, e chega em uma relação similar em termos gerais a da equação (11), percebendo que crescimento da população e disseminação da doença são termos bem próximos do ponto de vista matemático. Esta poderia ter sido a origem do R_0 em epidemiologia. No entanto, isto não aconteceu porque Lotka não transferiu a interpretação demográfica para a epidemiológica.

Lotka ainda continua seus estudos em doenças infecciosas, e especificamente em malária, faz estudos sobre os limiares para a persistência da doença, mas ainda não trabalha na teoria de R_0 . Em seus estudos sobre demografia, em um primeiro momento Lotka vê suas equações apenas como uma prova matemática, somente depois de algum tempo ele dá uma interpretação a elas.

Em seu livro *Elements of Physical Biology* de 1925 (Lotka, 1925 apud Heesterbeek (2002)**), ele escreve sobre os trabalhos de Ross sobre malária, e o de Willian Robin Thompson sobre as relações entre parasitas e hospedeiros. Em seu trabalho Thompson define o “poder reprodutivo” do parasita, que é definido como o número de ovos depositados por fêmea (Thompson W.R (1923) apud Heesterbeek (2002)***). Lotka percebe que a taxa de crescimento por geração é próxima da taxa total de nascimentos em sucessivas gerações, o que ele denota pela primeira vez por uma letra,

*Lotka AJ. Contribution to quantitative epidemiology. *Journal of the Washington Academy of Science*. 1919; 9: 73-7.

**Lotka AJ. *Elements of Physical Biology*. Baltimore: Williams and Wilkins Company; 1925.

*** Thompson WR. La théorie mathématique de l'action des parasites entomophages. *Revue Générale des Sciences Pures et Appliquées*. 1923; 34: 202-10.

$$R = \int_0^{\infty} b(a)p(a)da, \quad (12)$$

infelizmente, Lotka não dá muita atenção a isto em seu livro. Em seu artigo com Dublin “*On the true rate of natural increase*” (Dublin e Lotka, 1925), eles definem cálculos para uma tabela de fertilidade, em que calculam a taxa total de nascimentos de filhas em duas gerações sucessivas, e denotam isto com o símbolo R_0 . Mas ainda não dão um nome a este símbolo, mais tarde Lotka o chama de “fertilidade líquida” e depois de “taxa de reprodutibilidade líquida”. Inicialmente ele não envolveu o termo reprodução, mas já pensava em taxa de gerações sucessivas.

Em demografia diversos autores levaram a frente o estudo de dinâmica populacional, e deram uma base matemática mais rigorosa que a de Lotka, encerrando basicamente em 1925 o desenvolvimento do R_0 em demografia. Já na epidemiologia, foram necessários mais 50 anos para que se chegasse ao mesmo nível de desenvolvimento.

McKendrick e Ross se conheceram no exército britânico, Ross incentivou McKendrick a aprender matemática aplicada para resolver problemas médicos. Após algum tempo McKendrick relatou seus progressos a Ross, e este lhe falou sobre seu interesse em estabelecer uma teoria geral sobre a lei que rege as epidemias. Em 1911 ele escreveu uma carta a McKendrick em que diz: “*Nós acabaremos por estabelecer uma nova ciência. Mas primeiro você e eu iremos destrancar a porta e então qualquer um poderá entrar*”.

Em 1927, McKendrick e Kermack publicaram seu clássico artigo “*A contribution to the mathematical theory of epidemics*” (Kermack e McKendrick, 1927), no qual deram seguimento as ideias do próprio McKendrick, e também de Ross e Hudson. Eles provaram que existe um limiar para que um agente infeccioso seja mantido em uma população, generalizando assim o conceito de densidade crítica de Ross. Somente 25 anos depois George Macdonald publicou o artigo

“*The analysis of equilibrium in malaria*” (Macdonald, 1952), em que finalmente o R_0 aparece formalmente definido e calculado, e no apêndice de seu artigo ele define:

“*Taxa de reprodutibilidade basal da malária, definição: número de infecções distribuídas em uma comunidade como um resultado direto da presença de um simples caso não imune*”.

Ele deu uma expressão para a “taxa” de reprodutibilidade basal para malária baseada nos termos já identificados por Ross, e então explicou que reescrevendo a expressão resultante era possível obter o nível crítico obtido por Ross. Macdonald, foi a primeira pessoa a fazer isto em epidemiologia, e ele apresentou seus estudos a um assessor da Organização Mundial de Saúde (OMS) para que estes cálculos pudessem ser usados em campanhas antimalária. Em 1955, ele começou a utilizar o símbolo z_0 para identificar a “taxa” de reprodutibilidade basal.

Embora fosse esperado que a partir de então o conceito dado por Macdonald ganhasse força no campo da epidemiologia matemática, e que um número crescente de matemáticos se interessassem sobre isso, o que realmente aconteceu naquela época não foi exatamente isso.

Norman Bailey publicou em 1957, o primeiro livro inteiramente voltado ao estudo da epidemiologia matemática (Bailey N.J.T. (1957) apud Heesterbeek (2002)*). No entanto o conceito de R_0 foi deixado de lado. Seu livro deu aos matemáticos uma boa introdução à epidemiologia, e também mostrou a riqueza dos problemas matemáticos nesta área.

Bailey tinha conhecimento do trabalho de Macdonald e chegou a fazer algumas citações em seu livro, mas aparentemente ele não deu o devido crédito a “taxa de reprodutibilidade basal da malária”. Infelizmente, naquela época matemáticos dificilmente leriam um artigo publicado na revista *Tropical Diseases*

*Bailey NJT. *Mathematical Theory of Epidemics*. London:Griffin; 1957.

Bulletin, a menos que soubessem que existia uma ideia matemática interessante em um de seus artigos. Assim, o artigo de Macdonald ficou esquecido por algum tempo.

Em 1974, houve uma conferência em Utah que reuniu vários jovens pesquisadores incluindo Hethcote e Dietz que re-introduziram a ideia. Hethcote apud Heesterbeek (2002)* introduziu o conceito de “número de contatos infecciosos”.

Dietz (1975) definiu:

“A quantidade R é chamada de taxa de reprodutibilidade, representa o número de casos secundários que um caso inicialmente infectado produz se introduzido em uma população suscetível”.

Heesterbeek (2002) afirma que Dietz conhecia o trabalho de Macdonald, e escreve que Dietz usou o termo “taxa de reprodutibilidade basal” e o símbolo z_0 dados por Macdonald. Dentre os artigos que estudou, Dietz descobriu um estudo sobre arboviroses em que é utilizado o símbolo RR . Este estudo foi feito por Smith em 1964 **, ele foi a primeira pessoa a mostrar a relação entre o R_0 , e a fração de pessoas a ser vacinada. Dietz cita, e dá créditos a ambos estudos (Smith, 1964 apud (Heesterbeek, 2002)***).

E então finalmente foi obtida uma definição de R_0 , como um conceito de reprodutibilidade em epidemiologia, que pode ser usado para o estudo de controle de doenças infecciosas, e os modelos epidemiológicos começaram a crescer.

May e Anderson contribuíram muito para que o R_0 se tornasse mais conhecido, e defenderam com sucesso seu uso, mas mesmo eles levaram um certo

*Hethcote H.W. Mathematical models for the spread of infectious diseases. In: Ludwig D, Cooke KL [editor]. Epidemiology. *SIAM, Philadelphia*. 1975; 122-31.

**Esta citação de Smith não contém o ponto de interrogação por se tratar de uma pessoa diferente.

*** Smith CEG. Factors in the transmission of virus infections from animal to man. *Scientific Basics of Medicine Annual Review*. 1964: 125-50.

tempo para reconhecer seu potencial. Eles publicaram dois artigos na *Nature* sobre doenças infecciosas que revitalizaram as ideias sobre modelagem matemática em epidemiologia, chamados: “*Population biology of infectious diseases: Part I*” (Anderson e May, 1979a); e “*Population biology of infectious diseases: Part II*” (Anderson e May, 1979b), no entanto não escreveram nada sobre R_0 . Eles até chegaram a citar os artigos de Dietz e Macdonald, mas em outro contexto. Felizmente eles perceberam a importância do R_0 e, pouco tempo depois publicaram um artigo na *Science*, chamado “*Directly transmitted infectious: Control by vaccination*” (Anderson e May, 1982), em que compensaram sua falha e falaram largamente sobre o R_0 .

Eles também levaram algum tempo para fazer a ligação entre o conceito sobre R_0 em demografia e epidemiologia. Em um primeiro momento, eles os utilizam de forma separada, como se fossem assuntos diferentes, mas uma vez que perceberam que se tratava da mesma definição, eles não perderam tempo e apresentaram o R_0 em uma famosa conferência em Dahlem, essa conferência foi muito influente em revigorar o interesse científico no campo da modelagem matemática, como uma ferramenta para o estudo do controle e da disseminação de doenças infecciosas. No trabalho apresentado nesta conferência Anderson juntou o símbolo dado por Lotka, R_0 , ao nome dado por Macdonald “taxa de reprodutibilidade basal”, e finalmente este conceito teve a atenção merecida (Anderson e May (1982) apud Heesterbeek (2002)) *.

Uma observação curiosa e verdadeira feita por Heesterbeek (2002) é que quem lê toda a contribuição de Anderson e May, e o forte uso do R_0 é levado a crer que este conceito já havia sido bem estabelecido, estudado e usado em epidemiologia por décadas, no entanto isto está bem longe da realidade.

O termo correto para o R_0 é número de reprodutibilidade basal e

*Anderson RM, May RM. Population Biology of Infectious Diseases. Berlin:Springer-Verlag; 1982.

não taxa de reprodutibilidade basal, pois como é mostrado por Massad e Coutinho (2012), o R_0 é um número pois não tem dimensão.

1.3.2 O Número de Reprodutibilidade Basal e a Dengue

Nishiura (2006) faz uma boa revisão dos artigos publicados sobre o R_0 da dengue até o ano de 2006, onde ele afirma que até então, as técnicas mais utilizadas para seu cálculo eram: através da fase inicial da doença, através do tamanho final da epidemia, e também através de um estudo feito a partir do momento em que as pessoas adquiriram a doença.

Segundo Nishiura (2006), o primeiro estudo sobre o R_0 da dengue, utilizando dados de uma epidemia, foi feito por Koopman et al. em 1991 apud Nishiura (2006)*, em que o número de reprodutibilidade basal foi calculado a partir do tamanho final da epidemia no México,

$$R_0 = \frac{-\ln(1-p)}{p}, \quad (13)$$

em que p é a proporção final de pessoas que foram infectadas durante a epidemia (Koopman et al.(1991) apud Nishiura (2006)*). Neste estudo foi obtido que $R_0 = 1,3$, e conclui-se que o potencial de transmissão da dengue é baixo quando comparada a outras doenças (Nishiura, 2006).

Massad et al. (2001) calcularam o valor de R_0 considerando apenas a fase inicial de crescimento da dengue,

$$R_0 = \left(1 + \frac{r}{\gamma}\right) \left(1 + \frac{r}{\mu_v}\right), \quad (14)$$

em que r é a taxa de crescimento exponencial, μ_v a mortalidade do vetor, e γ a taxa

*Koopman JS, Prevots DR, Vaca Marin MA, Gomez DH, Zarate AML, Longini IM Jr e Sepulveda Amor J. Determinants and predictors of dengue infection in Mexico. *Am J Epidemiol.* 1991; 133: 1168-78.

de recuperação. Neste estudo, eles calculam o R_0 da dengue e o da febre amarela, e o R_0 da dengue variou entre 3,58-12,86, este cálculo foi feito para algumas cidades do Estado de São Paulo em agosto de 1990 (Massad et al., 2001). Mais tarde este modelo foi aprimorado por Favier et al. (2006), em que ele considerou os períodos de incubação extrínseco e intrínseco da dengue,

$$R_0 = \left(1 + \frac{r}{\gamma}\right) \left(1 + \frac{r}{\mu_v}\right) e^{r(\tau_e + \tau_i)}, \quad (15)$$

em que τ_e e τ_i são os períodos de incubação extrínseco e intrínseco da dengue respectivamente. Em seu estudo ele discute que os métodos de Marques et al. (1994) e Massad et al. (2001), falharam justamente por não considerarem os períodos de incubação. Favier et al. (2006), fez o cálculo do R_0 para algumas cidades brasileiras em que obteve seus resultados variando entre 2-5,6, com exceção de uma cidade em que foram obtidos valores extremamente altos. Esses três modelos assumem que a fase inicial da doença tem crescimento exponencial, o que não é assumido no estudo feito por Massad et al. (2010), em que ele utiliza o R_0 proposto por Ross-Macdonald (Macdonald, 1952) e mostra que existem casos que mesmo com $R_0 < 1$ é possível que exista um surto autolimitado da dengue.

Nishiura (2006) apresenta estudos de uma técnica baseada na proporção daqueles que foram expostos pelo menos uma vez à doença para calcular a força de infecção da doença, e o primeiro trabalho neste estilo foi feito para febre amarela por Muench em 1959. A seguinte expressão para o R_0 é apresentada por Dietz (1975),

$$R_0 = 1 + \frac{L}{A_{av}}, \quad (16)$$

em que L é a expectativa de vida, e A_{av} a média de idade no momento da infecção a qual é calculada a partir da força de infecção, esta expressão foi proposta para o cálculo de R_0 para doenças transmitidas por vetores.

Ferguson (1999) apud Nishiura (2006)* estudou este método e obteve um valor de R_0 entre 4-6 para a dengue, e neste caso ele fez algumas modificações neste modelo onde assumiu um estado não estacionário e calculou a força de infecção e potencial epidêmico ao mesmo tempo.

Alguns estudos calculam o R_0 a partir da matriz de próxima geração, sendo o exemplo mais recente deste caso o estudo feito por Yang (2014), em que é estudado não só este método como ele também calcula o R_0 através da análise de estabilidade de um modelo compartimental em que obtém um R_0 próximo ao de Ross-Macdonald (Macdonald, 1952). Para os cálculos utilizando a matriz de próxima geração ele obtém que o R'_0 é raiz quarta do R_0 obtido no cálculo da análise de estabilidade, onde $R'_0 \neq R_0$. Yang (2014) faz uma discussão teórica em torno dos aspectos matemáticos acerca dos modelos de R_0 calculados, e não calcula o valor do R_0 para nenhuma região.

1.3.3 Breve introdução ao cálculo de R_0 neste projeto

Neste trabalho foram estudados os métodos para o cálculo de R_0 , pela fase inicial de crescimento da dengue, sem supor crescimento exponencial. Foram estudados os métodos de: Ross-Macdonald (Macdonald, 1952), Nishiura (2010), e White e Pagano (2008).

O método proposto por Ross-Macdonald (Macdonald, 1952), foi estudado por Massad et al. (2010), onde não foi suposto crescimento exponencial dos casos, e bons resultados foram obtidos. Eles mostraram que é possível existir um surto epidêmico mesmo quando $R_0 < 1$, porém neste caso um surto autolimitado pode ser observado, o que já havia sido sugerido antes por outros autores. Neste estudo, foi calculado o valor de R_0 para a dengue nas cidades de São Paulo e Londrina, e para este cálculo foram utilizados dados de literatura e dados reais.

*Ferguson NM, Donnelly CA e Anderson M. Transmission dynamics and epidemiology of dengue: insights from age-stratified sero-prevalence surveys. *The Royal Society*. 1999; 354: 757-68.

Nishiura (2010), propôs um método para o cálculo do valor de R_0 , onde também não é assumido crescimento exponencial e propõe uma correção ao conceito de R_a (número de reprodutibilidade atual). Através dessa correção, os números de reprodutibilidade basal e atual são iguais, e propõe-se um método baseado em verossimilhança. Este método foi aplicado para calcular o valor de R_0 para AIDS em países europeus. Em seu cálculo Nishiura (2010) utiliza dados de incidência e informações sobre o valor do intervalo de geração da doença.

White e Pagano (2008), propõem um método relacionado a um processo de ramificação, em que eles desenvolvem duas expressões para o cálculo de R_0 em que também utilizam verossimilhança, nas quais são utilizadas informações sobre dados de incidência e intervalo de geração. Eles aplicam suas técnicas em situações hipotéticas e reais, no caso das situações reais são estudadas a Gripe Suína, Influenza e Ebola.

Este trabalho é um estudo teórico sobre o cálculo do R_0 . Uma análise comparativa de métodos para o seu cálculo, pela fase inicial da doença, em que foram utilizados dados reais de dengue.

1.4 Objetivo

O objetivo deste trabalho é comparar diferentes técnicas para cálculo do número de reprodutibilidade basal que utilizam a fase inicial de crescimento da doença, e não supõem crescimento exponencial, aplicando a dados de dengue (CVE, 2012).

Utilizando as seguintes ferramentas: Cálculo Diferencial e Integral e Simulação numérica a fim de comparar os resultados obtidos.

