

**A Matemática nas epidemias:
Introdução a alguns conceitos
básicos que podem nos ajudar a
entender como modelos
matemáticos fazem previsões e
avaliam eficácia de Possíveis
medidas de contenção. Parte I.**

Kleber Roberto Pirota

1. Motivação – Introdução

Estamos nos mês de abril de 2020 e ocorre uma grave pandemia de uma doença denominada de Covid-19 (*corona vírus disease – 19*) causada por uma nova cepa de um vírus do tipo Corona, o SARS-COV-2 (*severe acute respiratory syndrome coronavirus 2*), popularmente conhecido por “novo coronavírus”. Identificado em seres humanos pela primeira vez em dezembro de 2019 na cidade de Wuhan (China), no dia 20 de abril de 2020 já contabiliza ao redor de 2,5 milhões de casos confirmados com mais de 170 mil mortes. Ao estarmos no meio da pandemia, não é possível fornecer maiores dados introdutórios sobre a pandemia.

Devido ao imenso volume de informações produzido e à velocidade de sua disseminação, somos bombardeados sem trégua, e de forma contínua (celulares, tele-diários, jornais e revistas, internet...), com previsões a respeito da gravidade da pandemia, potencial de transmissão do vírus, número de casos aqui e acolá, número de mortos segundo local e data, efetividade de tal e qual intervenção de controle etc. Resultados de uma infinidade de estudos baseados em “modelos matemáticos”, quase sempre atribuídos de forma genérica a “especialistas” e comumente anunciados como arautos da ciência, à população são apresentados de forma seca, sem o devido cuidado e sem as devidas considerações. Sem informações sobre as hipóteses e/ou fonte e qualidade dos dados utilizados na modelagem matemática, discrepâncias expostas ao serem contrastadas tais previsões com a situação de fato, podem gerar ruído imprevisível na estima pelo público em geral pela ciência em si.

Tal contexto somado a minha formação como físico, o que me confere um alicerce básico para entender, ainda que com certo esforço, como modelos matemáticos simples são construídos, selecionei e li 5 artigos publicados sobre o tema em revistas científicas de altíssimo prestígio: (2 *Nature*, 1 *Science*, 1 *Proc. Royal Soc. B* e 1 *Am. J. Epidemiology*). Devido ao apelo do momento, todos os 5 artigos se encontram com acesso aberto e gratuito. Qualquer pessoa, a partir da página web da revista, pode ter acesso a esses e a muitos outros sobre o tema. Adianto que há uma média de 3 artigos diários sendo publicados na revista *Science* sobre a COVID-19.

Dessa forma, esse texto pretende expor o pouco do que aprendi desses 5 artigos sobre a matemática aplicada a epidemias. Conceitos introdutórios básicos, apresentados aqui sem as devidas demonstrações formais e rigorosas (ficam como exercícios para os leitores), podem ser úteis para nos mostrar minimamente sobre o

que especialistas falam quando falam em avaliações de gravidade e/ou previsões de cenários. Pretende também mostrar como a escolha de fontes de dados pode influenciar na estimativa do potencial devastador de uma epidemia (no caso, a de Influenza de 1918) e como estimar a eficácia de uma medida de intervenção de controle como vacinação ou isolamento (seletivo ou não). A última parte desse texto traz comentado um artigo (publicado na *Science*) que trata da aplicação dos conceitos apresentados para o caso da COVID-19 e discute a eficácia de protocolos de isolamento. Chamo a atenção que tal artigo foi escolhido com o intuito único de exemplificar como conceitos matemáticos simples e básicos são aplicados a um problema atual de capital importância mundial utilizando para isso um artigo publicado na *Science*, uma das revistas científicas mais importantes do mundo. Sendo assim, a escolha desse artigo, por ter sido um dos primeiros que li, foi por acaso e não sinaliza que as considerações nele contidas tenham relação com minha opinião.

Gostaria de deixar claro que não sou especialista em nenhuma das áreas pertencentes a macro área multidisciplinar relativa a epidemias. Pretendi ser correto na medida da minha ignorância e, nem de longe, pretendo dar a esse texto algum valor acadêmico. Sendo coerente, os 5 trabalhos consultados estão citados de forma amadora, de acordo com a conveniência, no decorrer do texto. No final estão todos os 5 trabalhos citados de forma completa para que o interessado possa facilmente chegar a eles com a ajuda mínima do Google.

O texto começa definindo o problema e, a partir do fenômeno modelo do crescimento populacional, apresentar uma linha de raciocínio que nos permite construir os conceitos matemáticos básicos e gerais que logo vão ser utilizados, com suas devidas adaptações, ao caso epidemiológico.

Mais adiante, exemplifica-se a aplicação de tais conceitos básicos no caso da famosa pandemia de Influenza de 1918 e definem-se, de forma geral, parâmetros que pretendem quantificar tanto o potencial de transmissão da epidemia quanto a eficácia de certas intervenções de controle. Na parte final, apresentamos os resultados de um trabalho que aplica todo o conteúdo discutido no caso da COVID-19 e analisa a eficácia da combinação de duas intervenções de controle no combate a epidemia.

2. O que é uma epidemia?

Podemos definir uma epidemia como sendo o surto transitório (com início, meio e fim) de uma doença infecciosa transmissível. Epidemias podem dar-se a partir da transmissão de um agente patológico de forma direta, entre hospedeiros, ou indireta, através do meio ambiente ou hospedeiros intermediários. No contexto de modelagem matemática, considerando apenas transmissão via contato direto com indivíduo infectado, a eficiência da transmissão da doença pode ser caracterizada por meio de dois aspectos importantes: i) A “Infecciosidade” relativa a um indivíduo infectado, definida principalmente por aspectos biológicos, comportamentais e ambientais; e ii) a “Susceptibilidade” a infecção para um indivíduo saudável, podendo sofrer influência das condições do sistema imunológico do indivíduo em contato com o agente infeccioso, por exemplo.