2 MÉTODOS

Nesta seção são apresentados os métodos estudados por Ross-Macdonald (Macdonald, 1952), Nishiura (2010), e White e Pagano (2008), para calcular o número de reprodutibilidade basal, sem assumir crescimento exponencial na fase inicial de crescimento da epidemia. Também são apresentadas as equações utilizadas para realizar a comparação dos métodos, e alguns cálculos adicionais decorrentes das necessidades exigidas para simulação computacional dos métodos.

2.1 Método para o cálculo do R_0 proposto por Ross e Macdonald.

A primeira expressão para o cálculo do R_0 apresentada nesta seção é a de Macdonald (1952), que apareceu pela primeira vez em um artigo escrito por George Macdonald em 1952, onde ele propõe um limiar para existência da Malária baseado nos estudos de Sir Ronald Ross (Ross, 1911). O conceito dado por Ross-Macdonald (Macdonald, 1952) foi estudado ao longo do tempo por vários pesquisadores como Anderson, May, Aron (Anderson e May, 1991; Aron e May, 1982), e alguns erros cometidos em relação a dimensão foram corrigidos (Massad e Coutinho, 2012), sendo a expressão utilizada neste estudo dada por:

$$R_0 = \frac{ma^2bce^{-\mu_V\tau}}{(\gamma + \mu_H)\mu_V}, \quad (17)$$

em que m é proporção entre mosquitos e humanos, a é a taxa de picadas, b é probabilidade de humano se tornar infectado, c é a probabilidade de um mosquito se tornar infectado, μ_H é taxa de mortalidade dos humanos, μ_V a taxa de mortalidade dos mosquitos, γ é taxa de recuperação dos humanos e τ é o período de incubação extrínseco no mosquito.

Para calcular o valor de R_0 , foram necessárias informações acerca de alguns parâmetros como taxa de mortalidade dos humanos e mosquitos, taxa de recuperação da dengue, período extrínseco de incubação. Entretanto, alguns parâmetros são difíceis de serem calculados, como a taxa de picadas pelo mosquito (a), a probabilidade de que um humano tornar-se infectado (b), a probabilidade de que um vetor se torne infectado (c) e a proporção entre mosquitos e humanos (m). A fim de determinar os valores desses parâmetros, o modelo de Ross-Macdonald (Macdonald, 1952) (este modelo é mostrado no próximo capítulo), foi ajustado a dados reais de dengue, os dados utilizados nesta pesquisa são dados na seção de Anexos.

Este ajuste foi feito em dois passos. No primeiro passo o número de novos casos por semana de dengue (dados reais) foi ajustado a seguinte função (Massad, 1987):

$$g(t) = a_1 \text{Sech}^2(a_2 t + a_3) \frac{1}{1 + e^{a_4(t-a_5)}}, \quad (18)$$

onde os parâmetros a_i com $i = 1, \dots, 5$ foram calculados utilizando a ferramenta *Curve Fit* do software Berkeley Madonna. Esta ferramenta calcula uma combinação destes parâmetros que reproduz o padrão dos dados reais, como pode ser visto na Figura 5.

Os dados reais dão informações apenas sobre o número de infectados em cada semana, ou seja, não mostram o número de pessoas que morrem (morte natural) ou se recuperam da doença. Portanto, no segundo passo, é necessário saber que o número de casos por semana é igual ao produto da força de infecção, $\lambda(t)$, pelo número de indivíduos de suscetíveis em cada instante de tempo, $S_H(t)$, tal que $g(t) = \lambda(t)S_H(t)$.

Então foi possível calcular a prevalência de dengue em cada instante

Figura 5: Em pontos azuis os dados reais, e na linha em vermelho o ajuste dado pela equação (18) no período de 2009-2010.

de tempo, $I_H(t)$, utilizando a mesma ideia do Modelo Ross-Macdonald (Macdonald, 1952),

$$\frac{dI_H(t)}{dt} = \lambda(t)S(t) - (\mu_H + \gamma)I_H(t). \quad (19)$$

Integrando a equação (19) tem-se,

$$I_H(t) = e^{-(\mu_H + \gamma)t} [I_H(0) + \int_0^t \lambda(s')S_H(s')e^{(\mu_H + \gamma)s'} ds']. \quad (20)$$

A dinâmica da equação (19) é dada pela entrada de novos indivíduos infectados dados por $\lambda(t)S_H(t)$, tendo como saída, os indivíduos que morrem devido a taxa de mortalidade natural, μ_H , e os indivíduos que se recuperam devido a taxa de recuperação, γ . Ao integrar (19), são obtidos os valores de indivíduos infectados no tempo t , dados pela expressão (20).

O Modelo de Ross-Macdonald (Macdonald, 1952) foi utilizado, para calcular o valor dos parâmetros a , b , c e m utilizando a ferramenta *Curve Fit* do software Berkeley Madonna, como pode ser visto na Figura 6. A fim de recuperar a evolução temporal, prevalência e incidência da dengue, ajustando a curva de infectados do Modelo de Ross-Macdonald (Macdonald, 1952) a expressão (20).

Figura 6: Na linha em azul indivíduos infectados dados pela equação (20), na linha em vermelho o ajuste obtido utilizando o Modelo de Ross-Macdonald (Macdonald, 1952) no período de 2009-2010.

2.2 Método para o cálculo do R_0 dado por Nishiura.

O segundo método apresentado neste estudo, foi proposto por Nishiura (2010), em que ele propõe uma técnica para calcular o valor do R_0 . Para isso, ele fez uma correção a definição de R_a , número de reprodutibilidade atual, o qual é definido como o produto da duração média de contágio, e a taxa de incidência dividida pela prevalência da doença (White et al., 2006; Nishiura, 2010), mostrando que $R_a = R_0$, e a partir desta correção ele desenvolve um método baseado em verossimilhança para calcular o valor do R_0 , dado por,

$$L(R_0) = \prod_{t=1}^T \left(\frac{j_t}{\sum_{s=0}^{\infty} g_s j_{t-s}} \right) \left(\frac{1}{R_0} \right)^{\sum_{s=0}^{\infty} g_s j_{t-s}} \left(1 - \frac{1}{R_0} \right)^{j_t - \sum_{s=0}^{\infty} g_s j_{t-s}}, \quad (21)$$

em que T é o ponto de tempo máximo de observação na fase inicial de crescimento da epidemia, ou seja, o fim da fase inicial da doença, j_t é a incidência da doença e g_s seu intervalo de geração. R_0 é obtido minimizando o negativo do logaritmo de (21).

O intervalo de geração é o tempo entre sucessivas infecções em uma

cadeia de transmissões, e ele depende da relação temporal entre a infecciosidade e início dos sintomas clínicos de um caso transmissor e o período de incubação do caso receptor (Fine, 2003).

Para calcular seu valor são necessárias informações sobre localização e início dos sintomas do indivíduo infectado, e uma das técnicas para seu cálculo consiste no rastreamento dos contatos. Esse procedimento é feito através de uma entrevista com pacientes doentes com o objetivo de identificar quem são seus contatos próximos. A definição de contatos próximos depende do tipo de doença a ser estudada. Por exemplo, no casos de doenças sexualmente transmissíveis, os contatos devem incluir os parceiros sexuais (Aldstadt, 2007).

2.2.1 Método para o cálculo de R_0 dado por White e Pagano.

Ao estudar o método dado por Nishiura (2010) foi encontrado um outro estudo que segue um caminho diferente e obtém uma expressão para o cálculo de R_0 similar à proposta de Nishiura (2010), e portanto este estudo foi incorporado a esta tese. Em seu artigo, White e Pagano (2008) estudaram duas expressões de R_0 , uma para calcular seu valor quando o intervalo de geração é conhecido, e outra expressão para quando ele não é conhecido. Neste segundo caso, eles utilizam verossimilhança e calculam ao mesmo tempo o valor de R_0 e os parâmetros referentes ao intervalo de geração. Suas expressões são dadas por:

Para intervalo de geração conhecido,

$$\widehat{R}_0 = \frac{\sum_{t=1}^T N_t}{\sum_{t=1}^T \sum_{s=1}^{\min(k,t)} p_j N_{t-j}}. \quad (22)$$

Para calcular R_0 e intervalo de geração em conjunto,

$$L(R_0, p) = \prod_{t=1}^T \frac{e^{-\mu_t} \mu_t^{N_t}}{N_t!}, \quad (23)$$

onde N_t são os dados de incidência e p_j o intervalo de geração e $\mu_t = R_0 \sum_{j=1}^{\min(k,t)} N_{t-j} p_j$.

2.3 Cálculos adicionais

Tanto para o método proposto por Nishiura (2010), quanto para o método proposto por White e Pagano (2008), foram necessários alguns cálculos adicionais para realizar a implementação computacional dos mesmos.

Em seu método, Nishiura (2010) calculou somente o valor do R_0 na equação (21), ou seja, manteve o intervalo de geração fixado. No trabalho desenvolvido nesta tese, foram feitos dois tipos de cálculos utilizando a expressão (21): cálculo do valor de R_0 em conjunto com o valor do intervalo de geração; e somente o cálculo do valor de R_0 com o valor de intervalo de geração fixado.

Ao realizar a implementação computacional do negativo do logaritmo da equação (21) para os casos em que $j_t > 100$ foi necessário calcular valores como o fatorial de 150 por exemplo, valor este que tende a infinito. Para contornar essa situação, foi necessária utilização da distribuição normal. Tem-se que para j_t grande e um valor de $\frac{1}{R_0}$ que não seja muito próximo de 0 ou 1, a distribuição Binomial se aproxima de uma distribuição normal, e quanto maior o valor de j_t melhor essa aproximação (Spiegel, 1976). Portanto o código dessa função foi dividido em duas partes para $j_t < 100$ a verossimilhança foi calculada utilizando o negativo do logaritmo da função (21), e para $j_t \geq 100$ foi utilizado o negativo do logaritmo da distribuição normal com média igual a $j_t(\frac{1}{R_0})$ e variância igual a $j_t(\frac{1}{R_0})(1 - \frac{1}{R_0})$.

Nishiura (2010) fez seus cálculos para a fase linear de seus dados e, devido a oscilação dos pontos do conjunto de dados da dengue o mesmo necessitou ser suavizado. Para isso foi calculada a média móvel dos casos da seguinte maneira: para 2009-2010 em cada conjunto de semanas foi calculada a média aritmética de

7 termos, seja Y_i a semana e S_i sua média aritmética, portanto $S_i^{2009-2010} = \frac{\sum_{i=i-6}^{i-1} Y_i}{7}$, para 2010-2011 a média aritmética foi calculada com 4 termos, portanto $S_i^{2010-2011} = \frac{\sum_{i=i-3}^{i-1} Y_i}{4}$ (Latorre, 2001), para cada conjunto de dados esse cálculo foi feito a partir da semana Y_i até a semana 1. Os valores de 7 e 4 termos, foram definidos em um processo de tentativa e erro para suavizar os dados reais de dengue.

Para calcular o valor de R_0 utilizando a expressão (21) com valor do intervalo de geração fixado, foi feita a seguinte modificação a fim de sanar o mesmo problema apresentado anteriormente em relação ao cálculo dos fatoriais. Sabe-se que,

$$L(R_0) = \prod_{t=1}^T \binom{j_t}{\sum_{s=0}^{\infty} g_s j_{t-s}} \left(\frac{1}{R_0}\right)^{\sum_{s=0}^{\infty} g_s j_{t-s}} \left(1 - \frac{1}{R_0}\right)^{j_t - \sum_{s=0}^{\infty} g_s j_{t-s}}. \quad (24)$$

Chamando $\sum_{s=0}^{\infty} g_s j_{t-s} = \mu_t$ para simplificar a notação, tem-se

$$L(R_0) = \prod_{t=1}^T \binom{j_t}{\mu_t} \left(\frac{1}{R_0}\right)^{\mu_t} \left(1 - \frac{1}{R_0}\right)^{j_t - \mu_t}. \quad (25)$$

Aplicando o logaritmo,

$$\text{Log}(L(R_0)) = \sum_{t=1}^T \text{Log} \left[\binom{j_t}{\mu_t} \left(\frac{1}{R_0}\right)^{\mu_t} \left(1 - \frac{1}{R_0}\right)^{j_t - \mu_t} \right], \quad (26)$$

$$\text{Log}(L(R_0)) = \sum_{t=1}^T \left[\text{Log} \left(\binom{j_t}{\mu_t} \right) + \mu_t \text{Log} \left(\frac{1}{R_0} \right) + (j_t - \mu_t) \text{Log} \left(1 - \frac{1}{R_0} \right) \right], \quad (27)$$

$$\text{Log}(L(R_0)) = \sum_{t=1}^T \text{Log} \begin{pmatrix} j_t \\ \mu_t \end{pmatrix} + \sum_{t=1}^T \mu_t \text{Log} \left(\frac{1}{R_0} \right) + \sum_{t=1}^T (j_t - \mu_t) \text{Log} \left(1 - \frac{1}{R_0} \right). \quad (28)$$

O termo $\sum_{t=1}^T \text{Log} \begin{pmatrix} j_t \\ \mu_t \end{pmatrix}$ não envolve R_0 e portanto pode ser considerado constante. Chamando-o de k tem-se,

$$\text{Log}(L(R_0)) = k + \sum_{t=1}^T \mu_t \text{Log} \left(\frac{1}{R_0} \right) + \sum_{t=1}^T (j_t - \mu_t) \text{Log} \left(1 - \frac{1}{R_0} \right). \quad (29)$$

Se o cálculo de j_t e μ_t não dependem de R_0 , então o termo k é independente de R_0 e é uma constante. Sendo k constante não é necessário calculá-lo pois maximizar $\text{Log}(L(R_0))$ é equivalente a maximizar $\text{Log}(L(R_0)) - k$, portanto,

$$\text{Log}(L(R_0)) - k = \sum_{t=1}^T \mu_t \text{Log} \left(\frac{1}{R_0} \right) + \sum_{t=1}^T (j_t - \mu_t) \text{Log} \left(1 - \frac{1}{R_0} \right). \quad (30)$$

E dessa forma, fixando os valores do intervalo de geração calculado pela implementação da distribuição binomial em conjunto com a distribuição normal, é possível calcular o valor de R_0 utilizando a equação (30).

Para os dados suavizados foi necessário fazer uma modificação na função (23) proposta por White e Pagano (2008). O software R traz funções como a distribuição Poisson já implementada em seu código mas essa função pronta não aceita receber números decimais em N_t da equação (23), sendo necessário implementar essa função manualmente, o que gerou o mesmo problema que houve com a distribuição Binomial (fatorial de números maiores que 100). Seguiu-se,

portanto, o mesmo raciocínio anterior. Sabe-se que:

$$L(R_0, p) = \prod_{t=1}^T \frac{e^{-\mu_t} \mu_t^{N_t}}{N_t!}, \quad (31)$$

em que $\mu_t = R_0 \sum_{j=1}^{\min(k,t)} N_{t-j} p_j$. Aplicando o logaritmo tem-se

$$\text{Log}(L(R_0, p)) = \sum_{t=1}^T \text{Log} \left[\frac{e^{-\mu_t} \mu_t^{N_t}}{N_t!} \right], \quad (32)$$

$$\text{Log}(L(R_0, p)) = \sum_{t=1}^T -\mu_t + \sum_{t=1}^T N_t \text{Log}(\mu_t) - \sum_{t=1}^T \text{Log}(N_t!). \quad (33)$$

Se o termo N_t não depende de R_0 e p , então este termo é uma constante e, portanto, não é necessário calculá-lo pois maximizar $\text{Log}((R_0))$ é equivalente a maximizar $\text{Log}((R_0)) + \sum_{t=1}^T \text{Log}(N_t!)$. Chamando $\sum_{t=1}^T \text{Log}(N_t!)$ de k , tem-se

$$\text{Log}(L(R_0, p)) = \sum_{t=1}^T -\mu_t + \sum_{t=1}^T N_t \text{Log}(\mu_t) - k, \quad (34)$$

$$\text{Log}(L(R_0, p)) + k = \sum_{t=1}^T -\mu_t + \sum_{t=1}^T N_t \text{Log}(\mu_t). \quad (35)$$

Para a implementação computacional dos métodos de Nishiura (2010) e White e Pagano (2008) foi utilizado o software R versão 2.14.1 e RStudio, a função de otimização *optim* e os métodos de “L-BFGS-B” para quando são calculados o valor de R_0 e intervalo de geração, e “Brent” para quando somente o valor de R_0 é calculado.

A função *optim* do software R é uma rotina de otimização que por padrão minimiza uma função com base nos parâmetros determinados no código. Essa função traz seis métodos de otimização sendo que os métodos “L-BFGS-B” e “Brent” permitem que sejam escolhidos os limites superiores e inferiores dos parâmetros a serem otimizados, o método de “Brent” é o único que permite que somente uma variável seja utilizada na otimização. Todos os outros métodos exigem que sejam utilizadas pelo menos duas variáveis na otimização, e é importante lembrar que o processo de otimização de uma função é utilizado para maximizar ou minimizar funções. Por esta razão este método pode ser utilizado nos problemas verossimilhança.