3. Infecciosidade e susceptibilidade

A Infecciosidade de um indivíduo infectado se refere, qualitativamente, ao potencial de tal indivíduo em infectar outros. Como fator biológico definidor dessa propriedade podemos citar, por exemplo, a taxa de excreção do agente infeccioso pelo indivíduo infectado, estreitamente ligada ao ciclo de vida do agente patológico ou à carga viral (ou bacteriana) em locais anatômicos do corpo infectado. Como aspecto comportamental, um exemplo seria a frequência de contatos com indivíduos saudáveis realizados pelo indivíduo infectado. Finalmente, o ambiente determina a sobrevivência do agente infeccioso e do eventual hospedeiro intermediário ou vetor transmissor, fatores que afetam a eficiência da transmissão. Por exemplo, variações climáticas de temperatura ou chuvas dão caráter sazonal a doenças como cólera, influenza ou pólio, sendo a temperatura um exemplo de aspecto ambiental que pode favorecer ou não a Infecciosidade.

A título ilustrativo, o artigo de Grassly & Fraser (*Nature reviews microbiology* n. 6, pp. 477, 13 de maio de 2008) traz três gráficos da componente biológica da Infecciosidade, ou Infecciosidade biológica, em função do tempo (contado a partir da infecção) para três patógenos humanos diferentes: Influenza A, HIV -1 e Malaria. Reproduzimos tal figura abaixo, como Figura 1. O artigo cita os trabalhos consultados para a construção das ilustrações da figura. Na Figura 1 (a), a Infecciosidade biológica para a Influenza A foi estimada a partir da taxa de emissão viral de um indivíduo infectado; já a Figura 1 (b), para o caso do HIV-1, foi feita baseada na carga viral de

indivíduos infectados; e a Figura 1 (c), para a malária, foi feita baseada na detecção de gametófitos em mosquitos que tiveram contato com pacientes de Sífilis uma vez que, antes da descoberta/invenção da penicilina, não havia tratamento específico para a Sífilis e o paciente era tratado contaminando-o por malária (Malarioterapia - prêmio Nobel de Medicina de 1927 a Julius Wagner-Jauregg).

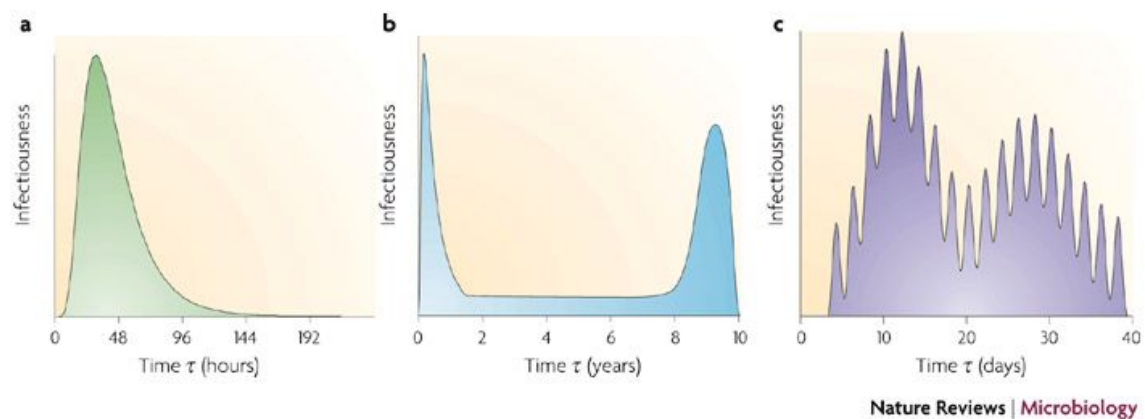


Figura 1. Reprodução da figura 1 do artigo de Grassly & Fraser, representa a Infeciosidade como função do tempo para: (a) Influenza A, (b) HIV-1 e (c) Malária. Os trabalhos originais que estudaram cada caso estão apropriadamente citados pelos autores.

Vemos que, do ponto de vista estritamente biológico, indivíduos com Influenza A apresentam um potencial de infecção máximo depois de dois dias de haverem sido infectados quando, provavelmente, ainda nem desenvolveram os sintomas. Doentes infectados por HIV-1 apresentam um alto potencial infeccioso logo nos primeiros meses depois da infecção, seguido de uma queda para os anos seguintes e voltando a subir perto de 10 anos após contraírem o vírus. O caso da Malária é bem mais curioso, vemos que a Infeciosidade apresenta uma oscilação de curto período (um dia) seguindo outra oscilação envelope de período mais longo (20 dias). O eixo horizontal dos gráficos da Figura 1 representa o tempo a partir do momento de infecção de um indivíduo saudável, respeitando a escala temporal de cada caso. Já o eixo vertical não especifica a grandeza que representa porém, do ponto de vista qualitativo, já sabemos, dá a idéia de potencial para infectar. Quantitativamente, no caso da Influenza A, por exemplo, pode ser uma quantidade proporcional a quantidade de vírus por unidade de tempo emitida pelo paciente.

Por outro lado, Susceptibilidade está estreitamente relacionada a dinâmica do agente no corpo do indivíduo em contato com o infectado e dependerá, por sua vez, de uma série de fatores: sistema imunológico do indivíduo, aspectos particulares da replicação do patógeno e seu espalhamento pelo corpo do infectado, dose inicial do agente infeccioso no momento inicial da infecção, interação do patógeno com drogas etc. A susceptibilidade de um indivíduo a ser infectado igualmente depende de componentes biológicas, comportamentais e ambientais. Por exemplo, a probabilidade de um indivíduo ser infectado por HIV-1 durante uma relação heterossexual sem proteção dependerá da carga viral do indivíduo infectado assim como será maior caso algum dos indivíduos apresente uma doença ulcerativa sexualmente transmissível.

4. Bases para um modelo baseadas no crescimento populacional

Tal como fazem os autores Wallinga & Lipsitch, em um de seus trabalhos (*Proceedings of the Royal Society* n. 274, pp. 599, 28 de novembro de 2008) , a partir de um modelo de crescimento populacional, formularemos hipóteses para derivar os conceitos básicos importantes para a construção de modelos matemáticos epidemiológicos que serão discutidos a seguir. No desenvolvimento que segue, nos referiremos a nascimentos de uma forma geral, não importando o gênero, como se todos os indivíduos fossem capazes de se reproduzir. Tal simplificação não tem nenhuma influência nas conclusões, uma vez o raciocínio não mudaria se considerássemos apenas nascimentos femininos e supondo que sempre haveria um número suficiente de masculinos para garantir as reproduções.

Partimos então do seguinte raciocínio. Em uma hipotética população, o número de novos nascimentos em um determinado ano (caracterizado pelo instante de tempo t , considerando o presente sendo $t = 0$) cujas mães têm uma determinada idade a é igual a quantidade de nascimentos (possíveis mães, sobreviventes ou não) no instante de tempo $t - a$, multiplicado pelo número esperado de descendentes gerados (*offspring* em inglês) por ano, de mães com idade a .

Somando todos os nascimentos para todas as possíveis idades das mães, teremos como resultado o número total de nascimentos no ano t em particular. Matematicamente, teríamos:

$$b(t) = \sum_{a=0}^{\infty} b(t-a)n(a) \quad (1)$$

Onde $b(t)$ é o número de nascimentos em um ano t e $n(a)$ é a taxa de nascimentos anual para mães com idade a . Apesar de indicar uma somatória infinita, a partir de uma certa idade (100 anos, por exemplo) todos os n (para $a > 100$) vão ser nulos, limitando na prática o número de termos na somatória. Por outro lado, a expressão pôde ser escrita em termos de uma somatória discreta pois nosso intervalo temporal é de um ano, razoável nesse caso uma vez que essa é a escala temporal característica para esse evento. Se o tempo característico entre um nascimento e outro, para uma mesma mãe, fosse muito menor em comparação ao tempo de observação t (segundos em comparação com anos, por exemplo), de tal forma que pudéssemos inclusive considerar a variável a como sendo contínua (qualquer número real) e não discreta (números naturais: 0, 1, 2,...,69,70,...), nossa somatória seria impraticável pois teríamos infinitos termos de fato. Nesse caso, a somatória se transforma em uma integral e a expressão 1 se torna:

$$b(t) = \int_0^{\infty} b(t-a)n(a)da \quad (2)$$

A equação 2 é conhecida como equação de renovação para o caso de nascimentos em um crescimento populacional.

A taxa de nascimentos para uma idade a , $n(a)$, evidentemente varia com a idade a da mãe. A forma como varia depende de fatores biológicos e comportamentais da mãe. Poderia ser bem representada, supomos, pela fertilidade (indicada pelo potencial para engravidar e medido via algum indicador que não faço idéia) multiplicada pela taxa (número de vezes por unidade de tempo) de relações sexuais sem métodos anticonceptivos realizados pela mãe (supondo que a fertilidade do parceiro é sempre a mesma, para simplificar).

Uma quantidade importante que podemos introduzir é número de *offspring* produzido por uma mãe ao longo de toda sua vida, ou seja, o número total de descendentes femininos de uma mãe. Esse número tem um nome e um símbolo especiais: número de reprodução e R , respectivamente). Usando o mesmo raciocínio,

podemos escrever esse numero como uma integral da taxa de nascimentos como função da idade (agora uma variável contínua) $n(a)$ longo de toda sua vida:

$$R = \int_0^{\infty} n(a) da \quad (3)$$

A função $n(a)$ nos fornece o numero de nascimentos por unidade de tempo (taxa de nascimentos) para uma mãe de idade a . A partir dessa função podemos derivar outra de extrema importância, denominada $g(a)$, simplesmente dividindo $n(a)$ por R :

$$g(a) = \frac{n(a)}{R} \quad (4)$$

Na expressão 4 acima, $g(a)$ leva o nome de “distribuição de geração de tempo” (na literatura especializada em inglês: *distribution generation time*) e representa, entre todos os conceitos aqui apresentados, o que talvez requeira o maior esforço para se entender e digerir. Mas nós vamos explicá-lo aqui. A função distribuição de geração de tempo $g(a)$ é o que chamamos em matemática de uma função densidade de probabilidade. Ou seja, como qualquer outra função, a partir de um argumento particular (no nosso caso, um número que representa uma certa idade a), a função devolve, a partir de operações com o argumento, outro número. Se os cálculos em questão estão sendo aplicados a um evento natural, geralmente o argumento da função representa uma grandeza que tem, portanto, sua unidade específica a depender de sua dimensão (no nosso caso a pode ser medido em anos - 53,378 anos, por exemplo - pois tem dimensão de tempo). Dessa forma, uma função de a , $f(a)$, qualquer, pode devolver outro número que representa uma grandeza de dimensão diferente da representada por a e que, portanto, leva outra unidade. Como exemplo, a , como já dito, possui dimensão de tempo e pode ser dada em anos. Já $n(a)$, por ter dimensão de taxa temporal ou frequência, pode ser medida em nascimentos por ano (ou qualquer unidade de dimensão 1/tempo). Uma função densidade de probabilidade possui dimensão de probabilidade/dimensão do argumento da função e seu valor para um determinado valor de seu argumento não tem sentido prático. Vejamos. A partir de uma simples análise dimensional, vemos que a dimensão de $g(a)$

é a mesma da de $n(a)$, pois ela é obtida a partir da divisão de $n(a)$ por um número absoluto (uma normalização apenas). Desse modo, supomos que temos a forma explícita e analítica da função $n(a)$ para um indivíduo de uma população em estudo. Se colocarmos o argumento $a = 23,678$ anos na expressão 4 e calcularmos $g(23,678 \text{ anos})$ podemos obter, digamos, $g(23,678 \text{ anos}) = 0,3/\text{ano}$ (unidade 1/ano), o que não nos fornece nenhuma informação prática pois não se trata de uma taxa de nascimento mas sim de uma taxa de probabilidade!

O que sim tem sentido é obter de $g(a)$ a fração de filhos, relativa a quantidade total R de filhos, que uma determinada mãe tem quando atingir a idade específica a .