Método L-BFGS-B

O método de Newton é um método com uma boa convergência no entanto problemas podem surgir se a solução do problema de otimização não ocorre em uma vizinhança próxima do ponto de partida da otimização. A fim de sanar esse tipo de problemas surgem os métodos quase-Newton. Estes métodos tentam resolver este problema mantendo os bons resultados pelo Método de Newton, e um exemplo de método quase-Newton é o Método BFGS (Mendonça, 2005).

O método de BFGS foi desenvolvido simultaneamente por Broyden, Fletcher, Goldfarb e Shanno na década de 60, e consiste em resolver problemas de otimização não linear sem restrições, o qual usa os valores da função e gradientes para construir uma imagem da superfície a ser otimizada (Mendonça, 2005). Este método não requer o cálculo das derivadas segundas o que torna o cálculo mais fácil e rápido.

O método L-BFGS-B foi desenvolvido por Byrd et al. (1994) que utiliza uma matriz BFGS de memória limitada para aproximar a matriz Hessiana da função objetivo. Neste algoritmo é construída uma caixa de restrições, onde

para cada variável são impostos limites superiores e inferiores, e no algoritmo em R deve se dar um palpite inicial para cada variável o qual é claro deve obedecer os limites impostos.

Método Brent

É um método para minimização unidimensional de funções sem que seja necessário o cálculo de derivadas, em que são combinados dois algoritmos para o cálculo de mínimos de funções, o algoritmo da razão áurea e o de interpolação parabólica inversa. O procedimento adotado neste método é rápido e robusto (Azevedo et al., 1997).

2.4 Determinação da fase inicial da doença

A fase inicial da doença foi determinada calculando-se o ponto de inflexão utilizando a função (18). Matematicamente, o ponto de inflexão é o ponto onde a derivada da equação (18) muda de sinal (Guidorizzi, 2001), ou seja, é o ponto onde a curva muda de comportamento. Este cálculo foi feito utilizando o software Mathematica 9.0.

Biologicamente o ponto de inflexão indica o ponto em que a curva de infectados perde sua força de crescimento, ou seja, o fim da fase inicial.

Este cálculo foi feito encontrando a segunda derivada da função (18), igualando-a a zero e encontrando suas raízes (Guidorizzi, 2001).

3 RESULTADOS

Nesta seção são apresentados os cálculos do valor de R_0 para a dengue na cidade de Ribeirão Preto/SP, nos períodos de 2009-2010 e 2010-2011, em ambos os casos a cidade apresentou uma situação epidêmica.

Foi procurado o ponto no tempo equivalente ao início da epidemia, ou seja, a semana em que houve apenas um caso de dengue. Por esta razão o conjunto de dados não foi feito com apenas um ano, pois tanto em 2010 como em 2011, ocorreram mais de cem casos de dengue na primeira semana desses anos, o que naturalmente não corresponde ao início de uma epidemia. O que foi feito em ambos os casos, foi verificar o número de casos nas semanas antecedentes à primeira semana de cada ano, procurando pela semana em que havia apenas um caso de dengue.

Para o período de 2009-2010, tem-se que o ponto de máximo ocorre na 14^a semana do ano de 2010. Já o ponto de inflexão se dá na 21^a semana do conjunto de dados, o que equivale à 5^a semana de 2010. O mesmo foi calculado utilizando a função (18) com $a_1 = 4316690$, $a_2 = -0,268801$, $a_3 = 4,97401$, $a_4 = -0,606375$ e $a_5 = 34,4913$. Para o período de 2010-2011, tem-se que o ponto de máximo ocorre na 14^a semana de 2011. Já o ponto de inflexão se dá na 24^a semana do conjunto de dados, o que equivale à 11^a semana de 2011. O mesmo foi calculado utilizando a função (18) com $a_1 = 15943,8$, $a_2 = -0,168003$, $a_3 = 4,83185$, $a_4 = 0,0543326$ e $a_5 = -8,7884$.

3.1 Cálculo do valor de R_0 utilizando o método de Ross-Macdonald.

A história do R_0 para doenças causadas por vetores começa com os estudos de Sir Ronald Ross que, enquanto trabalhava no Serviço Médico Indiano

nos anos de 1890, demonstrou o ciclo de vida do parasita da malária em um mosquito. Ele foi um dos primeiros a publicar artigos usando funções matemáticas para estudar a transmissão da malária. Ross (1911) desenvolveu um modelo simples que explicava o relacionamento entre o número de mosquitos, e a incidência de malária em humanos. Hoje, este modelo é conhecido como o clássico “Modelo de Ross” (Mandal et al., 2011).

Ross mostrou através de seu modelo, que a redução do número de mosquitos abaixo de um certo limiar era suficiente para conter a malária, o que era um conceito bem a frente de sua época. Depois de cerca de 40 anos George Macdonald modificou o modelo proposto por Ross acrescentando informações biológicas a respeito do período de latência do mosquito (Mandal et al., 2011).

Macdonald (1952), define o R_0 , no apêndice de seu artigo como:

“Taxa de reprodutibilidade basal da malária: Número de infecções distribuídas em uma comunidade como resultado direto da presença de um simples caso primário não imune.”

Ele afirma que o caso primário pode permanecer infeccioso durante o período de $\frac{1}{r}$ dias, em que r é taxa de recuperação da parasitemia. Em cada um desses dias este indivíduo infectado pode ser picado por ma mosquitos, em que m é a densidade de mosquitos em relação aos hospedeiros humanos e a a taxa de picadas. Cada mosquito tem como expectativa de vida após sobreviver n dias,

$$\frac{p^n}{-\log_e p}, \quad (36)$$

em que $\frac{1}{-\log_e p}$ é a expectativa de vida do mosquito e p^n é a fração de mosquitos que sobrevivem após o período de incubação extrínseca.

Cada mosquito pode picar um humano a vezes por dia, e desde que b é a probabilidade de que o humano se torne infectado após ser picado, segue que o número total de infecções é dado por:

$$\frac{ma^2bp^n}{-r\log_e p}. \quad (37)$$

Macdonald (1952) não atribui nenhum símbolo para a representação do R_0 neste artigo, isto só acontece algum tempo depois em 1955 no artigo "*The Measurement of Malaria*", em que ele utiliza o símbolo z_0 .

A equação (37) apresenta um problema de dimensionalidade no termo $\frac{1}{-\log_e p}$, pois se p tem uma dimensão então $\log_e p$ é mal definido, e se p não tem dimensão então R_0 tem dimensão de T^{-1} , o que faz de R_0 uma taxa e não um número (Massad e Coutinho, 2012).

Fazendo algumas modificações na notação utilizada por Macdonald (1952), mas não em seu significado $\frac{1}{-\log_e p} = \frac{1}{\mu_V}$, $r = \gamma$, e $p^n = e^{-\mu_V \tau}$, tem-se:

$$R_0 = \frac{ma^2be^{-\mu_V \tau}}{\gamma\mu_V}. \quad (38)$$

Isto foi feito no artigo de Aron e May (1982), em que eles estudaram a dinâmica da malária utilizando o modelo de Ross-Macdonald (Macdonald, 1952), e obtendo é claro o mesmo R_0 de Ross-Macdonald (Macdonald, 1952). Essas modificações são importantes porque a equação (38) é dimensionalmente correta (Massad e Coutinho, 2012).

Aron e May (1982) fazem uma observação em relação a ausência do parâmetro c que é a probabilidade do mosquito se tornar infectado. Eles observam que seu valor não é 1, ou seja, que a probabilidade de que um mosquito se torne infectado ao picar uma pessoa infectada não é de 100%, e portanto este parâmetro não deveria ser omitido, no entanto, eles não o incluem em seu estudo.

Macdonald (1952) também não considera a taxa de mortalidade dos humanos μ_H , e a ausência deste parâmetro é justificada por Aron e May (1982). Eles afirmam que no caso dos humanos, a saída de humanos infectados da classe dos infectados devido à taxa de recuperação, é muito mais rápida do que pela

taxa de mortalidade, e então ela pode ser omitida.

Incluindo estas taxas e seguindo a mesma linha de raciocínio de Ross-Macdonald (Macdonald, 1952), Massad et al. (2010) apresentam a seguinte derivação intuitiva do R_0 .

Considere que o caso índice da doença é um humano. Seja N_V o número total de vetores, e a sua taxa de picadas nos humanos, então o número de picadas recebidas pelos humanos por unidade de tempo é $N_V a$. Seja N_H a população de humanos e γ sua taxa de recuperação devido a doença, então o caso índice produz $\frac{N_V a}{N_H(\gamma + \mu_H)} c$ vetores infectados, em que c é a probabilidade do vetor se tornar infectado ao picar um humano infectado, e μ_H a taxa de mortalidade dos humanos. Os $\frac{N_V a}{N_H(\gamma + \mu_H)} c$ vetores infectados produzem $\frac{N_V a}{N_H(\gamma + \mu_H)} c \frac{a}{\mu_V} b e^{-\mu_V \tau}$ novos casos na primeira geração da doença. Em que $\frac{1}{\mu_V}$ é a expectativa de vida do mosquito, b a probabilidade do humano se tornar infectado ao ser picado por um vetor infectado, e $e^{-\mu_V \tau}$ é a fração de vetores infectados que sobrevivem ao período de incubação extrínseca, e seja $m = \frac{N_V}{N_H}$,

$$R_0 = \frac{ma^2 b c e^{-\mu_V \tau}}{(\gamma + \mu_H) \mu_V}. \quad (39)$$

O qual é a mesma expressão dada por Ross-Macdonald (Macdonald, 1952) apenas reescrita em outros termos, a qual é a equação (17), já mostrada anteriormente no capítulo Métodos.

Massad et al. (2010), estudaram o número de reprodutibilidade basal utilizando o modelo de Ross-Macdonald (Macdonald, 1952) para doenças causadas por vetores, onde as populações são divididas em: população de humanos N_H a qual é dividida em: suscetíveis $S_H(t)$, infectados $I_H(t)$ e recuperados $R_H(t)$; população de vetores N_V a qual é dividida em: suscetíveis $S_V(t)$, vetores em estado latente $L_V(t)$ e infectados $I_V(t)$.

A dinâmica da população de humanos se dá da seguinte forma: os

humanos suscetíveis, S_H , podem se tornar infectados, I_H , através da taxa de picadas, a , com probabilidade de ficar infectado, b , e se tornam recuperados, R_H , com taxa de recuperação, γ , a população de humanos tem taxa de mortalidade, μ_H . A dinâmica da população de vetores se dá da seguinte forma: os vetores suscetíveis, S_V , se tornam latentes, L_V , após picar um humano infectado, com probabilidade de se tornar infectado, c , a fração de vetores latentes que sobrevivem ao período de incubação extrínseco de incubação, $e^{-\mu_V\tau}$, se tornam infectados, I_V , a população de vetores tem taxa de mortalidade, μ_V . Na Figura (7), tem-se o diagrama para o modelo compartimental de Ross-Macdonald (Macdonald, 1952).

Figura 7: Diagrama do modelo compartimental referente ao Modelo de Ross-Macdonald (Macdonald, 1952), o sistema (40).

As equações abaixo descrevem a dinâmica da doença:

$$\left\{ \begin{array}{l} \frac{dS_H(t)}{dt} = -abI_V(t)\frac{S_H(t)}{N_H} + \mu_H(N_H - S_H(t)), \\ \frac{dI_H(t)}{dt} = abI_V(t)\frac{S_H(t)}{N_H} - (\mu_H + \gamma)I_H(t), \\ \frac{dR_H(t)}{dt} = \gamma I_H(t) - \mu_H R_H(t), \\ \frac{dS_V(t)}{dt} = -acS_V(t)\frac{I_H(t)}{N_H} + \mu_V(N_V - S_V(t)), \\ \frac{dL_V(t)}{dt} = acS_V(t)\frac{I_H(t)}{N_H} - \mu_V L_V(t) - acS_V(t - \tau)\frac{I_H(t - \tau)}{N_H}e^{-\mu_V\tau}, \\ \frac{dI_V(t)}{dt} = acS_V(t - \tau)\frac{I_H(t - \tau)}{N_H}e^{-\mu_V\tau} - \mu_V I_V(t). \end{array} \right. \quad (40)$$

A população é mantida constante portanto tem-se que:

$$N_H = S_H + I_H + R_H.$$

$$N_V = S_V + L_V + I_V.$$

Em seu estudo Massad et al. (2010), mostram que R_0 dado pela equação (39), pode ser dividido em dois componentes: $T_{(V-H)}$ (componente vetor - humano) que representa a transmissão da doença do vetor para o humano, e $T_{(H-V)}$ (componente humano - vetor) que representa a transmissão da doença do humano para o vetor.

$$T_{(H-V)} = \frac{N_V a}{N_H(\gamma + \mu_H)} c, \quad (41)$$

e

$$T_{(V-H)} = \frac{a}{\mu_V} b e^{-\mu_V \tau}. \quad (42)$$

Massad et al. (2010), fizeram uma análise do R_0 , onde foi demonstrado que quando $R_0 < 1$, a concavidade da curva de crescimento da fase inicial é voltada para baixo, e ao longo do tempo tende a zero. Quando $R_0 > 1$, a concavidade da curva de crescimento da fase inicial é voltada para cima, e tende a um estado endêmico ao longo do tempo.

Eles mostraram que se $T_{(V-H)} > 1$, e $T_{(H-V)} < 1$, obtém-se $R_0 < 1$, e portanto a introdução de um certo número de vetores pode provocar um surto epidêmico, mas a concavidade dessa curva é voltada para baixo e ao longo do tempo tende a zero.

Uma questão levantada a respeito desse modelo, é a respeito da inclusão ou não do parâmetro de retardo, τ , que corresponde ao período de incubação extrínseco. Tem-se que a expectativa de vida do mosquito é: $\frac{e^{-\mu_V \tau}}{\mu_V}$. No entanto, este termo é igual a $\frac{\gamma_V}{\mu_V(\gamma_V + \mu_V)}$, em que γ_V é taxa de recuperação do mosquito e poderia substituir o termo com retardo. Porém, os resultados finais obtidos utilizando uma expressão ou outra é a mesma. No presente estudo foi

escolhido utilizar o modelo com retardo.

O Modelo de Ross-Macdonald (Macdonald, 1952), foi implementado utilizando o Método Runge-Kutta de 4ª Ordem e o software Berkeley-Madonna. Na Tabela (1), são dados os parâmetros que foram mantidos fixos durante a simulação, e seus valores biológicos (Massad et al., 2010).

Tabela 1: Valores biológicos utilizados na simulação.

Parâmetro	Valor (Massad et al., 2010)
γ	$0,98 \text{ semanas}^{-1}$
τ	1 semana
μ_H	$2,38 \times 10^{-4} \text{ semanas}^{-1}$
μ_V	$0,175 \text{ semanas}^{-1}$

Os parâmetros ajustados foram: a variando entre $1\text{-}3 \text{ semanas}^{-1}$ (Favier et al., 2006; Massad et al., 2010), b, c variando entre $0\text{-}1$ e m variando entre $0\text{-}1$ (Rodrigues et al., 2015). Apesar das informações acerca desses parâmetros serem escassas na literatura, tomou-se o cuidado de não serem tomados valores não físicos. A escolha da variação dos termos b e c foi fácil de ser definida já que esses parâmetros são probabilidades e portanto só poderiam variar entre 0 e 1 . Em relação a taxa de picadas acredita-se que os mosquitos em geral se alimentam duas ou até três vezes durante seu ciclo gonotrófico (Gubler e Kuno, 1997). Sobre a densidade de mosquitos existem estudos que dizem que em geral o número de mosquitos *Aedes aegypti* por pessoa é baixo (Scott e Morrison, 2010), e em uma pesquisa feita na cidade de São Sebastião, região litorânea do estado de São Paulo foi obtido que a densidade de fêmeas por residentes variou de $0,02\text{-}0,64$ considerando a variação total mostrada no estudo (Rodrigues et al., 2015). As condições iniciais foram dadas por $I_H = 1$, $R_H = 0$, $S_V = m * N_H$, $L_V = 0$ e $I_V = 0$.

A fim de se obter o tamanho aproximado da população de suscetíveis, foi feita a soma da população de pessoas infectadas de 10 anos anteriores ao período em estudo, e este valor foi subtraído da população total de suscetí-

veis. Para o período de 2009-2010 a população de suscetíveis variou entre 550000-589168, e para o período de 2010-2011 a população de suscetíveis variou entre 550000-582546.

A Figura 8 (a) mostra o ajuste da equação (18) aos dados reais, e a Figura 8 (b) mostra o ajuste do Modelo de Ross-Macdonald (Macdonald, 1952) a equação (20). Foi obtido como resultado o valor de $R_0 = 5,44$, com $a = 1,82257$ semanas⁻¹, $b = 0,692187$, $c = 0,601626$, $m = 0,80429$, e $S_H = 589162$.