Para entender, imaginemos que $g(a)$ não varie com a , ou seja, é igual a um valor G para qualquer idade a dessa mãe em questão. Se multiplicarmos G (supondo que G seja dado em 1/ano) por uma idade em concreto, digamos, 20 anos, obteremos a fração de filhos que essa mãe vai ter (em relação ao numero total R) aos 20 anos. Atenção! Para que $g(a)$ seja uma constante ao longo de toda a vida de uma mulher, necessariamente $n(a)$ também tem que ser constante pois R é um numero que não varia com a idade a . Estaríamos dizendo que a mãe gera filhos a uma taxa constante de G filhos por ano a vida toda, por exemplo. Suposição totalmente irreal porém, útil para se entender a tal função $g(a)$. Como $n(a)$ obviamente varia com a idade, para obter a informação acima, não podemos fazer uma simples multiplicação mas sim temos que integrar $g(a)$ entre zero a idade específica a' para a qual queremos obter a informação:

$$F(a') = \int_0^{a'} g(a) da \quad (5)$$

Em 5, $F(a')$ é a fração de filhos, em relação ao total, que a mãe em estudo terá aos a' anos. Agora, se considerarmos a plausível hipótese que o crescimento populacional é um fenômeno “markoviano”(não se preocupem com isso agora, não influenciará o entendimento geral do assunto), ou seja, que medias ao longo do tempo para um indivíduo fornecem o mesmo resultado que medias para uma população toda em um determinado instante de tempo, a expressão 5 pode significar algo fabuloso para as análises vindouras: a probabilidade do primeiro filho de uma mãe nascer quando essa tenha a idade a . Por isso o nome de “distribuição de geração de tempo”. O significado mais útil de $g(a)$ é que sua integral de zero (nascimento da mãe) até a'

(idade qualquer) fornece a probabilidade dessa mãe ter seu primeiro filho com essa idade a' . Interessante, não? Sim!

Vamos dar um passo a mais no modelo populacional antes de saltarmos para o epidemiológico. Supomos agora que a população cresce exponencialmente respeitando uma distribuição de idades estável (a fração de pessoas com uma determinada idade não muda com o tempo). Assim, o número de nascimentos em um determinado instante de tempo t é igual ao número de nascimentos a unidades de tempo atrás multiplicado pelo fator exponencial de crescimento referente a esse tempo, ou seja:

$$b(t) = b(t - a)e^{ra} \quad (6)$$

Sendo r é a taxa de crescimento.

Aqui cabe uma abertura de parênteses.

Você pode ter lido que o crescimento não é exponencial com base e mas sim com base no número de descendentes reproduzidos a partir de um indivíduo. Por exemplo, se um indivíduo dá origem a dois, depois de um tempo, esses dois darão origem a 4, que se transformarão em 8 depois do mesmo tempo característico de reprodução. E assim vai. Logo, o correto seria um crescimento exponencial na base 2 e não na base e ! Mas veja:

$$b(t) = b(t - a)e^{ra} = b(t - a)2^{sa}$$

desde que

$$\begin{aligned} \ln(b(t - a)e^{ra}) &= \ln(b(t - a)2^{sa}) \Rightarrow \\ \ln(b(t - a)) + \ln(e^{ra}) &= \ln(b(t - a)) + \ln(2^{sa}) \Rightarrow \\ \ln(e^{ra}) &= \ln(2^{sa}) \Rightarrow \\ ra &= sa \ln(2) \Rightarrow \\ r &= s \ln(2) \end{aligned}$$

Ou seja, tanto faz descrever o crescimento populacional na base e , na base 2 ou em qualquer outra base. Apenas deveremos utilizar a taxa de crescimento adequada para a descrição, ou seja, aquela que reproduz os dados experimentais. O fato de escrevermos na base e é simplesmente pelo fato de que as operações do cálculo diferencial e integral serem enormemente facilitadas nessa base. Aqui fechamos os parênteses.

Voltando a expressão 6 para $b(t)$ e multiplicando ambos lados por e^{-ra} , podemos escrever $b(t-a)$ como:

$$b(t-a) = b(t)e^{-ra} \quad (7)$$

Substituindo agora $b(t-a)$ da expressão 7 na equação 2 para $b(t)$ previamente construída a partir de nossas primeiras hipóteses, teremos uma expressão transcendental para $b(t)$, ou seja, uma equação com $b(t)$ a direita e a esquerda da igualdade:

$$b(t) = \int_0^{\infty} b(t)e^{-ra}n(a)da \quad (8)$$

Estou de acordo que não é evidente, em um primeiro momento, o porque dessas passagens matemáticas porém, garanto que o caminho nos levará a resultados úteis que poderemos trasladar diretamente para o caso de epidemias, escopo desse texto.

A expressão 8 para o número de nascimentos no momento t possui a mesma informação que a expressão 2 para essa mesma quantidade, apenas reescrita considerando fator exponencial para o fenômeno populacional e reescrita de uma forma muito menos intuitiva. O motivo disso fica claro, matematicamente quando, nessa expressão, substituímos $n(a)$ pelo produto de $g(a)$ por R , a partir da expressão 4 para a função distribuição de geração de tempo. Sendo assim, teremos:

$$b(t) = \int_0^{\infty} b(t)e^{-ra} Rg(a)da \quad (9)$$

Como tanto $b(t)$ quanto R são fatores comuns na soma infinita implícita na operação Integral, temos:

$$b(t) = b(t)R \int_0^{\infty} e^{-ra} g(a)da \Rightarrow$$

$$\frac{1}{R} = \int_0^{\infty} e^{-ra} g(a)da \quad (10)$$

Pra quem tem maior envolvimento com matemática (não necessário para seguir com a leitura) identifica a integral do lado direito da última equação como sendo a famosa “transformada de Laplace” da função $g(a)$. Transformadas de Laplace (em homenagem a Pierre-Simon Laplace – Beaumont en Auge França, 23 de março de 1749 – 5 de março de 1827) é uma genial concepção do intelecto humano (referente ao cálculo diferencial e integral), ao mesmo tempo útil e linda, que nos permite dizer, nesse caso, que o inverso do número de reprodução R é a transformada de Laplace da função distribuição de geração de tempo $g(a)$. Existe todo um ramo de conhecimento por trás disso, com métodos de cálculos de transformadas de Laplace das mais variadas funções, tabelas etc. O Fato é que, no fim do dia, se tivermos a função $g(a)$ característica do crescimento populacional que desejamos estudar, teremos R , como função da taxa de crescimento r , a partir da transformada de Laplace de $g(a)$. A relação entre função e sua transformada de Laplace (quando existir) é única e de duplo sentido. Sugiro ao aluno de cálculo buscar estudante o significado dessas afirmações.

Sendo assim, no perfeito mundo da matemática, podemos determinar a função $R(r)$, ou seja, o número de reprodução R para uma dada taxa de crescimento exponencial de uma população a partir das mais variadas funções distribuição de probabilidades (classe que se enquadra as $g(a)$) contínuas e bem comportadas. O artigo de Wallinga & Lipsitch trás, como exemplos, gráficos de $R(r)$ para funções $g(a)$ representando comportamentos gaussiano, delta, geométrico etc. O leitor que queira se aprofundar deveria conferir o citado artigo.

No entanto, ao se estudar um crescimento populacional, dispomos de dados empíricos e não de funções matemáticas ideais. Nesse caso, o lado direito da expressão 10 pode ser inferido numericamente, a partir de métodos computacionais simples. Veremos exemplos disso a seguir, voltando ao tema protagonista desse texto: epidemias.

5. Como usar o que desenvolvemos até aqui para o caso das epidemias de doenças transmissíveis?

A analogia entre o crescimento populacional baseado em nascimentos, discutida nos parágrafos anteriores, e o crescimento de doentes infectados por um agente transmissível via contato entre indivíduo infectado e saudável (ou via hospedeiro intermediário) parece direta. No caso da epidemia, a variável a será representada pelo tempo τ , que poderá ser contado a partir da infecção de um indivíduo ou a partir do início da epidemia (infecção do caso índice), dependendo do contexto. A depender da escala temporal característica definida pela epidemia, τ pode estar contado em dias, horas ou minutos, vide Figura 1. A função $b(t)$ representará aqui o número de novos infectados no instante de tempo τ e levará o símbolo $I(\tau)$. A função $n(a)$ é substituída pela função Infeciosidade $\beta(\tau)$ que, analogamente a $n(a)$, pode ser dada pela Infeciosidade biológica multiplicada pela frequência de contatos realizados pelo indivíduo infectado. Essa frequência de contatos pode variar de indivíduo para indivíduo de acordo com a natureza da doença em estudo. No caso de doenças sexualmente transmissíveis a dispersão no padrão de contatos individuais é bastante alta, uma vez que alguns indivíduos apresentarão alta frequência de contatos de natureza íntima enquanto que outros estarão no outro extremo. Já para o caso de contatos de natureza social (aperto de mão, conversas, aproximações etc.), relevantes para Influenza A, a variação da frequência de contatos, considerando um número elevados de indivíduos, tende a ser mais próxima de uma média comum. Já a função $g(a)$ será, no contexto das epidemias, $\omega(\tau)$, se chamará igual e caracterizará a distribuição de intervalos de tempo entre a infecção de um indivíduo e uma infecção secundária causada pelo mesmo. Logo, responderá a seguinte pergunta: qual a probabilidade de um indivíduo infectar seu primeiro caso secundário em τ unidades de tempo a partir de sua infecção? Por fim, $R(r)$ será número total de casos secundários

gerados por um indivíduo infectado e, em analogia com o crescimento populacional, dependerá da taxa de exponencial r de crescimento de infectados e poderá ser determinado diretamente e de forma única, a partir da função $\omega(\tau)$ que caracteriza a epidemia.

Uma das diferenças entre os dois fenômenos tal como descritos (crescimento populacional e transmissão de doença infecciosa) está no fato da transmissão não depender somente da Infeciosidade do indivíduo infectado mas também da Susceptibilidade do indivíduo saudável. Assumindo, para simplificar, que a Susceptibilidade média de uma população não varia com o tempo e que possa ser caracterizada por um número (digamos s) variando de zero a um (0 a 1), sendo $s = 0$ totalmente imune e $s = 1$ totalmente susceptível, a expressão análoga a equação 3 para R seria:

$$R = s \int_0^{\infty} \beta(\tau) d\tau \quad (11)$$

Para o caso de uma população cujos indivíduos são todos totalmente susceptíveis ($s = 1$ – epidemia causada por um vírus novo), teremos o que se chama de Numero de reprodução básico, denominado por R_0 :

$$R_0 = \int_0^{\infty} \beta(\tau) d\tau \quad (12)$$

Ou seja, o número total R de infecções secundárias provocadas por um indivíduo é a rebaixado, com relação a R_0 , por um fator igual a susceptibilidade média da população.

Como a Susceptibilidade está embutida em R , todas as outras expressões derivadas para o caso populacional são diretamente trasladadas para o estudo das epidemias, necessitando apenas fazer as substituições de conceitos e símbolos apropriadamente.

6. Passando de indivíduos para epidemias

Até aqui todos os conceitos foram introduzidos para um indivíduo da população em estudo. Por exemplo, R informa o número esperado de infecções secundárias causadas por um indivíduo infectado (possuindo sua particular função Infeciosidade $\beta(\tau)$) em uma população de indivíduos igualmente susceptíveis. Poderíamos extrapolar e dizer que essa informação se refere a epidemia e não mais a um indivíduo, apenas se considerássemos que todos os infectados apresentassem a mesma Infeciosidade (mesma função $\beta(\tau)$ para todos) o que na prática não acontece. Tanto a Infeciosidade biológica quanto a Susceptibilidade variam de indivíduo para indivíduo de acordo com uma série de fatores determinantes para eficiência da transmissão, sendo alguns deles já mencionados no início do texto. Uma forma simplificada de tratar esse fato e que pode ser de utilidade matemática ao construir-se modelos, é dividir a população em grupos de indivíduos semelhantes no que se refere a manifestação desses fatores. Teríamos, dessa forma, n grupos (grupo 1, grupo 2, grupo 3, ..., grupo n) cada qual representado por função Infeciosidade e Susceptibilidade comum aos indivíduos do grupo porém diferente das dos outros grupos.