Figura 8: Figura (a) em pontos azuis os dados reais, na linha em vermelho o ajuste dado pela equação (18) no período de 2009-2010, na Figura (b) a linha azul apresenta os indivíduos infectados dados pela equação (20), na linha em vermelho o ajuste obtido utilizando o Modelo de Ross-Macdonald (Macdonald, 1952) (sistema de equações (40)) no período de 2009-2010.

A Figura 9 (a) mostra o ajuste da equação (18) aos dados reais, e a Figura 9 (b) mostra o ajuste do Modelo de Ross-Macdonald (Macdonald, 1952) a equação (20). Foi obtido como resultado o valor de $R_0 = 5,41$, com $a = 2,36399$ semanas⁻¹, $b = 0,314763$, $c = 0,823785$, $m = 0,762838$, e $S_H = 550000$ *

*As taxas obtidas para a, b, c, m são quase puramente decimais ao arredondá-las e aplicá-las a expressão de R_0 dada por Ross-Macdonald (Macdonald, 1952) foi obtido um valor muito próximo, porém diferente do valor obtido pelo programa, por isso foram escritas todas as casas decimais no texto da tese.

Figura 9: Figura (a) em pontos azuis os dados reais, na linha em vermelho o ajuste dado pela equação (18) no período de 2010-2011, na Figura (b) a linha azul apresenta os indivíduos infectados dados pela equação (20), na linha em vermelho o ajuste obtido utilizando o Modelo de Ross-Macdonald (Macdonald, 1952) (sistema de equações (40)) no período de 2010-2011.

3.2 Cálculo do valor de R_0 utilizando os métodos de Nishiura e White e Pagano.

Para os cálculos utilizando os métodos de Nishiura (2010) e White e Pagano (2008) foram necessárias informações sobre os dados de incidência e sobre o intervalo de geração. No caso da dengue, o mesmo varia entre 14-19 dias, e considerando-se a variação total apresentada nos estudos de Bennett et al. (2003), Aldstadt et al. (2012), e Aldstadt (2007), assumiu-se o valor médio de 2,35 semanas compreendendo os períodos intrínseco (dentro de humano) e extrínseco (dentro de mosquito), com variância no intervalo entre 0,28-0,43, o que corresponde ao intervalo de 2-3 dias.

3.2.1 Método de Nishiura.

Segundo Nishiura (2010), para calcular o valor de R_0 a partir de dados epidêmicos deve-se transformar r , que é definido como a taxa de crescimento da doença, em R_0 através do procedimento dado por ele a seguir.

Seja $j(t)$ o número de novas infecções no tempo t , e seja $A(\tau)$ a taxa de geração de casos secundários, em que τ é o momento no tempo em que a doença é adquirida,

$$j(t) = \int_0^\infty A(\tau)j(t-\tau)d\tau. \quad (43)$$

Na equação (43) é possível observar que o número de indivíduos infectados no tempo t é igual ao número de infectados τ unidades de tempo atrás, multiplicado pela taxa de geração de casos secundários (Wallinga e Lipsitch, 2007).

Uma vez que R_0 representa o total de casos secundários que são gerados durante todo o período de uma doença, tem-se que

$$R_0 = \int_0^\infty A(\tau)d\tau. \quad (44)$$

Nesta parte do artigo Nishiura considera que $j(t)$ segue um crescimento exponencial, portanto,

$$j(t) = Ke^{rt}, \quad (45)$$

em que K é constante, e além disso,

$$j(t-\tau) = Ke^{rt}e^{-r\tau}. \quad (46)$$

Substituindo (46) em (43),

$$1 = \int_0^\infty e^{-r\tau} A(\tau)d\tau. \quad (47)$$

Seja $g(\tau)$ definido como a frequência de casos secundários, então,

$$g(\tau) = \frac{A(\tau)}{\int_0^\infty A(s)ds} = \frac{A(\tau)}{R_0}. \quad (48)$$

Isto implica que,

$$A(\tau) = R_0g(\tau). \quad (49)$$

Substituindo (49) do lado direito da equação de (47), obtém-se

$$\frac{1}{R_0} = \int_0^\infty g(\tau) e^{-r\tau} d\tau. \quad (50)$$

O termo do lado direito da equação (50) é conhecido como função geratriz de momentos $M(z)$ da distribuição $g(\tau)$ em probabilidade (Wallinga e Lipsitch, 2007).

$$M(z) = \int_0^\infty e^{za} g(a) da. \quad (51)$$

Usando a equação (51) para simplificar a notação de (50). Fazendo $z = -r$,

$$\frac{1}{R_0} = M(-r), \quad (52)$$

$$R_0 = \frac{1}{M(-r)}. \quad (53)$$

Substituindo (53) em (50), obtém-se o estimador

$$\hat{R}_0 = \frac{1}{\int_0^\infty e^{-r\tau} g(\tau) d\tau}. \quad (54)$$

Segundo Wallinga e Lipsitch (2007), tem-se que $M(z)$ determina $g(\tau)$ e $g(\tau)$ determina $M(z)$. O corolário biológico da função geratriz de momentos é o relacionamento entre a taxa de crescimento r e o número de reprodutibilidade basal R_0 , que caracteriza unicamente a forma da distribuição do intervalo de geração e determina a apropriada relação entre R_0 e r .

Neste caso o valor de R_0 é encontrado através do valor de r e também assumindo $g(\tau)$ ser conhecido ou calculado separadamente (Nishiura, 2010).

R_a é definido como o número médio de novas infecções originadas por indivíduos infectados já existentes, e ele se aplica à média da vida infecciosa dos indivíduos infectados para fazer uma avaliação em um período específico de tempo, a taxa de incidência à prevalência é utilizada para determinar seu valor (White et al., 2006).

A probabilidade de estar infectado no tempo t é dada por,

$$p(t) = \int_0^\infty \Psi(\tau)j(t-\tau)d\tau, \quad (55)$$

em que $\Psi(\tau)$ é a probabilidade de estar infectado no tempo τ .

Seja $\beta(\tau)$ a taxa de transmissão da doença, que depende da frequência de contatos e da infecciosidade no tempo τ . Portanto, $A(\tau)$ pode ser decomposto como $A(\tau) = \beta(\tau)\Psi(\tau)$ (Chowell e Nishiura, 2008). R_a é dado por,

$$R_a = \frac{j(t)}{p(t)}D = \frac{\int_0^\infty \beta(\tau)\Psi(\tau)j(t-\tau)d\tau}{\int_0^\infty \Psi(\tau)j(t-\tau)d\tau}D, \quad (56)$$

onde D é descrito como a média do intervalo de geração. Se $\beta(\tau)$ é constante, ou seja, se ele não depende do tempo, tem-se

$$R_0 = \int_0^\infty \beta(\tau)\Psi(\tau)d\tau = \beta D. \quad (57)$$

Portanto, neste caso com β considerado constante, Nishiura (2010) conclui que o número de reprodutibilidade basal e o número de reprodutibilidade atual são iguais.

Nishiura diz que no entanto R_a tende a ser um cálculo enviesado, e que R_a não é robusto como o R_0 obtido pela equação (54).

Então Nishiura propõe uma correção à definição de R_a a fim de solucionar este problema.

O conceito de R_a é corrigido e a equação (56) é reescrita como,

$$R_a = \frac{j(t)}{\frac{1}{D} \int_0^\infty \Psi(\tau) j(t - \tau) d\tau}, \quad (58)$$

onde o numerador representa o número de novas infecções no tempo t , e o denominador que antes representava o número total de indivíduos infectados, aqueles que poderiam potencialmente ser casos primários no tempo t , passa a ter a seguinte definição: número total de contatos efetivos, aqueles que podem potencialmente levar a transmissões secundárias. Esta modificação foi feita a fim de que o estimador R_a coincida com R_0 .

Portanto R_a é escrito como:

$$\bar{R}_a = \frac{j(t)}{\int_0^\infty g(\tau) j(t - \tau) d\tau}. \quad (59)$$

Onde $g(\tau)$ é escrita como a densidade normalizada de transmissões secundárias:

$$g(\tau) = \frac{\beta(\tau)\Psi(\tau)}{\int_0^\infty \beta(\tau)\Psi(\tau) d\tau} = \frac{\beta(\tau)\Psi(\tau)}{R_0}. \quad (60)$$

Substituindo (60) em (59), tem-se,

$$\bar{R}_a = \frac{j(t)}{\frac{1}{R_0} \int_0^\infty \beta(\tau)\Psi(\tau) j(t - \tau) d\tau} = R_0. \quad (61)$$

Portanto Nishiura (2010) mostra que $R_a = R_0$, ou seja, o valor de R_0 pode ser calculado utilizando dados de incidência e o intervalo de geração sem assumir crescimento exponencial dos casos na fase inicial da doença.

Uma vez que os dados utilizados para o cálculo do R_0 são discretos

a equação (59), precisa ser reescrita em tempo discreto

$$\hat{R}_0 = \frac{j_t}{\sum_{s=0}^{\infty} g_s j_{t-s}}, \quad (62)$$

onde g_s é função de densidade discretizada dada por,

$$g_s = G(s+1) - G(s), \quad (63)$$

$G(s)$ é a função de probabilidade acumulada do intervalo de geração de tamanho s . Assumindo que o intervalo de geração segue uma Distribuição Gama tem-se que,

$$g_s = \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^{s+1} x^{\alpha-1} e^{-\lambda x} dx - \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^s x^{\alpha-1} e^{-\lambda x} dx, \quad (64)$$

$$g_s = \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_s^{s+1} x^{\alpha-1} e^{-\lambda x} dx. \quad (65)$$

O próximo passo para obtenção da expressão de R_0 em tempo discreto é inverter ambos os lados da equação (62),

$$\frac{1}{R_0} = \frac{\sum_{s=0}^{\infty} g_s j_{t-s}}{j_t}, \quad (66)$$

o numerador do lado direito indica o número total de contatos efetivos feitos por potenciais casos primários no tempo t , que são aqueles que tem a mesma probabilidade de resultar em uma transmissão secundária, e o denominador é o número de casos secundários no tempo t .

Dado que o número total de contatos efetivos tem o mesmo potencial para resultar em uma transmissão secundária, a probabilidade de um caso sozinho resultar em uma transmissão secundária é dada por $\frac{1}{R_0}$. Este é um processo

binomial, e então R_0 pode ser calculado pela função de verossimilhança,

$$L(R_0) = \prod_{t=1}^T \binom{j_t}{\sum_{s=0}^{\infty} g_s j_{t-s}} \left(\frac{1}{R_0} \right)^{\sum_{s=0}^{\infty} g_s j_{t-s}} \left(1 - \frac{1}{R_0} \right)^{j_t - \sum_{s=0}^{\infty} g_s j_{t-s}}, \quad (67)$$

em que T é o ponto de tempo máximo de observação na fase inicial de crescimento da epidemia, e R_0 é obtido minimizando o negativo do logaritmo de (67), a qual é a equação (21), já mostrada anteriormente no capítulo Métodos.

O termo $\sum_{s=0}^{\infty} g_s j_{t-s}$ não tem significado físico pois, o tempo assumido pelo intervalo de geração, g_s , é finito, e também porque ao executar o cálculo de $\sum_{s=0}^{\infty} g_s j_{t-s}$, para $s > t$ tem-se que j_{t-s} assume valores de tempo negativo. Portanto, este termo foi substituído por $\sum_{s=1}^{\min(s_{max}, t)} g_s j_{t-s}$ na equação (67), como é feito por White e Pagano (2008). Em que s_{max} é o valor máximo do intervalo de geração, e a condição $\min(s_{max}, t)$ garante que j_{t-s} não assuma valores de tempo negativo.

Em seu estudo Nishiura (2010) calcula o valor de R_0 para HIV/AIDS, no qual relata que esta doença apresenta três intervalos de geração diferentes ao longo do tempo. A dengue apresenta apenas um intervalo de geração, portanto substituiu-se (63) por,

$$g_s = \frac{G(s) - G(s-1)}{G(s_{max})}, \quad (68)$$

em que $g_0 = 0$ e s_{max} é o valor máximo para o tempo de infecção causado por transmissões secundárias, como dito anteriormente. Esta expressão para o intervalo de geração é dada por Roberts e Nishiura (2011) para o cálculo do intervalo de geração da H1N1, que assim como a dengue apresenta apenas um intervalo de geração.

Portanto g_s é dado por,

$$g_s = \frac{\frac{\lambda^\alpha}{\Gamma(\alpha)} \int_{s-1}^s x^{\alpha-1} e^{-\lambda x} dx}{\frac{\lambda^\alpha}{\Gamma(\alpha)} \int_1^{s_{max}} x^{\alpha-1} e^{-\lambda x} dx}. \quad (69)$$

O valor de s_{max} é finito e deve ser escolhido de forma que seja suficientemente grande para garantir que g_s segue uma Distribuição Gama completa (White e Pagano, 2008).

Neste estudo foram feitos dois cálculos: no primeiro foram calculados intervalo de geração e R_0 , utilizando a fórmula da distribuição Binomial em conjunto com a distribuição normal, com R_0 variando entre 1.2-10; e no segundo cálculo foi calculado somente o valor de R_0 utilizando a expressão (30) que é equivalente a (21), com R_0 variando entre 1-10. O intervalo de geração foi calculado utilizando a equação (69) com $s_{max} = 8$, este valor garante que a Distribuição Gama assume sua forma completa, naturalmente foram feitos teste de tentativa e erro para encontrar este valor. Em todos os cálculos feitos utilizando o método de Nishiura (2010) foram utilizados dados suavizados.

Para o período de 2009-2010 foram obtidos os valores de $R_0 = 3,10$ e variância de 0,43 utilizando a fórmula da Binomial em conjunto com a distribuição normal, e $R_0 = 3,11$ utilizando a equação (30) com variância fixada em 0,43.

No período de 2010-2011 foram obtidos os valores de $R_0 = 1,84$ e variância de 0,43 utilizando a fórmula da Binomial em conjunto com a distribuição normal, e $R_0 = 1,83$ utilizando a equação (30) com variância fixada em 0,43.

3.2.2 Método de White e Pagano.

Seja N_t o número de novos casos de uma doença no tempo t , sendo t uma unidade de tempo discreta podendo ser dias, semanas, meses ou anos. Foi suposto que o número de casos secundários gerados por um indivíduo infectado segue uma Distribuição de Poisson, tendo R_0 como valor esperado, e que o intervalo de geração é dado por uma distribuição multinomial. A suposição de que o

intervalo de geração segue uma distribuição multinomial implica que após algum tempo a probabilidade de casos secundários é desprezível, e foi assumido também que os sintomas dos casos primários sempre surgem antes de seus casos secundários. Não há movimento de saída e entrada de indivíduos do sistema de casos infectados, existe uma mistura homogênea dos indivíduos e os surtos ocorrem da seguinte maneira: seja N_0 o número de indivíduos que foram infectados no início da epidemia, sabe-se que cada caso gera casos secundários de maneira independente segundo uma distribuição de Poisson com média R_0 . Seja X_0 o número total de casos produzidos por N_0 , então $X_0 \sim Pois(N_0 R_0)$. N_i representa o número de casos no dia i , X_{ij} representa o número de casos no dia j , que foram gerados por pelos N_i casos, e X_i é o número total de casos produzidos por casos primários no dia i .

White e Pagano (2008) dão um exemplo do procedimento anterior, em que é suposto que o intervalo máximo do intervalo de geração é 3 unidades de tempo,

$$\begin{aligned}
 N_0 & \\
 N_1 &= X_{01} \\
 N_2 &= X_{02} + X_{12} \\
 N_3 &= X_{03} + X_{13} + X_{23} \\
 N_4 &= X_{14} + X_{24} + X_{34} \\
 \vdots & \qquad \qquad \qquad \vdots \quad \vdots
 \end{aligned} \tag{70}$$

No entanto (70) não traz informações referentes ao momento em que a interação que provocou a infecção aconteceu, e traz apenas a informação do momento no qual os casos se tornaram sintomáticos. White e Pagano (2008) dizem que não observaram X_{ij} . Se isto fosse possível seria fácil calcular o valor de R_0 e do vetor de probabilidades p da distribuição multinomial, e que poderiam

fazer isto construindo a verossimilhança de X_{ij} , dada por

$$\begin{aligned}
L(R_0, \mathbf{p} | \mathbf{N}, \mathbf{X}) = & \left[\frac{e^{-N_0 R_0} (N_0 R_0)^{X_{0\cdot}}}{X_{0\cdot}!} \right] \left[\begin{pmatrix} X_{0\cdot} \\ X_{01} \dots X_{0,1+k} \end{pmatrix} p_1^{X_{01}} \dots p_k^{X_{0,k}} \right] \\
& \times \left[\frac{e^{-N_1 R_0} (N_1 R_0)^{X_{1\cdot}}}{X_{1\cdot}!} \right] \left[\begin{pmatrix} X_{1\cdot} \\ X_{12} \dots X_{1,1+k} \end{pmatrix} p_1^{X_{12}} \dots p_k^{X_{1,k}} \right] \dots \\
& \times \left[\frac{e^{-N_T R_0} (N_T R_0)^{X_{T\cdot}}}{X_{T\cdot}!} \right] \left[\begin{pmatrix} X_{T\cdot} \\ X_{T,T+1} \dots X_{T,T+k} \end{pmatrix} p_1^{X_{T,T+1}} \dots p_k^{X_{T,T+k}} \right].
\end{aligned} \tag{71}$$

Eles fazem algumas considerações sobre (71), em que rearranjaram os termos tal que $X_{i,T+l}$ com $l > 0$ pode ser normalizado e resumido a uma Poisson com variáveis aleatórias. Ainda rearranjaram os termos restantes resumindo os casos não observados $X_{i,j}$ como variáveis aleatórias binomiais e multinomiais, então (71) é reduzido a uma Poisson:

$$L(R_0, p) = \prod_{t=1}^T \frac{e^{-\mu_t} \mu_t^{N_t}}{N_t!}, \tag{72}$$

a qual é a equação (23), já mostrada anteriormente no capítulo e Métodos. Em que $\mu_t = R_0 \sum_{j=1}^{\min(k,t)} N_{t-j} p_j$, através desta forma simplificada é possível usar técnicas de máximo verossimilhança para calcular o valor de R_0 e p_j , com $j = 1, \dots, k$. Neste ponto eles enfatizam a importância do valor de k (s_{max} em Nishiura (2010)), pois seu valor está diretamente ligado ao cálculo correto do valor de R_0 , se k não for suficientemente grande ele não representa a distribuição de probabilidade completa assumida pelo intervalo de geração.

A seguir são dados os cálculos para quando o intervalo de geração é conhecido. Aqui eles estão interessados em descrever um método para o cálculo do R_0 em tempo real que possa ser feito enquanto a infecciosidade ainda está em curso, o que neste caso tem a particularidade de ser de fácil implementação. No primeiro passo White e Pagano (2008) mostram que o R_0 pode ser derivado através

de cálculos de verossimilhança a partir de (72). Eles mostram que esta expressão relaciona-se a um Processo de Ramificação, e que descreve resultados pertinentes às aplicações de interesse deles.

De (72), é obtido,

$$U_{R_0} = \sum_{t=1}^T \frac{(N_t - \mu_t)}{R_0}, \quad (73)$$

em que $\mu_t = R_0 \sum_{j=1}^{\min(k,t)} N_{t-j} p_j$, e resolvendo-se esta equação chega-se a seguinte expressão:

$$\widehat{R_0} = \frac{\sum_{t=1}^T N_t}{\sum_{t=1}^T \sum_{s=1}^{\min(k,t)} p_j N_{t-j}}. \quad (74)$$

A qual é a equação (22), já mostrada anteriormente no capítulo Métodos. a expressão (74) pode ser comparado ao Processo de Ramificação da média de descendentes. Em Processos de Ramificação os dados podem ser agrupados em gerações, chegando-se a uma expressão para R_0 que é bem próxima de (74). Por exemplo, se a média do intervalo de geração é dada por μ dias e o vetor de dados em dias é dado por $N = N_1, N_2, \dots, N_T$, este vetor pode ser agrupado segundo a média dada pelo intervalo de geração como $N^* = \{N_0, \sum_{t=1}^{\mu} N_t, \sum_{t=\mu+1}^{2\mu} N_t, \dots, \sum_{t=(m-1)\mu+1}^{m\mu} N_t\}$, em que $T/\mu = m$, então os dados do vetor de parâmetros N^* pode ser usado para calcular o valor de R_0 ,

$$R_0 = \frac{\sum_{i=1}^m N_i^*}{\sum_{i=1}^m N_{i-1}^*}. \quad (75)$$

Nos dois casos são necessárias apenas informações acerca dos dados de incidência e do intervalo de geração. No entanto, (75) requer apenas informações sobre a média do intervalo de geração, mas a desvantagem é que, quando o valor do intervalo de geração não é um número inteiro, este método se torna difícil de implementar, sendo necessário arredondar a média ou redistribuir os dados de

alguma maneira (White e Pagano, 2008).

A proximidade entre (74) e (75) ajuda a entender melhor a ideia dada pela equação (74). Processos de Ramificação fornecem informações acerca da probabilidade de extinção ou permanência de uma epidemia (White e Pagano, 2008) e são frequentemente usados em epidemiologia para fazer uma aproximação do estágio inicial de um surto epidêmico, onde o esgotamento do número de suscetíveis é negligenciável (Brauer et al., 2008).

As expressões são bem próximas e a diferença é que em (74) é levado em consideração as probabilidades ligadas ao tempo do intervalo de geração, enquanto que em (75) os dados são apenas agrupados com relação ao valor médio do intervalo de geração.

White e Pagano (2008) também descrevem um método para o cálculo do valor de R_0 para o caso em que o intervalo de geração não é conhecido, e quando se tem poucas informações a seu respeito, pois alguns problemas podem surgir se os valores assumidos para o intervalo de geração forem incorretos, levando a um valor de R_0 ruim. Então eles ampliam o método descrito anteriormente e calculam ao mesmo tempo R_0 e o intervalo de geração, considerado a verossimilhança dada por (72).

Eles utilizam técnicas de máxima verossimilhança para calcular R_0 e p_j , $j = 1, \dots, k$, o intervalo de geração pode ser modelado através de uma Distribuição Gama, a qual tem flexibilidade suficiente para modelar um grande número de conjunto de dados. Isso implica em calcular apenas três parâmetros: R_0 , α , e λ , sendo que α , e λ são respectivamente os parâmetros de forma e escala da Distribuição Gama. Da mesma forma que é apresentado em Nishiura (2010).

White e Pagano (2008) apresentam duas expressões para calcular o valor de R_0 , uma para o intervalo de geração conhecido e outra para calcular R_0 e intervalo de geração ao mesmo tempo. Foram feitos os cálculos utilizando as

duas expressões, e, assim como para o método de Nishiura (2010) para o caso em que foi calculado somente o valor de R_0 , foram utilizados os valores do intervalo de geração dados pelos cálculos da verossimilhança em que foram calculados R_0 e intervalo de geração. O intervalo de geração foi calculado utilizando a equação (69) com $k = 8$, este valor garante que a Distribuição Gama assume sua forma completa, naturalmente foram feitos teste de tentativa e erro para encontrar este valor. R_0 foi variado entre 1-10.

Para o período de 2009-2010 foram obtidos os valores de $R_0 = 2,67$ e variância de 0,43 utilizando a equação (23) para dados não suavizados, e de $R_0 = 3,11$ e variância de 0,43 utilizando a equação (35) com dados suavizados, $R_0 = 2,67$ para a expressão (22) para dados não suavizados e $R_0 = 3,12$ para a expressão (22) com os dados suavizados, ambos com variância fixada em 0,43.

No período de 2010-2011 foram obtidos os valores de $R_0 = 1,78$ e variância de 0,43 utilizando a equação (23) para dados não suavizados, e de $R_0 = 1,83$ e variância de 0,43 utilizando a equação (35) com dados suavizados, $R_0 = 1,78$ para a fórmula (22) para dados não suavizados e $R_0 = 1,84$ para a expressão (22) com os dados suavizados, ambos com variância fixada em 0,43.

A seguir são dados os cálculos para o valor de R_0 utilizando todos os métodos, com o objetivo de comparar seus resultados ao longo do tempo. Calculando-se seu valor a partir da primeira semana do conjunto de dados 2009-2010 até a primeira semana de 2010, depois a partir da primeira semana do conjunto de dados até a segunda semana de 2010, e assim por diante até a semana que apresenta o maior valor de infectados neste período o que corresponde a 14^a semana de 2010. O mesmo procedimento foi feito para o período de 2010-2011, neste caso, a semana que apresentou o maior número de infectados coincidentemente corresponde a 14^a semana de 2011.

Tabela 2: Valores de R_0 para o período de 2009-2010.

Número de Semanas	Eq.(23)	Eq.(30)	Binom+Normal	Eq.(35)	Eq.(22)	Eq.(22)*	Eq.(17)
17	2,96	2,58	2,58	2,58	2,97	2,62	4,70
18	3,79	2,99	2,97	2,99	3,80	3,02	5,05
19	3,90	3,32	3,31	3,32	3,90	3,34	5,35
20	3,25	3,34	3,33	3,34	3,25	3,35	5,46
21	2,67	3,11	3,10	3,11	2,67	3,12	5,44
22	2,54	2,91	2,90	2,91	2,54	2,92	5,40
23	2,42	2,75	2,74	2,75	2,42	2,75	5,38
24	2,29	2,59	2,58	2,59	2,29	2,59	5,32
25	2,10	2,41	2,40	2,41	2,10	2,41	5,25
26	1,96	2,24	2,23	2,24	1,96	2,24	5,17
27	1,82	2,09	2,09	2,09	1,82	2,09	5,06
28	1,68	1,96	1,96	1,96	1,68	1,96	4,94
29	1,58	1,84	1,84	1,84	1,58	1,84	4,81
30	1,52	1,73	1,75	1,73	1,52	1,73	4,69

Binom+Normal= distribuição binomial em conjunto com a distribuição normal,eq.(22)=eq.(22) com dados suavizados.

Tabela 3: Valores de R_0 para o período de 2010-2011.

Número de Semanas	Eq.(23)	Eq.(30)	Binom+Normal	Eq.(35)	Eq.(22)	Eq.(22)*	Eq.(17)
14	2,25	2,18	2,18	2,18	2,26	2,24	5,97
15	2,35	2,26	2,26	2,26	2,35	2,30	5,98
16	2,54	2,39	2,38	2,39	2,55	2,42	5,99
17	2,53	2,47	2,47	2,47	2,53	2,50	5,99
18	2,81	2,63	2,63	2,63	2,81	2,65	6,03
19	2,79	2,72	2,72	2,72	2,79	2,74	6,09
20	2,37	2,60	2,60	2,60	2,37	2,61	6,04
21	1,97	2,35	2,34	2,35	1,97	2,36	5,84
22	1,88	2,13	2,12	2,13	1,88	2,13	5,84
23	1,78	1,94	1,94	1,94	1,78	1,94	5,49
24	1,78	1,83	1,84	1,83	1,78	1,84	5,41
25	1,75	1,79	1,80	1,79	1,75	1,79	5,23
26	1,66	1,73	1,74	1,73	1,66	1,73	5,09
27	1,66	1,70	1,71	1,70	1,66	1,70	4,99

Binom+Normal= distribuição binomial em conjunto com a distribuição normal,eq.(22)=eq.(22) com dados suavizados.

4 DISCUSSÃO

Neste trabalho foram estudados alguns métodos para o cálculo do valor de R_0 pela fase inicial de crescimento de doenças causadas por vetores, em que não é suposto o crescimento exponencial dos casos. Foram analisados dados de dengue, sendo que um dos métodos é específico para doenças causadas por vetores, e os outros são métodos gerais para o cálculo do R_0 .

Inicialmente seriam estudados apenas dois métodos, o de Ross-Macdonald (Macdonald, 1952), o qual é específico para doenças causadas por vetores e foi inicialmente desenvolvido para malária, e o método de Nishiura (2010) que é um método geral para o cálculo de R_0 baseado em verossimilhança. No entanto, ao estudar este método, foi encontrado um estudo feito por White e Pagano (2008), o qual é semelhante ao proposto por Nishiura (2010), e os dois estudos utilizam dados de incidência e informações sobre o intervalo de geração e são baseados em verossimilhança.

Nishiura (2010), tem como ponto inicial de sua metodologia, o Número de Reprodução Atual (R_a), o qual é corrigido a fim de demonstrar que R_a , e R_0 são iguais. Já White e Pagano (2008), desenvolvem a ideia do cálculo de R_0 relacionado a um processo de ramificação. Os dois trabalhos findam em expressões de verossimilhança, que utilizam dados de incidência, e intervalo de geração, expressões estas que diferem basicamente na distribuição utilizada, Nishiura (2010) utiliza a Distribuição Binomial, e White e Pagano (2008) utilizam a Distribuição de Poisson.

Neste trabalho, foram utilizados métodos diferentes, e foram obtidos resultados diferentes, principalmente se olharmos apenas para seus valores numéricos. Analisando os resultados calculados utilizando o ponto de inflexão, obteve-se que todos os métodos cumpriram a “tarefa” do R_0 , que é dizer se a doença irá se manter ou não na população. Neste caso, todos os métodos concordaram que para

os dois períodos estudados, haveria um surto epidêmico de dengue em Ribeirão Preto, o que realmente aconteceu nesta cidade.

Para avaliar melhor o comportamento de todos os métodos, ao longo do curso inicial da epidemia, foram calculados os valores de R_0 para a cidade de Ribeirão Preto, começando da primeira semana do conjunto de dados de 2009-2010 até a primeira semana de 2010, depois a partir da primeira semana do conjunto de dados até a segunda semana de 2010, e assim por diante, e terminando no ponto de máximo que corresponde à 14^a semana de 2010. O mesmo procedimento foi feito para o período de 2010-2011. Apesar dos valores apresentados serem diferentes, ao compará-los entre si foi possível perceber que eles apresentaram um mesmo padrão, e todos os métodos começaram com o valor de R_0 crescendo nos primeiros conjuntos de semanas, e tendo seu valor máximo entre 18 e 20 semanas. Após esse valor máximo de R_0 , em todos os métodos o valor de R_0 decresceu até a última semana do conjunto de dados.

Mais especificamente notou-se que para o período de 2009-2010 pelos métodos Nishiura (2010), e White e Pagano (2008), para dados suavizados, o maior valor de R_0 ocorreu com 20 semanas, e para os dados não suavizados, ocorreu com 19 semanas e para Ross-Macdonald (Macdonald, 1952) ocorreu com 20 semanas. Já para o período de 2010-2011 pelos métodos Nishiura (2010), e White e Pagano (2008), para dados suavizados, o maior valor de R_0 ocorreu com 19 semanas, e para os dados não suavizados, ocorreu com 18 semanas e para Ross-Macdonald (Macdonald, 1952) ocorreu com 19 semanas.

O cálculo semana a semana, também foi interessante quando olhou-se para o momento em que a doença está em curso. Neste caso, todos os métodos anteviram já nos conjuntos das semanas iniciais estudadas que haveria um surto, o que é importante para alertar a população em relação a prevenção da doença, neste caso específico da dengue, com os dados das primeiras semanas do ano já

seria possível informar a população de que deveriam se prevenir, usar repelentes, eliminar criadouros, etc.

Ainda não existe uma vacina completamente eficiente em relação a dengue, mas os estudos já estão bem avançados nessa área. Em relação à proteção da população através da vacinação quanto maior a cobertura melhor, desde de que a vacina não tenha efeitos adversos, é claro. Portanto, os cálculos para calcular qual é a proporção mínima da população a ser vacinada, devem ser feitos com base na expressão de Ross-Macdonald (Macdonald, 1952).

Como já mencionado anteriormente, Massad et al. (2010) analisam o R_0 proposto por Ross-Macdonald (Macdonald, 1952), e demonstram que se $R_0 > 1$ a curva inicial de crescimento da dengue tem uma forma exponencial, do contrário essa curva é côncava para baixo e tende a zero ao longo do tempo. Foi observado nos resultados apresentados neste trabalho, para o cálculo do R_0 proposto por Ross-Macdonald (Macdonald, 1952), que em ambos os casos, obteve-se uma situação em que a curva de crescimento dos dados teve formato exponencial, resultando em $R_0 > 1$.

Um bom ajuste de curvas, depende das variáveis envolvidas na simulação numérica. No entanto, quanto mais parâmetros envolvidos, mais custoso é esse procedimento computacionalmente, e mesmo assim isso não implica que o ajuste será perfeito. Portanto, as variáveis escolhidas foram as mais “imprevisíveis”, e os ajustes calculados neste estudo foram bons, visto que para estes conjuntos de dados o ajuste perfeito só poderia ser alcançado para os parâmetros escolhidos, se estes não fossem mantidos em uma variação biológica coerente com a realidade.

Tanto Nishiura (2010), quanto White e Pagano (2008), aplicam suas técnicas para doenças de transmissão direta. Nishiura (2010), aplica seu método para dados HIV/AIDS em países Europeus, e ele compara seu método a uma

outra técnica dada por Wallinga e Lipsitch (2007), no qual a ideia de crescimento exponencial está implícita em seu cálculo, e obtém resultados muito próximos. No entanto, em sua discussão, Nishiura (2010) observa, que um outro estudo sobre o valor do R_0 para HIV/AIDS, calculado para os EUA, tem um valor bem acima do calculado por seu estudo, com R_0 variando entre 13,9-54,5, enquanto que, em seu trabalho R_0 variou entre 3,59-3,74. Porém, neste estudo, foi utilizado um tempo médio de infecciosidade bem maior do que o usado por Nishiura. Ele diz que em seu estudo, o valor de R_0 menor, se dá em virtude do uso de um cálculo mais preciso do intervalo de geração. A respeito de desafios futuros, Nishiura (2010), defende seu trabalho dizendo que ao menos este estudo satisfaz a necessidade de um estudo com uma aproximação para calcular R_0 baseada em verossimilhança.