Em uma população cujos indivíduos pertencentes aos diferentes grupos estejam distribuídos de forma randômica no espaço geográfico, o número de infecções secundárias respeitará uma distribuição dada pela mistura (média) das obtidas em cada grupo ponderada pelo tamanho do grupo. Um grupo que contem um número grande de indivíduos semelhantes quanto a Infeciosidade/susceptibilidade terá maior influência na forma da distribuição de casos secundários que um grupo de poucos indivíduos.

Um aspecto emergente interessante surge aqui. Mesmo que não se possa prever o comportamento de um indivíduo em particular, a dinâmica da epidemia em uma população é uma propriedade emergente e segue um comportamento que pode ser previsto matematicamente, desde que modelado de forma correta. Esse fato pode ser entendido a partir da lei dos números grandes, que prevê que diferenças aleatórias entre comportamentos de indivíduos tendem a se cancelarem para uma população suficientemente grande.

7. Modelos epidêmicos ou observações empíricas?

Cabe aqui fazer uma observação bastante importante. Quando consideramos, na sessão 4, que uma população apresenta crescimento exponencial (vinculado a distribuição estável de idades) não atrelamos nenhuma outra propriedade aos indivíduos a não ser estarem vivos. Ou seja, uma população aumenta quando o número de indivíduos que a compõe aumenta, e ponto final. Por outro lado, os modelos matemáticos mais comuns utilizados para descrever a condição da população perante a incidência de uma epidemia pretende, na verdade, descrever a dinâmica (evolução) da condição de saúde dos indivíduos que a compõe. Sendo assim, introduziremos alguns aspectos de um modelo bastante utilizado, que considera a população em estudo composta por três tipos de indivíduos vivos: susceptíveis (S), infectados (I) ou recuperados (R). As hipóteses do modelo, que resultarão em equações diferenciais, cujas soluções nos fornecerá a evolução temporal das frações S, I e R da população, são as seguintes:

Para um determinado instante de tempo concreto durante a epidemia:

i) A fração S (susceptível) da população está decrescendo. Tal decréscimo instantâneo é tão maior quanto maior for a fração I (infectados) naquele instante de tempo (infectados infectam). Ao mesmo tempo, o decréscimo temporal de S é proporcional a própria fração S de indivíduos susceptíveis (em um determinado momento, quanto mais indivíduos S existirem, maior a taxa deles que deixam de ser S para ser I). Matematicamente, com λ sendo a constante de proporcionalidade, temos:

$$\frac{dS}{dt} = -\lambda SI \quad (13)$$

ii) A fração recuperada R da população está aumentando. Seu aumento é maior quanto maior for, no dado instante de tempo, a fração I de infectados. Se η for a constante de proporcionalidade, matematicamente teremos:

$$\frac{dR}{dt} = \eta I \quad (14)$$

iii) Por fim, a fração infectada I da população pode estar aumentando ou diminuindo, dependendo se estamos mais pro início ou pro fim da epidemia. O que leva a seu aumento é exatamente o decréscimo de S (13), ou seja, todos os que estão deixando de ser S necessariamente, evoluem para I . Por outro lado, nesse dado instante, vários I 's estão evoluindo para R (14). Logo, a taxa de variação de I , matematicamente, é dado por:

$$\frac{dI}{dt} = \lambda SI - \eta I \quad (15)$$

Nesse caso, uma das equações é redundante, uma vez que a soma das três frações S, I e R necessariamente é igual a 1:

$$S + I + R = 1 \quad (16)$$

Esse modelo despreza nascimentos, mortes, imigração ou emigração. São aproximações razoáveis uma vez que a escala de tempo da epidemia é muito menor que a necessária para se observar grandes variações populacionais. A taxa de contatos feitos pelos indivíduos também é assumida constante pelo modelo SIR.

Ao lançarmos mão de modelos como o SIR acima, implicitamente estamos definindo a função distribuição de geração de tempo $\omega(\tau)$ definida na sessão 4 (expressão 4). O que é de se esperar uma vez que a função do modelo é exatamente “fabricar” a dinâmica da epidemia por meio das equações diferenciais envolvidas. No caso 13, 14 e 15. Mais, se leva implícito $\omega(\tau)$, estará nele também $\beta(\tau)$. Dessa forma, o modelo concreto SIR leva implícito (e não demonstrado aqui – acreditem ou façam o exercício da demonstração, recomendado) uma distribuição de densidade de probabilidades exponencial para $\omega(\tau)$ (distribuição de tempos para o primeiro caso secundário gerado por um indivíduo infectado), com um tempo médio dado por $\tau_m = 1/\eta$. A transformada de Laplace para essa $\omega(\tau)$ (que pode ser calculada ou, devido ao grau de ordinariedade, encontrada em tabelas que habitam livros ou páginas web didáticos) nos leva a seguinte relação entre o número de reprodução R , taxa de crescimento exponencial de casos novos (r) e a constante η :

$$R = 1 + \frac{r}{\eta} \quad (17)$$

Outro modelo, um pouco mais elaborado e que não vamos nos estender aqui, considera um período intermediário entre o susceptível e o infectado, o exposto (E). Tal modelo ficou conhecido por SEIR. Não colocarei aqui as equações diferenciais para o SEIR, apenas dizer que ele é caracterizado e constantes (λ , η_1 e η_2) e não duas como o SIR. A terceira constante (η_2) está relacionada a distribuição de geração de tempo pré sintomas, ou seja, fornece a distribuição de probabilidades relacionadas ao tempo entre infecção e surgimento dos sintomas. A função geração de tempo total, nesse modelo é dada pela convolução (sugiro como mais uma palavra motivadora para o aprofundamento nos estudos – Googleável e Wikipediável em primeiro momento) de duas distribuições exponenciais com tempo médio dado por $\tau_m = 1/\eta_1 + 1/\eta_2$. O raciocínio continua o mesmo. No modelo SEIR a relação entre R , r , η_1 e η_2) fica:

$$R = \left(1 + \frac{r}{\eta_1}\right) \left(1 + \frac{r}{\eta_2}\right) \quad (18)$$