É válido ressaltar que na maioria dos estudos sobre o R_0 , está implícita a ideia de que $R_0 \geq 1$, pois dinâmicas que não apresentam esta característica, são de pouco interesse. Neste estudo, tem-se que o método proposto por Nishiura (2010), traz esta ideia explícita em sua formulação. Uma vez que, ele considera que a probabilidade de um indivíduo se tornar infectado é de $\frac{1}{R_0}$, o que implica que o valor mínimo assumido pelo R_0 , neste caso, é 1.

White e Pagano (2008), restringem seu problema a doenças de transmissão direta, pois eles dizem que para outros tipos de doença, o cálculo do intervalo de geração pode ser incorreto, o que acarretaria em um erro no cálculo do valor de R_0 , e eles afirmam que dessa forma, eles obtêm um sistema fechado. Neste estudo, o cálculo do R_0 foi feito para a cidade de Ribeirão Preto, em que foi suposto um sistema fechado. Foram impostos limites físicos para o intervalo de geração, lembrando que o intervalo de geração para doenças causadas por vetores, inclui os períodos intrínseco e extrínseco da doença, ou seja, inclui, humanos e vetores.

White e Pagano (2008), fazem um estudo com dados simulados para

a equação (22), em que eles obtêm que um valor errado do intervalo de geração, pode acarretar em um cálculo errado de R_0 , o que em seus estudos leva a um valor superestimado. Quando utilizam a verossimilhança para calcular R_0 , e intervalo de geração, desse mesmo conjunto de dados, eles obtêm bons valores de ambos, ou seja, valores coerentes, o que comprova que o método é bem posto, tanto para a equação (22), quanto para a verossimilhança (23). White e Pagano (2008), fazem cálculos para situações reais utilizando o método da verossimilhança (23), em que estudam Ebola no Congo, e H7N7, e Gripe Suína na Holanda. Eles mostram que, o valor calculado por eles, para Ebola é próximo a outro estudo em que foi modelado um sistema determinístico SEIR para Ebola. O mesmo acontece para influenza, em que os resultados obtidos em seu estudo tanto para R_0 , quanto para o intervalo de geração, são coerentes com estudos anteriores.

Para os cálculos apresentados neste trabalho, obteve-se que os valores de R_0 , calculados tanto para o método de Nishiura (2010), quanto para o método de White e Pagano (2008), são consistentes entre si, o que já era esperado, pois, ao calcular analiticamente o negativo do logartimo da expressão (21), foi obtida a expressão (22), o que significa que, mesmo utilizando distribuições diferentes, e caminhos diferentes para o cálculo do R_0 , eles possuem o mesmo estimador de máximo verossimilhança, para intervalo de geração fixado.

Para o intervalo de geração, foram utilizados dados calculados em campo para Porto Rico, e Tailândia (Aldstadt, 2007; Aldstadt et al., 2012). A literatura sobre o valor do intervalo de geração da dengue é escassa, no entanto, tem-se por definição que o intervalo de geração da dengue é a soma dos períodos extrínseco e intrínseco da dengue (Aldstadt et al., 2012), o que corrobora os valores utilizados nas simulações.

O trabalho publicado por Li et al. (2011), discute as falhas no cálculo do R_0 , em que eles estudam um modelo simples para malária, de onde são

obtidas várias expressões para o cálculo do R_0 . Mais precisamente, eles utilizam o operador de próxima geração, o termo constante do polinômio característico, uma expressão exponencial, a matriz jacobiana e a fórmula do termo constante do polinômio característico onde os autores adicionam e subtraem 9. Os métodos coincidem apenas no limiar ($R_0 = 1$). Em todos os outros casos, eles apresentam valores diferentes tanto para $R_0 > 1$, quanto para $R_0 < 1$. Contudo, todos foram concordantes na predição da evolução da doença.

Li et al. (2011), discutem vários casos em que os valores calculados para R_0 são errados, pelo menos no sentido de que eles não dizem ao certo o número de casos secundários gerados na população. No entanto, eles concluem que, apesar de todos os erros que podem estar implícitos nas várias formas de calcular R_0 , esta é a única ferramenta que existe com este propósito.

“Não existe outro conceito tão geral que possa ser entendido em termos de modelos compartimentais, modelos de rede, equações diferenciais parciais, e modelos de metapopulação. O número de infecções secundárias, tem um apelo intuitivo que ultrapassa em duração até mesmo a imprecisão de sua afirmação, quando aplicado ao conceito.” Li et al. (2011).

Com base nos resultados obtidos com os cálculos aqui apresentados, pôde-se concluir, como já dito anteriormente, que todos os métodos estavam corretos em prever que haveria uma epidemia de dengue. No entanto, não foi possível concluir, se alguns dos métodos diz corretamente qual seria o número de infecções secundárias causadas enquanto a doença ainda está em curso.

As limitações dos modelos foram em geral as mesmas, se não houver acesso aos dados reais, não existe o cálculo do valor do R_0 . No caso dos estudos propostos por Nishiura (2010), e White e Pagano (2008), ainda seriam necessárias informações mais precisas sobre o intervalo de geração. Um estudo feito no Brasil, por exemplo, apesar de os intervalos utilizados serem biologicamente coerentes,

se não houver qualquer informação sobre a sua duração, o cálculo pode ter um resultado completamente diferente da realidade.

Uma dificuldade encontrada é em relação à cidade/região estudada. O cálculo para o valor do R_0 da dengue, em uma cidade como São Paulo, por exemplo, é complicado pois a área da cidade é muito grande, e o que se tem na realidade são vários “patches” de dengue, cada um com sua dinâmica própria.

Outro problema encontrado, é delimitar de forma correta o final da fase inicial de crescimento da doença. Apesar desse problema ter sido resolvido neste trabalho, utilizando o cálculo do ponto de inflexão, esta escolha ainda é difícil pois, neste caso, se escolhermos uma função diferente da utilizada neste trabalho, os resultados podem ser diferentes. Foram feitos testes com a função de Fourier, mas seus resultados não foram satisfatórios, visto que, para o período em 2010-2011 foram encontrados 2 pontos de inflexão, o que portanto tornava impossível uma escolha de qual seria o melhor ponto.

Ao observar somente os valores calculados semana a semana, foi possível notar que o desaceleramento do valor de R_0 , ocorre entre 19 e 20 semanas para 2009-2010, e entre 18 e 19 semanas para 2010-2011, o que não coincide com o ponto de inflexão calculado. Para o período de 2009-2010, obteve-se que o ponto de inflexão aconteceu com 21 semanas, o qual é próximo do observado nos cálculos semanais, já para o período de 2010-2011 não aconteceu o mesmo, pois neste caso o ponto de inflexão aconteceu com 24 semanas. Este conjunto de dados teve um ajuste mais difícil, o que poderia explicar esta diferença.

Mesmo com resultados diferentes, os métodos previram que haveria um surto de dengue em ambos os períodos estudados, para a cidade de Ribeirão Preto, o que ajuda na prevenção da dengue. Um dos objetivos de se calcular o valor do R_0 , pela fase inicial de uma doença, e também todos os métodos puderam ser utilizados para a dengue enquanto essa doença ainda está em curso.

Os resultados apresentados pelos métodos, não podem ser simplesmente generalizados para descrever o comportamento de uma epidemia/endemia para um período completo, uma vez que, os cálculos foram feitos para os casos em que a doença ainda estava crescendo. No caso de Ross-Macdonald (Macdonald, 1952), é possível fazer um ajuste para o período completo de estudo. Nos estudos de Nishiura (2010), e White e Pagano (2008), os métodos foram desenvolvidos especialmente para a fase inicial da doença. No caso da expressão dada por Nishiura (2010), não é possível fazer os cálculos quando o número de infectados decresce, pois surgem valores negativos no termo $j_t - \sum_{s=0}^{\infty} g_s j_{t-s}$ da expressão (21), o que impossibilita seus cálculos. Segundo White e Pagano (2008), sua metodologia pode ser estendida para estudos mais gerais. Um estudo futuro seria tentar propor uma modificação ou complementação desses métodos a fim de generalizá-los um período completo, e analisar as implicações no resultados que o ajuste completo gera.

A proposta deste trabalho foi estudar métodos diferentes para o cálculo do R_0 , a fim de compará-los através da fase inicial da doença. Mesmo com as aplicações feitas neste trabalho, ele não deixa de ser um estudo teórico, em que sua finalidade foi aprofundar os estudos em cálculo do R_0 , utilizando dados reais de dengue, onde foram comparados métodos específicos e gerais para seu cálculo.

5 CONCLUSÃO

Neste trabalho foram estudadas técnicas para o cálculo do R_0 , para doenças causadas por vetores, sem supor crescimento exponencial, com o objetivo de compará-las através da fase inicial da doença, onde foram utilizados dados reais de dengue.

Foram estudados os métodos propostos por: Ross-Macdonald (Macdonald, 1952), Nishiura (2010), e White e Pagano (2008). Em que Ross-Macdonald (Macdonald, 1952), é um método específico para doenças causadas por vetores, e Nishiura (2010), e White e Pagano (2008), são métodos gerais para o cálculo do R_0 .

Comparar técnicas tão diferentes é uma tarefa difícil, pois é impossível afirmar que uma seja melhor que a outra, já que não existe um padrão ouro para o R_0 . No entanto, foi possível concluir que:

- Todos os métodos acertaram a previsão de que haveria um surto epidêmico, em ambos os períodos estudados;
- apesar de serem numericamente diferentes todos os métodos utilizados para o cálculo de R_0 , apresentaram um mesmo padrão ao longo dos cálculos semanais;
- o valor do R_0 proposto por Ross-Macdonald (Macdonald, 1952) foi sistematicamente maior do que o valor do R_0 calculados por Nishiura (2010), e White e Pagano (2008).

Em relação a implementação computacional todos os métodos exigem conhecimento de programação; no caso de Ross-Macdonald (Macdonald, 1952) é necessário conhecer algum software que realize ajustes, ou saber implementá-los o que nem sempre é uma tarefa fácil ou rápida; no caso de Nishiura

(2010), e White e Pagano (2008), são necessários alguns conhecimentos estatísticos e computacionais para poder implementar a verossimilhança da Distribuição Binomial e da Distribuição de Poisson.

Anexo 1

Tabela 4: Dados de dengue na cidade de Ribeirão Preto
no período de 2009-2010(CVE, 2012)

Semana	Número de Infectados
1	1
2	2
3	5
4	3
5	2
6	4
7	3
8	11
9	2
10	12
11	19
12	18
13	23
14	36
15	59
16	67
17	191
18	349
19	436
20	516
21	619

continua

Tabela 4 – continuação

Semana	Número de Infectados
22	1019
23	1131
24	1426
25	1601
26	1856
27	1920
28	1909
29	2119
30	2280
31	2061
32	1943
33	2273
34	2005
35	1362
36	851
37	652
38	377
39	310
40	155
41	112
42	78
43	43
44	50
45	24

continua

Tabela 4 – continuação

Semana	Número de Infectados
46	22
47	12
48	17
49	17
50	17
51	5
52	14
53	9
54	10
55	3
56	1
57	11
58	7
59	7
60	10
61	13
62	14
63	15
64	21
65	26
66	45
67	50
68	65

Tabela 5: Dados de dengue na cidade de Ribeirão Preto
no período de 2010-2011(CVE, 2012)

Semana	Número de Infectados
1	1
2	11
3	7
4	7
5	10
6	13
7	14
8	15
9	21
10	26
11	45
12	50
13	65
14	119
15	143
16	249
17	294
18	608
19	681
20	549
21	647
22	1025

continua

Tabela 5 – continuação

Semana	Número de Infectados
23	827
24	1316
25	1471
26	1254
27	2117
28	1633
29	1924
30	2029
31	1580
32	1044
33	1037
34	559
35	342
36	239
37	132
38	64
39	45
40	36
41	16
42	13
43	11
44	9
45	4
46	4

continua

Tabela 5 – continuação

Semana	Número de Infectados
47	5
48	5
49	8
50	7
51	21
52	10
53	10
54	3
55	2
56	3
57	2
58	2
59	2
60	2
61	1
62	1
63	1
64	5
65	23

Como dito anteriormente, a semana 1 não corresponde a primeira semana do ano.

No período de 2009-2010, tem-se que a semana 1 corresponde a 37^a semana do ano de 2009, e a semana 68 corresponde a 52^a semana de 2010.

No período de 2010-2011, tem-se que a semana 1 corresponde a 40^a semana do ano de 2010, e a semana 65 corresponde a 52^a semana de 2011.

6 REFERÊNCIAS

Aldstadt J. An incremental Knox test for the determination of the serial interval between successive cases of an infectious disease. *Stoch Environ Res Risk Assess*; 2007, 21, 487–500.

Aldstadt J, Yoon I, Tannitisupawong D, Jarman R. G, Thomas S. J, Gibbons R. V, Uppapong A, Iamsirithaworn S, Rothman A. L, Scott T. W, Endy T. Space-time analysis of hospitalised dengue patients in rural Thailand reveals important temporal intervals in the pattern of dengue virus transmission. *Trop Med Int Health*; 2012, 17, 1076–85.

Anderson R. M, May R. M. Population biology of infectious diseases: Part I. *Nature*; 1979a, 280, 361–7.

Anderson R. M, May R. M. Population biology of infectious diseases: Part II. *Nature*; 1979b, 280, 455–61.

Anderson R. M, May R. M. Directly transmitted infectious: Control by vaccination. *Science*; 1982, 215, 1053–60.

Anderson R. M, May R. M. *Infectious diseases of humans - Dynamics and Control*. New York: Oxford Science Publications; 1991.

Aron J. L, May R. M. 5-The population dynamics of malaria. In: Anderson R. M [editor]. *Population dynamics of infectious diseases: Theory and applications*. New York: Chapman and Hall; 1982. p.139–79.

Azevedo C. A. V, Merkley G. P, Walker W. R. Otimização não-linear na solução do problema inverso da irrigação por sulcos em tempo real. *Rev Brasileira de Engenharia Agrícola e Ambiental*; 1997, I, 43–9.

Barreto M. L, Teixeira M. G. Dengue in Brazil: epidemiological situation and contribution to a research. *Estud av*; 2008, 22, 53–72.

Bennett S. N, Holmes E. C, Chirivella M, Rodriguez D. M, Beltran M, Vorndam V, Gubler D. J, McMillan W. O. Selection-driven evolution of emergent dengue virus. *Mol Biol Evol*; 2003, 20, 1650–8.

Bhatt S, Gething P. W, Brady O. J, Messina J. P, Farlow A. W, Moyes C. L, Drake J. M, Brownstein J. S, Hoen A. G, Sankoh O, Myers M. F, George D. B, Jaenisch T, Wint G. R. W, Simmons C. P, Scott T. W, Farrar J. J, Hay S. I. The global distribution and burden of dengue. *Nature*; 2013, 496, 504–507.

Boyce W. E, DiPrima R. C. *Equações diferenciais e problemas de valores de contorno*. 7. ed. Rio de Janeiro: LTC; 2002.

Brauer F, Castillo-Chavez C. *Mathematical Models in Populations Biology and Epidemiology*. 2nd. ed. New York: Springer; 2012.

Brauer F, van den Driessche P, Wu J. *Mathematical Epidemiology*. New York: Springer; 2008.

Byrd R. H, Lu P, Nocedal J, Zhu C. A limited memory algorithm for bound constrained optimization. Rel. téc., Northwestern University-Department of Electrical Engi; 1994.

Chowell G, Clark G. G. Dengue/Dengue Hemorrhagic Fever: The Emergence of a Global Health Problem. *Emerg Infect Dis*; 1995, 1, 55–7.

Chowell G, Nishiura H. Quantifying the transmission potential of pandemic influenza. *Physics of Life Reviews*; 2008, 5, 50–77.

Cummings D. A. T, Lesser J. 6-Infectious diseases dynamics. In: Nelson K. E,

Williams C. M [editor]. *Infectious Disease Epidemiology: Theory and Practice*. 3rd. ed. Burglinton: Jones & Bartlett Learning; 2013. p.131-66.

da Cunha R. V, Nogueira R. M. R. 151 Dengue e dengue hemorrágico. In: Coura J. R [editor]. *Dinâmica das Doenças Infecciosas e Parasitárias*. Rio de Janeiro: Guanabara Koogan; 2005. p.1767-81.

CVE. Centro de Vigilância Epidemiológica. Disponível em:

http://www.cve.saude.sp.gov.br/htm/cve_dengue.html; 2012.

CVE. Centro de Vigilância Epidemiológica- Informe Técnico: Dengue.

Disponível em:

http://www.cve.saude.sp.gov.br/htm/zoo/dengue_inf2103.htm; Acesso em 10 de maio de 2015.

Nature. Dengue fever climbs the social ladder. *Nature*; 2007, 448, 734–5.

Diekmann O, Heesterbeek J. A. J, Metz J. A. J. On the definition and the computation of the basic ratio R_0 in models for infectious diseases in heterogeneous populations. *J Math Biol*; 1980, 28, 365–82.

Dietz K. Transmission and control of arbovirus diseases. *Epidemiol*; 1975, 104–21.

van den Driessche P, Watmough J. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Math Biosci*; 2002, 180, 29–48.

Dublin L. I, Lotka A. J. On the true rate of natural increase. *Journal of the American Statistical Association*; 1925, 20, 305–39.

FAPESP A. CTNBio aprova vacina contra dengue do Butantan. Disponível em:

http://agencia.fapesp.br/ctnbio_aprova_vacina_contra_dengue_do_butantan/21658/; Acesso em 19 de setembro de 2015.

Favier C, Degallier N, Rosa-Freitas M. G, Boulanger J. P, Lima J. R. C, Luitgards-Moura J. F, Menke's C. E, Mondet B, Oliveira C, Weimann E. T. S, Tsouris P. Early determination of the reproductive number for vector-borne diseases: the case of dengue in Brazil. *Trop Med Int Health*; 2006, 11, 332–40.

Fine P. E. M. The interval between successive cases of an infectious disease. *Am J Epidemiol*; 2003, 158, 1039–47.

Gubler D. J. Dengue and Dengue Hemorrhagic Fever. *Clin Microbiol Rev*; 1998, 11, 480–96.

Gubler D. J, Clark G. G. Dengue/Dengue Hemorrhagic Fever: The Emergence of a Global Health Problem. *Emerg Infect Dis*; 1995, 1, 55–7.

Gubler D. J, Kuno G. *Dengue and dengue hemorrhagic fever*. Colorado: CAB International; 1997.

Guidorizzi H. L. *Um curso de cálculo- Volume 1*. 5. ed. Rio de Janeiro: LTC; 2001.

Guzman A, Istúriz R. E. Update on the global spread of dengue. *Int J Antimicrob Agents*; 2010, 36S, S40–2.

Heesterbeek H. R_0 [tese]. Leiden : University of Leiden; 1992.

Heesterbeek J. A. P. A brief history of R_0 and a recipe for its calculation. *Acta Biotheor*; 2002, 50, 189–204.

Hethcote H. W. The Mathematics of Infectious Diseases. *SIAM Review*; 2000, 42, 599–653.

Holmes E. C, Bartley L. M, Garnett G. P. 10. The emergence of Dengue; past, present, and future. In: Krause R. M [editor]. *Emerging Infections Biomedical Research Reports*. New York: Academic Press; 1990. p.301–23.

Kermack W. O, McKendrick A. G. A contribution to the mathematical theory of epidemics. *Proc R Soc Lond A Math Phys Eng Sci*; 1927, 115, 700–21.

Latorre M. R. D. O. Câncer em Goiânia: análise da incidência e da mortalidade no período de 1988 a 1997 [tese livre docência]. São Paulo : Faculdade de Saúde Pública da Universidade de São Paulo; 2001.

Li J, Blakeley D, Smith? R. The failure of R_0 . *Comput Math Methods Med*; 2011, 2011, 1–17.

Lopez L. F, Coutinho F. A. B, Burattini M. N, Massad E. Threshold conditions for infection persistence in complex host-vectors interactions. *C R Biol*; 2002, 325, 1073–84.

Macdonald G. The analysis of equilibrium in malaria. *Trop Dis Bull*; 1952, 49, 813–29.

Mandal S, Sarkar R. R, Sinha S. Mathematical models of malaria - a review. *Malar J*; 2011, 10, 1–19.

Marques C. A, Forattini O. P, Massad E. The basic reproduction number for dengue fever in São Paulo state, Brazil:1990-1991 epidemic. *Trans R Soc Trop Med Hyg*; 1994, 88, 58–9.

Martelli C. M. T, Siqueira J. B, Parente M. P. P. D, Zara A. L. S. A, Oliveira C. S, Braga C, Pimenta F. G, Cortes F, Lopez J. G, Bahia L. R, Mendes M. C. O, da Rosa M. Q. M, Siqueira N. T. F, Constenla D, de Souza W. V. Economic Impact of Dengue: Multicenter Study across Four Brazilian Regions. *PLoS Neglected Tropical Diseases*; 2015, 1–19.

Martín J. L. S, Brathwaite O, Zambrano B, Solórzano J. O, Bouckennooghe A, Dayan G. H, Guzmán M. G. The Epidemiology of Dengue in the Americas Over the Last Three Decades: A Worrisome Reality. *Am J Trop Med Hyg*; 2010, 82, 128–35.

Massad E. Transmission rates and the evolution of pathogenicity. *Evolution*; 1987, 41, 1127–30.

Massad E. Mathematical and computer models in epidemiology. *MD Comput*; 1996, 13, 147–54.

Massad E, Burattini M. N, Azevedo-Neto R. S, Yang H. M, Coutinho F. A. B, Zanetta D. M. T. A model-based design of a vaccination strategy against rubella in a non-immunized community of São Paulo State, Brazil. *Epidemiol Infect*; 1994, 112, 579–94.

Massad E, Coutinho F. A. B. Dimensional analysis of vector-borne infections equations. *Mem Inst Oswaldo Cruz*; 2012, 107, 564–67.

Massad E, Coutinho F. A. B, Burattini M. N, Lopez L. F. The risk of yellow fever in a dengue-infested area. *Transactions of the royal society of tropical medicine and hygiene*; 2001, 95, 370–4.

Massad E. M, Coutinho F. A. B, Burattini M. N, Amaku M. Estimation of R_0 from the initial phase of an outbreak of a vector-borne infection. *Trop Med Int Health*; 2010, 15, 120–6.

McBride W. J, Bielefeldt-Ohmann H. Dengue viral infections; pathogenesis and epidemiology. *Microbes Infect*; 2000, 2, 1041–50.

Mendonça M. W. O Método L-BFGS com Fatoração Incompleta para a Resolução de Problemas de Minimização [dissertação]. Florianópolis : Universidade Federal de Santa Catarina; 2005.

Ministério da Saúde. Ministério da Saúde -Boletim Epidemiológico. Secretaria de Vigilância em Saúde. Ministério da Saúde; 2015.

Murray N. E. A, Quam M. B, Smith A. W. Epidemiology of dengue: past, present and future prospects. *Clin Epidemiol*; 2013, 5, 299–309.

Nishiura H. Mathematical and statistical analyses of the spread of dengue. *Dengue Bulletin*; 2006, 30, 51–67.

Nishiura H. Correcting the Actual reproduction Number: A Simple Method to Estimate R_0 from Early Epidemic Growth Data. *Int J Environ Res Public Health*; 2010, 7, 291–302.

Ooi E. E, Gubler D. J. Dengue in Southeast Asia: epidemiological characteristics and strategic challenges in disease prevention. *Cad Saude Publica*; 2008, 25, S115–24.

PAHO. Number of Reported Cases - Organização Pan-Americana de Saúde. Disponível em: http://www.paho.org/hq/index.php?option=com_topics&view=readall&cid=2158&Itemid=40734&lang=ptl; 2015.

Roberts M. G, Nishiura H. Early Estimation of the reproduction number in the presence of imported cases: pandemic influenza H1N1-2009 in New Zealand. *PloS One*; 2011, 6, 1–9.

Rodrigues M. M, Marques G. R. A. M, Serpa L. L. N, Arduino M. B, Voltolini J. C, Barbosa G. L, Andrade V. R, Lima V. L. C. Density of *Aedes aegypti* and *Aedes albopictus* and its association with number of residents and meteorological variables in the home environment of dengue endemic area, São Paulo, Brazil. *Parasites & Vectors*; 2015, 8, 1–9.

Rodriguez-Barraquer I, Cordeiro M. T, Braga C, Souza W. V, Marques E. T,

Cummings D. A. T. From re-emergence to hyperendemicity: the natural history of the dengue epidemic in Brazil. *PLoS Neglected Tropical Diseases*; 2008, 5, 1–6.

Ross R. *The prevention of Malaria*. 2nd. ed. London: John Murray; ????

Ross R. An Application of the Theory of Probabilities to the Study of a priori Pathometry. Part I. *Proc R Soc Lond A Math Phys Eng Sci*; 1916, 92, 204–30.

Ross R, Hudson H. P. An Application of the Theory of Probabilities to the Study of a priori Pathometry. Part II. *Proc R Soc Lond A Math Phys Eng Sci*; 1917a, 93, 212–25.

Ross R, Hudson H. P. An Application of the Theory of Probabilities to the Study of a priori Pathometry. Part II. *Proc R Soc Lond A Math Phys Eng Sci*; 1917b, 93, 225–40.

Sang R. C. Dengue in Africa. In: *Scientific Working Group Report on Dengue*; 2006.

Scott T. W, Morrison A. C. Vector Dynamics and Transmission of Dengue Virus: Implications for Dengue Surveillance and Prevention Strategies: Vector Dynamics and Dengue Prevention. *Curr Top Microbiol Immunol*; 2010, 338, 115–28.

Simmons C. P, Farrar J. J, Chau N. V. V, Wills B. Dengue current concepts. *N Engl J Med*; 2012, 366, 1423–32.

Smith? R. *Modelling disease ecology with Mathematics*. Springfield: American Institute of Mathematical Sciences; 2008.

Spiegel M. R. *Estatística*. São Paulo: McGraw-Hill do Brasil, LTDA; 1976.

Teixeira M. G, Barreto M. L. Diagnosis and management of dengue. *BMJ*; 2009, 339, 1189–93.

Teixeira M. G, Barreto M. L, Guerra Z. Epidemiologia e Medidas de Prevenção do Dengue. *Informe Epidemiológico do SUS*; 1999, 8, 5–33.

Toledo K. Aprovação da vacina da dengue pode ser antecipada em dois anos. Disponível em: http://agencia.fapesp.br/aprovacao_da_vacina_da_dengue_pode_ser_antecipada_em_dois_anos/20882/; Acesso em 19 de setembro de 2015.

WHO. *Dengue guidelines for diagnosis, treatment, prevention and control*. WHO-World Health Organization; 2009.

WHO. Vector-Borne Diseases -Dengue. Available from: <http://www.who.int/mediacentre/factsheets/fs387/en/index2.html>; Acesso em 19 de setembro de 2015.

Wallinga J, Lipsitch M. How generation intervals shape the relationship between growth rates and reproductive numbers. *Proc R Soc Lond B Biol Sci*; 2007, 274, n.1609, 599–604.

White L. F, Pagano M. A likelihood-based method for real-time estimation of the serial interval and reproductive number of an epidemic. *Stat Med*; 2008, 27, 2999–3016.

White P. J, Ward H, Garnett G. P. Is HIV out of control in the UK? An example of analysing patterns of HIV spreading using incidence-to-prevalence ratios. *AIDS*; 2006, 20, 1898–1901.

Wu J. Y, Lun Z. R, James A. A, Chen X. G. Review: Dengue Fever in Mainland China. *Am J Trop Med Hyg*; 2010, 83, 664–71.

Yang H. M. The basic reproduction number obtained from Jacobian and next generation matrices? A case study of dengue transmission modelling. *Biosystems*; 2014, 126, 52–75.

APÊNDICES

Apêndice A

Código em R, e algoritmo, para o cálculo da minimização do logaritmo negativo da equação (21) utilizando a Distribuição Binomial em conjunto com a normal.

```
rm(list=ls())
f2<-function (p) {
xm<-2.35
var<-p[1]
r0<-p[2]
alpha=(xm*xm)/var
beta=xm/(var)
pxvec<-c(6.5,8.75,9.25,11,13,15.75,19,26.75,35.5,46.5,69.75)
pn<-length(pxvec)
k=8
ps<-rep(0,k)

for(i in 1:k){
ps[i] <- (pgamma(i,shape=alpha,rate=beta)-pgamma(i-1,shape=alpha,rate=beta))
}

ps1 <- ps/sum(ps)

vector<-rep(0,pn)
for(t in 1:pn){
for(s in 1:min(t,k)){
```

```

if((t-s)<=0){
vector[t]=0+vector[t]
}
if((t-s)>0){
vector[t]=(ps1[s]*pxvec[t-s])+vector[t]
}
}
}

pr<-rep(0,pn)
for(ii in 1:pn){
if(vector[ii]>0){
if(pxvec[ii]< 100){
pr[ii]=-(log(factorial(pxvec[ii]))/(factorial(vector[ii])*
factorial(pxvec[ii]-vector[ii])))+(vector[ii]*log(1/r0))
+((pxvec[ii]-vector[ii])*log(1-(1/r0))))
}
if(pxvec[ii ]>= 100){
pr[ii]=-dnorm(vector[ii],pxvec[ii]*(1/r0),
sqrt(pxvec[ii]*(1/r0)*(1-(1/r0))),log=TRUE)
}
}
}

soma=sum(pr)
}

optim(par=c(0.4,2), fn=f2,method="L-BFGS-B",lower=c(0.28,1.2),upper=c(0.43,10))

```

Algoritmo "calculo de verossimilhança: binomial + normal"

Var

```
// Seção de Declarações das variáveis
pn,i,ll,posicao1,vl,K,jj,m,n,cont,a,q,qq,t,tt,ii,s : inteiro
xm,blal,soma: real
alpha : vetor[1..16] de inteiro // vetor de alpha
beta : vetor[1..16] de inteiro // vetor de beta
vari : vetor[1..16] de real // vetor da variância
vector: vetor[1..N] de real// vetor de somatorio
ps: vetor[1..K] de real // vetor da distribuição gama
psl: vetor[1..K] de real // vetor da distribuição gama normalizado
pr:vetor[1..N]de real// vetor das probabilidades
r : vetor[1..14096] de inteiro // somatória da probabilidade
rr : vetor[1..14096] de inteiro // guarda a variancia da distribuição gama
rrrr: vetor[1..14096] de inteiro // guarda r0
pxvec : vetor[1..N] de real //dados reais suavizados
```

Inicio

```
// Seção de Comandos, procedimento, funções, operadores, etc...
```

```
pn <- N // tamanho do vetor de dados
xm <- 2.35// média da distribuição gama
K <- 8 // valor do intervalo máximo de intervalo de geração
jj <- 1 // índice
```

```
// Atribuição de valores para a variância
cont:=1 //contador
PARA n DE 0.28 ATÉ 0.43 PASSO 0.01 FAÇA
    vari[cont] := n
    cont:=cont+1
FIMPARA
    vl:=cont-1// tamanho do vetor da variancia
// cálculo dos alphas e betas para média fixada e variância variando
PARA i DE 1 ATÉ vl FAÇA
    alpha[i] <- ((xm*xm)/vari[i])
    beta[i] <- (xm/(vari[i]))
FIMPARA
```

```
PARA m DE 1 ATÉ 14096 FAÇA
    r[m] := 0 // guarda a soma das probabilidades
    rr[m] := 0 // guarda a variancia
    rrrr[m] := 0 // guarda o valor do R0
FIMPARA
```

```
PARA ll DE 1 ATÉ vl FAÇA //variação da variância
    PARA r0 DE 1.2 ATÉ 10 PASSO 0.01 FAÇA // variação do R0
```

```
        PARA qq DE 1 ATÉ K FAÇA//limpando os vetores
            ps[qq] :=0
            psl[qq] := 0
        FIMPARA
        // intervalo de geração
        PARA q DE 1 ATÉ K FAÇA//cálculo dos valores do intervalo de geração
            //em que pgamma é a distribuição gama acumulada
            ps[q] <- (pgamma(q,alpha[ll],beta[ll]) - pgamma(q-1,alpha[ll],beta[ll]))
            psl[q] <- ps[q]/pgamma(K,alpha[ll],beta[ll])// normalização
        FIMPARA
```

```
        PARA tt DE 1 ATÉ pn FAÇA
            vector[tt] :=0
        FIMPARA
```

```
        PARA t DE 1 ATÉ pn FAÇA
            PARA s DE 1 ATÉ min(t,K) FAÇA // calculo do mínimo entre t e K
                LEIA(t)
                LEIA(s)
                SE ((t-s)<=0) ENTÃO
```