Para outros modelos ou situações nas quais a função $\omega(\tau)$ seja gaussiana (normal) com tempo médio τ_m e desvio padrão σ teríamos:

$$R = e^{r\tau_m - 1/2(r^2\sigma^2)} \quad (19)$$

Para uma distribuição de geração de tempo $\omega(\tau)$ para a qual todos primeiros casos secundários acontecem para o mesmo tempo τ_m (chamado de função delta):

$$R = e^{r\tau_m} \quad (20)$$

E, claro, para o caso de que seja possível (desejável) obter $\omega(\tau)$ a partir de dados empíricos, por exemplo, observando com lupa a epidemia no seu início (quando apresenta crescimento exponencial do número de incidências), a relação

acima é obtida a partir de métodos diversos de cálculos numéricos. Por exemplo, caso possamos fazer um histograma da frequência (y_i) com que ocorre cada valor específico de τ_i , em função do tempo τ_i , em uma população de infectados, R pode ser calculado como segue (histograma com n valores de τ_i , ou seja, $i = n$. Quanto maior o valor de n melhor nosso histograma):

$$R = \frac{r}{\sum_{i=1}^n y_i (e^{-r\tau_{i-1}} - e^{-r\tau_i}) / (\tau_i - \tau_{i-1})} \quad (21)$$

O artigo de J. Wallinga & M. Lipsitch traz uma figura ilustrando o caso empírico de Influenza estudado por Hirotsu e colaboradores, em 2004 (ver a referencia completa do trabalho de Hirotsu no artigo de J. Wallinga & M. Lipsitch. Essa figura está reproduzida abaixo como sendo Figura 2 (lembro que o artigo é de acesso livre mas reconheço que não pedi permissão para isso – pedirei caso queira usar esse texto de forma mais abrangente).

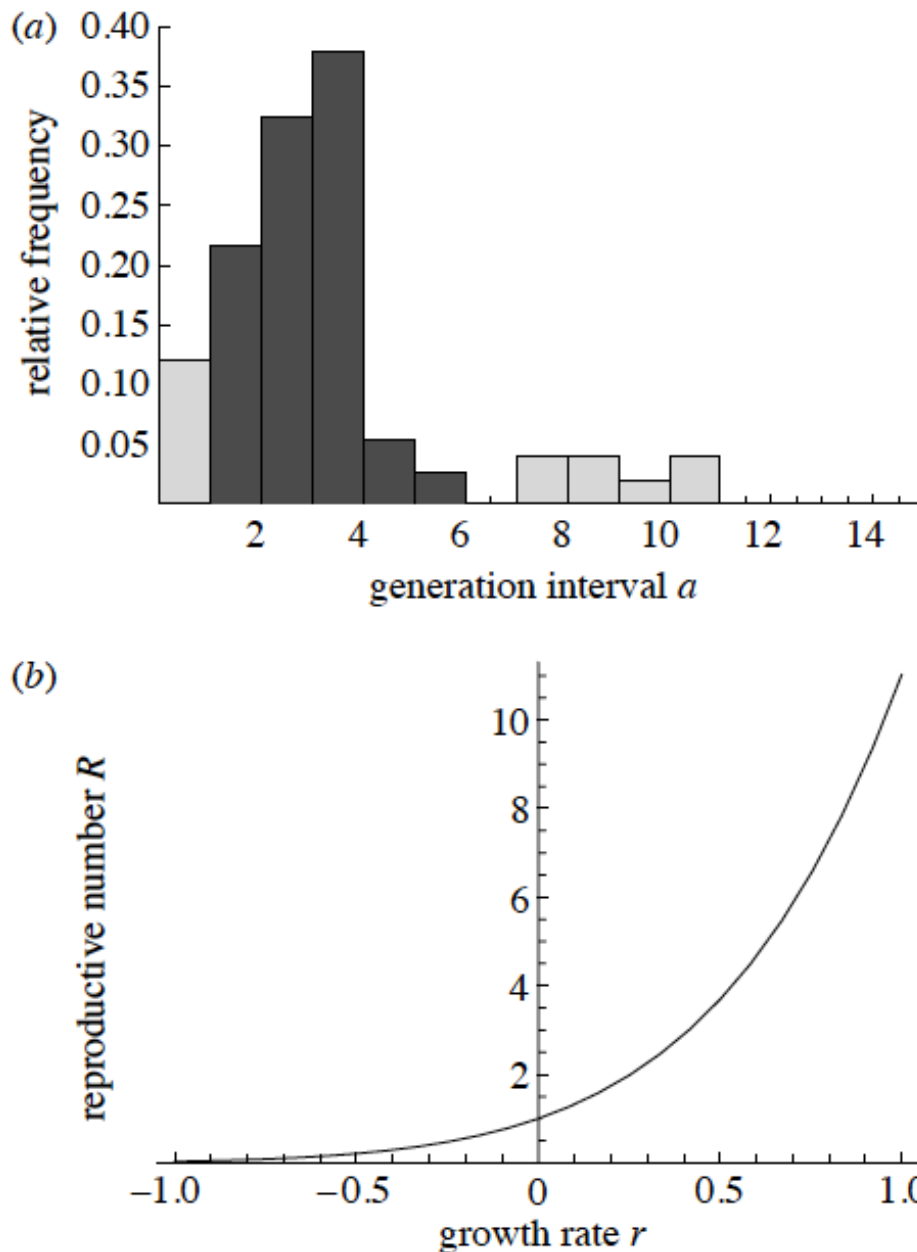


Figure 2. Observed generation intervals for influenza A virus infections and the empirical relationship between reproductive number and growth rate: (a) histogram of observed generation intervals between two subsequent influenza cases in a household (after Hirotsu *et al.* 2004), the lighter shading indicates observations that could be due to co-primary or tertiary infections and (b) the corresponding empirical relationship between R and r , according to equation (3.6).