Apêndice B

Código em R, e algoritmo, para o cálculo da minimização do logaritmo negativo da equação (30).

```
rm(list=ls())

f2<-function (p) {

  xm<-2.35

  var<-0.43

  r0<-p

  alpha=(xm*xm)/var

  beta=xm/(var)

  pxvec<-c(6.5,8.75,9.25,11,13,15.75,19,26.75,35.5,46.5,69.75)

  pn<-length(pxvec)

  k=8

  ps<-rep(0,k)

  for(i in 1:k){
    ps[i] <- (pgamma(i,shape=alpha,rate=beta)-pgamma(i-1,shape=alpha,rate=beta))
  }

  ps1 <- ps/sum(ps)

  vector<-rep(0,pn)

  for(t in 1:pn){
```

```

for(s in 1:min(t,k)){
  if((t-s)<=0){
    vector[t]=0+vector[t]
  }
  if((t-s)>0){
    vector[t]=(ps1[s]*pxvec[t-s])+vector[t]
  }
}

pr<-rep(0,pn)
for(ii in 1:pn){
  if(vector[ii]>0){
    pr[ii]=-((vector[ii]*log(1/r0))+((pxvec[ii]-vector[ii])*log(1-(1/r0))))
  }
}

soma=sum(pr)

}

optim(par=2, fn=f2, method="Brent",lower=1,upper=10)

```

Algoritmo "calculo de verossimilhança:binomial com variância fixada"

Var

```
// Seção de Declarações das variáveis
pn,i,ll,posicao1,vl,K,jj,m,n,cont,a,q,qq,t,tt,ii,s : inteiro
xm,blal,soma, vari,alpha,beta: real
vector: vetor[1..N] de real// vetor de somatorio
ps: vetor[1..K] de real // vetor da distribuição gama
psl: vetor[1..K] de real // vetor da distribuição gama normalizado
pr:vetor[1..N]de real// vetor das probabilidades
r : vetor[1..901] de inteiro // somatória da probabilidade
rrrr: vetor[1..901] de inteiro // guarda r0
pxvec : vetor[1..N] de real //dados reais suavizados
```

Inicio

```
// Seção de Comandos, procedimento, funções, operadores, etc...
```

```
pn <- N // tamanho do vetor de dados
xm <- 2.35// média da distrbuição gama
K <- 8 // valor do intervalo máximo de intervalo de geração
jj <- 1 // índice
vari<-0.43//variancia fixada
// Atribuição de valores para a variância
```

```
// cálculo de alpha e beta para média e variância fixada
```

```
alpha <- ((xm*xm)/vari)
beta <- (xm/(vari))
```

```
PARA m DE 1 ATÉ 901 FAÇA
  r[m] := 0 // guarda a soma das probabilidades
  rrrr[m] := 0 // guarda o valor do R0
FIMPARA
```

```
PARA r0 DE 1.0 ATÉ 10 PASSO 0.01 FAÇA // variação do R0
```

```
PARA qq DE 1 ATÉ K FAÇA//limpando os vetores
  ps[qq] :=0
  psl[qq] := 0
FIMPARA
// intervalo de geração
PARA q DE 1 ATÉ K FAÇA//cálculo dos valores do intervalo de geração
  //em que pgamma é a distribuição gama acumulada
  ps[q] <- (pgamma(q,alpha,beta) - pgamma(q-1,alpha,beta))
  psl[q] <- ps[q]/pgamma(K,alpha,beta)// normalização
FIMPARA
```

```
PARA tt DE 1 ATÉ pn FAÇA
  vector[tt] :=0
FIMPARA
```

```
PARA t DE 1 ATÉ pn FAÇA
  PARA s DE 1 ATÉ min(t,K) FAÇA // calculo do mínimo entre t e K
    LEIA(t)
    LEIA(s)
    SE ((t-s)<=0) ENTÃO
      vector[t] = 0+vector[t]
    SENÃO
      SE ((t-s)>0) ENTÃO
        vector[t] = (psl[s]*pxvec[t-s])+vector[t]
      FIMSE
    FIMSE
  FIMPARA
FIMPARA
```

```
PARA m DE 1 ATÉ pn FAÇA
  pr[m] := 0
```


FIMPARA

```
PARA ii de 1 ATÉ pn FAÇA // binominal para variancia fixada
    pr[ii] = -((vector[ii]*log(1/r0))+((pxvec[ii]-
        vector[ii])*log(1-1/r0))))
```

FIMPARA

soma := 0

```
PARA ii DE 1 ATÉ pn FAÇA
    soma := soma+pr[ii]
```

FIMPARA

```
r[jj] <- soma
rrrr[jj] <- r0
jj := jj + 1
```

FIMPARA

bla1 := 10000000000000000000000000000000

PARA a DE 1 ATÉ 901 FAÇA // calculo de qual r0 que gerou a menor probabilidade

SE(r[a]<bla1) ENTÃO

bla1 := r[a] //guarda a menor soma

posicao1 := a // guarda a posição no vetor da menor soma

FIMSE

FIMPARA

Escreva("r0=",rrrr[posicao1])

Fimalgoritmo

Apêndice C

Código em R, e algoritmo, para o cálculo da minimização do logaritmo negativo da equação (23).

```
rm(list=ls())
f2<-function (lambda) {
xm<-2.35
var<-lambda[1]
r0<-lambda[2]
alpha=(xm*xm)/var
beta=xm/(var)

pxvec<-c(1,11,7,7,10,13,14,15,21,26,45,50,65,119)
pn<-length(pxvec)
k=8
ps<-rep(0,k)

for(i in 1:k){
ps[i] <- (pgamma(i,shape=alpha,rate=beta)-pgamma(i-1,shape=alpha,rate=beta))
}
ps1 <- ps/sum(ps)

vector<-rep(0,pn)
for(t in 1:pn){
for(s in 1:min(t,k)){
if((t-s)<=0){
vector[t]=0+vector[t]
}
}
```

```

if((t-s)>0){
vector[t]=(ps1[s]*pxvec[t-s])+vector[t]
}
}
}

pr<-rep(0,pn)
for(ii in 1:pn){
if(vector[ii]>0){
pr[ii]=-dpois(pxvec[ii],lambda=r0*vector[ii],log=TRUE)
}
}
soma=sum(pr)
}

optim(par=c(0.4,2), fn=f2,method="L-BFGS-B",lower=c(0.28,1),upper=c(0.43,10))

```

Algoritmo "calculo de verossimilhança:poisson"

Var

```
// Seção de Declarações das variáveis
pn,i,ll,posicao1,vl,K,jj,m,n,cont,a,q,qq,t,tt,ii,s : inteiro
xm,blal,soma: real
alpha : vetor[1..16] de inteiro // vetor de alpha
beta : vetor[1..16] de inteiro // vetor de beta
vari : vetor[1..16] de real // vetor da variância
vector: vetor[1..N] de real// vetor de somatorio
ps: vetor[1..K] de real // vetor da distribuição gama
psl: vetor[1..K] de real // vetor da distribuição gama normalizado
pr:vetor[1..N]de real// vetor das probabilidades
r : vetor[1..14416] de inteiro // somatória da probabilidade
rr : vetor[1..14416] de inteiro // guarda a variancia da distribuição gama
rrrr: vetor[1..14416] de inteiro // guarda r0
pxvec : vetor[1..N] de real //dados reais suavizados
```

Inicio

```
// Seção de Comandos, procedimento, funções, operadores, etc...
```

```
pn <- N // tamanho do vetor de dados
xm <- 2.35// média da distribuição gama
K <- 8 // valor do intervalo máximo de intervalo de geração
jj <- 1 // índice
```

```
// Atribuição de valores para a variância
cont:=1 //contador
PARA n DE 0.28 ATÉ 0.43 PASSO 0.01 FAÇA
    vari[cont] := n
    cont:=cont+1
FIMPARA
    vl:=cont-1// tamanho do vetor da variancia
// cálculo dos alphas e betas para média fixada e variância variando
PARA i DE 1 ATÉ vl FAÇA
    alpha[i] <- ((xm*xm)/vari[i])
    beta[i] <- (xm/(vari[i]))
FIMPARA
```

```
PARA m DE 1 ATÉ 14416 FAÇA
    r[m] := 0 // guarda a soma das probabilidades
    rr[m] := 0 // guarda a variancia
    rrrr[m] := 0 // guarda o valor do R0
FIMPARA
```

```
PARA ll DE 1 ATÉ vl FAÇA //variação da variância
    PARA r0 DE 1.2 ATÉ 10 PASSO 0.01 FAÇA // variação do R0
```

```
        PARA qq DE 1 ATÉ K FAÇA//limpando os vetores
            ps[qq] :=0
            psl[qq] := 0
        FIMPARA
            // intervalo de geração
        PARA q DE 1 ATÉ K FAÇA//cálculo dos valores do intervalo de geração
            //em que pgamma é a distribuição gama acumulada
            ps[q] <- (pgamma(q,alpha[ll],beta[ll]) - pgamma(q-1,alpha[ll],beta[ll]))
            psl[q] <- ps[q]/pgamma(K,alpha[ll],beta[ll])// normalização
        FIMPARA
```

```
        PARA tt DE 1 ATÉ pn FAÇA
            vector[tt] :=0
        FIMPARA
```

```
        PARA t DE 1 ATÉ pn FAÇA
            PARA s DE 1 ATÉ min(t,K) FAÇA // calculo do mínimo entre t e K
                LEIA(t)
                LEIA(s)
                SE ((t-s)<=0) ENTÃO
```


Apêndice D

Código em R, e algoritmo, para o cálculo da minimização do logaritmo negativo da equação (35).

```
rm(list=ls())
f2<-function (lambda) {
xm<-2.35
var<-lambda[1]
r0<-lambda[2]
alpha=(xm*xm)/var
beta=xm/(var)

pxvec<-c(6.5,8.75,9.25,11,13,15.75,19,26.75,35.5,46.5,69.75)
pn<-length(pxvec)
k=8
ps<-rep(0,k)

for(i in 1:k){
ps[i] <- (pgamma(i,shape=alpha,rate=beta)-pgamma(i-1,shape=alpha,rate=beta))
}
ps1 <- ps/sum(ps)

vector<-rep(0,pn)
for(t in 1:pn){
for(s in 1:min(t,k)){
if((t-s)<=0){
vector[t]=0+vector[t]
}
}
```

```

if((t-s)>0){
vector[t]=(ps1[s]*pxvec[t-s])+vector[t]
}
}
}

pr<-rep(0,pn)
for(ii in 1:pn){
if(vector[ii]>0){
pr[ii]=-(-r0*vector[ii]+pxvec[ii]*log(r0*vector[ii]))
}
}
soma=sum(pr)
}

optim(par=c(0.4,2), fn=f2,method="L-BFGS-B",lower=c(0.28,1),upper=c(0.43,10))

```

Algoritmo "calculo de verossimilhança:poisson com variância fixada"

Var

```
// Seção de Declarações das variáveis
pn,i,ll,posicao1,vl,K,jj,m,n,cont,a,q,qq,t,tt,ii,s : inteiro
xm,blal,soma, vari,alpha,beta: real
vector: vetor[1..N] de real// vetor de somatorio
ps: vetor[1..K] de real // vetor da distribuição gama
psl: vetor[1..K] de real // vetor da distribuição gama normalizado
pr:vetor[1..N]de real// vetor das probabilidades
r : vetor[1..901] de inteiro // somatória da probabilidade
rrrr: vetor[1..901] de inteiro // guarda r0
pxvec : vetor[1..N] de real //dados reais suavizados
```

Inicio

```
// Seção de Comandos, procedimento, funções, operadores, etc...
```

```
pn <- N // tamanho do vetor de dados
xm <- 2.35// média da distrbuição gama
K <- 8 // valor do intervalo máximo de intervalo de geração
jj <- 1 // índice
vari<-0.43 //variancia fixada
// Atribuição de valores para a variância
```

```
// cálculo de alpha e beta para média e variância fixada
```

```
alpha <- ((xm*xm)/vari)
beta <- (xm/(vari))
```

```
PARA m DE 1 ATÉ 901 FAÇA
  r[m] := 0 // guarda a soma das probabilidades
  rrrr[m] := 0 // guarda o valor do R0
FIMPARA
```

```
PARA r0 DE 1.0 ATÉ 10 PASSO 0.01 FAÇA // variação do R0
```

```
PARA qq DE 1 ATÉ K FAÇA//limpando os vetores
  ps[qq] :=0
  psl[qq] := 0
FIMPARA
  // intervalo de geração
PARA q DE 1 ATÉ K FAÇA//cálculo dos valores do intervalo de geração
  //em que pgamma é a distribuição gama acumulada
  ps[q] <- (pgamma(q,alpha,beta) - pgamma(q-1,alpha,beta))
  psl[q] <- ps[q]/pgamma(K,alpha,beta)// normalização
FIMPARA
```

```
PARA tt DE 1 ATÉ pn FAÇA
  vector[tt] :=0
FIMPARA
```

```
PARA t DE 1 ATÉ pn FAÇA
  PARA s DE 1 ATÉ min(t,K) FAÇA // calculo do mínimo entre t e K
    LEIA(t)
    LEIA(s)
    SE ((t-s)<=0) ENTÃO
      vector[t] = 0+vector[t]
    SENÃO
      SE ((t-s)>0) ENTÃO
        vector[t] = (psl[s]*pxvec[t-s])+vector[t]
      FIMSE
    FIMSE
  FIMPARA
FIMPARA
```

```
PARA m DE 1 ATÉ pn FAÇA
  pr[m] := 0
```


FIMPARA

```
    PARA ii de 1 ATÉ pn FAÇA // poisson para dados suavizados  
        pr[ii] = -(-r0*vector[ii]+pxvec[ii]*log(r0*vector[ii]))
```

FIMPARA

soma := 0

PARA ii DE 1 ATÉ pn FAÇA

soma := soma+pr[ii]

FIMPARA

r[jj] <- soma

rrrr[jj] <- r0

jj := jj + 1

FIMPARA

bla1 := 10000000000000000000000000000000

PARA a DE 1 ATÉ 901 FAÇA // calculo de qual r0 gerou a menor probabilidade

SE(r[a]<bla1) ENTÃO

bla1 := r[a] //guarda a menor soma

posicao1 := a // guarda a posição no vetor da menor soma

FIMSE

FIMPARA

Escreva("r0=",rrrr[posicao1])

Fimalgoritmo

Apêndice E

Código em R, e algoritmo, para a equação (22). O código em R foi cedido por Laura White, com uma pequena modificação para que a entrada de dos dados na distribuição Gamma fosse feita diretamente com a média e a variância desejadas.

```
getR0.WP <- function(N,pars,distn="gamma",k=8){
  ps <- rep(0,k)
  xm<- pars[1]
  var<-pars[2]
  if(distn=="gamma"){
    alpha <- (xm*xm)/var
    rate <- xm/(var)
    for(i in 1:k){
      ps[i] <- pgamma(i,shape=alpha,rate=rate)-pgamma(i-1,shape=alpha,rate=rate)
    }
  }
  if(distn=="exp"){
    for(i in 1:k){
      ps[i] <- pexp(i,pars)-pexp(i-1,pars)
    }
  }
  if(distn=="MN"){
    ps <- pars
  }
  ps1 <- ps/sum(ps)
  N <- N[1:max(which(N!=0))]
  r<- length(N)
```

```

N.long <- c(rep(0,(k-1)),N)
l <- length(N.long)
if(l > k){
  sumK <- 0
  for(i in k:(l-1)){
    sumK <- ps1%%N.long[i:(i-k+1)] + sumK
  }
  R0 <- sum(N)/sumK
} else R0 <- NA
return(R0))
}

```

Algoritmo "Cálculo do valor de R0"

Var

```
// Seção de Declarações das variáveis
i,j,k,l,m:inteiro
R0,somajt, xm,vari,alpha,beta, soma2:real
pxvec : vetor[1..N] de real //dados reais
vector: vetor[1..N] de real
ps : vetor[1..K] de real // vetor para armazenamento
psl : vetor[1..K] de real
```

Inicio

```
K=8 // valor máximo assumido pela função gama
xm:=2.35 //média do intervalo de geração
vari:=0.43 //variância do intervalo de geração
alpha:=(xm*xm)/vari //parâmetro de forma da função gama
beta:=xm/vari //parametro de escala da função gama
```

```
// Soma de todos os dados reais
```

```
somajt:=0
PARA i DE 1 ATÉ N FAÇA
somajt:= pxvec[i]+somajt
FIMPARA
```

```
// cálculo dos valores do intervalo de geração em que pgamma é a distribuição
```

```
//gama acumulada
PARA q DE 1 ATÉ K FAÇA
ps[q] <- (pgamma(q,alpha,beta) - pgamma(q-1,alpha,beta))
psl[q] <- ps[q]/pgamma(K,alpha,beta)
FIMPARA
```

```
// somatório total do intervalo de geração
```

```
PARA t DE 1 ATÉ N FAÇA
PARA s DE 1 ATÉ min(t,k) FAÇA // calculo do mínimo
LEIA(t)
LEIA(s)
SE ((t-s)<=0) ENTÃO
vector[t] = 0+vector[t]
SENÃO
SE ((t-s)>0) ENTÃO
vector[t] = (psl[s]*pxvec[t-s])+vector[t]
FIMSE
FIMSE
FIMPARA
FIMPARA
```

```
soma2:=0
```

```
PARA j DE 1 ATÉ N FAÇA
soma2= vector[j]+soma2
FIMPARA
```

```
R0=somajt/soma2/// calculo do R0
```

```
ESCREVA R0
```

Fimalgoritmo