Figura 2. Figura 2 do artigo de J. Wallinga & M. Lipsitch. (a) Histograma descrevendo a distribuição de geração de tempos $\omega(\tau)$ empírica para infecções de

Influenza A estudadas por Hirotsu e colaboradores em 2004 para casos domésticos (voltaremos a falar disso mais adiante). As barras mais claras no histograma são atribuídas a possíveis duplo casos índice (infecção co-primária) ou a infecções terciárias (detalhes no artigo original). (b) Relação empírica entre R e r obtida a partir da equação 21. Os valores para os diferentes tempos τ_i usados na expressão 21 são os limites superiores dos intervalos do histograma. Vemos que J. Wallinga & M. Lipsitch utilizam o símbolo a_i para os tempos entre caso primário e primeiro caso secundário (estamos usando τ_i), tal como foi feito na sessão 4 desse texto para o caso populacional. Na verdade, eles não mudaram a letra ao mudar o fenômeno estudado.

8. Eficácia de intervenções de contenção de uma epidemia

Como já dito no início do texto, agora em linguagem matemática, uma epidemia só ocorre, e se mantém, se R for maior ou igual a 1 (um indivíduo infectado deve, em média infectar a um número maior ou igual a 1 indivíduo saudável).

Dessa forma, um objetivo plausível de uma intervenção seria reduzir R para um número menor que 1. Se R_0 (número de reprodução básico) é o número máximo esperado (sinônimo de médio) de infecções secundárias causadas por um indivíduo infectado em uma população 100% susceptível, uma intervenção deve evitar um número (R_0-1) de cada R_0 infecções para que seja eficaz. Por exemplo, se um indivíduo infecta 3 pessoas em média, uma intervenção teria que evitar 2 dessas infecções (R_0-1) para reduzir seu R a 1, ou seja, o equivalente a ou $2/3$ delas. O mesmo que $(R_0-1)/R_0$. Claro que não há sentido em analisar uma intervenção em nível individual pois, nesse caso específico, o R seria 0, 1, 2 ou 3 e nunca maior que zero e menor que 1. Mas ao tratar da população como um todo, se uma intervenção evitar uma fração minimamente maior que $(R_0-1)/R_0$ dos casos, no momento de sua implementação, terá reduzido R para valores menores que 1 e, assim, contido a epidemia.

Dessa forma, o impacto crítico de uma intervenção para conter uma epidemia pode ser medido a partir do número relativo de casos que pode evitar em relação a R_0 , parâmetro característico da epidemia em questão. Chamaremos essa fração característica de impacto crítico de contenção (ICC):

$$ICC = \frac{R_0 - 1}{R_0} \quad (22)$$

A partir da expressão 22 fica evidente que os esforços para conter uma determinada epidemia não é proporcional ao seu valor de R_0 . Por exemplo, uma epidemia com $R_0 = 2$ requer $ICC = 1/2$. Já para uma epidemia com $R_0 = 1,5$ o ICC cai para $1/3$. Ou seja, os esforços para conter uma epidemia com $R_0 = 2$ devem ser 50% maiores que para conter uma cujo R_0 seja igual a 1,5 (de $1/3$ a $1/2$), apesar de ter R_0 apenas 33% maior (de 1,5 a 2,0). Esse simples raciocínio está corroborado com trabalhos reportados por vários artigos (ver referências no artigo original) que comprovam a ineficácia de protocolos de intervenção em conter epidemias com R_0 maiores que 2.

Para muitas epidemias, um indivíduo só saberá que está infectado após a manifestação dos sintomas. Porém, há um período pré-sintomático para o qual sua Infectiosidade pode ser elevada, como é o caso da Influenza A. Se introduzirmos aqui a função $S(\tau)$, que descreve a probabilidade de um indivíduo infectado não apresentar sintomas após um tempo τ a partir da infecção, o isolamento imediato desse indivíduo evitará um número de infecções igual a $\int_0^{\infty} S(\tau)\beta(\tau)d\tau$.

Nesse ponto os autores mencionam uma forma de se avaliar a eficácia de protocolos de intervenção como, por exemplo, o isolamento e/ou o acompanhamento dos contatos realizados por indivíduos infectados. Dessa forma, se $S(\tau)$ agora descreve a fração de infectados que, após um tempo τ da infecção, ainda não desenvolveram sintomas. Que, como já mencionado, é o mesmo que a probabilidade de um indivíduo não ter desenvolvido os sintomas após um tempo τ a partir da infecção, o isolamento imediato desse indivíduo, caso fosse possível sua identificação, evitaria uma fração de casos secundários pré sintomáticos (utilizaremos a letra grega θ como símbolo) igual a:

$$\theta = \frac{\int_0^{\infty} S(\tau)\beta(\tau)d\tau}{R_0} \quad (23)$$

Como vemos, θ pode ser um bom indicativo da efetividade do isolamento de indivíduos pré sintomáticos (pode ser comparada com o *ICC*). Uma forma hipotética de isolar indivíduos pré sintomáticos seria identificando-os a partir dos contatos recentes feitos por um infectado identificado.

A partir dessa análise, podemos também inferir que uma epidemia com um valor característico alto de θ é bastante mais difícil de ser controlada com o isolamento de indivíduos apenas depois de apresentarem sintomas, enquanto que outra epidemia com baixo valor de θ tal protocolo seria bastante mais efetivo. Como exemplo de epidemia com θ baixo podemos citar a SARS e, com θ alto, a Influenza A.

9. Saturação

Uma infecção pode produzir algum dos seguintes macro efeitos: levar o indivíduo a morte, induzir uma resposta imune ou ativar um efeito memória, simplesmente reduzindo as chances de re-infecção. Se indivíduos infectados se recuperam e assumem uma condição de imunidade ($s = 0$), a epidemia, numa população fechada, não afetará necessariamente a totalidade dos indivíduos susceptíveis, levando ao efeito de saturação. Após atingir uma certa fração da população, a transmissão se torna ineficiente e a epidemia termina. Para o caso de uma população fechada, na qual o número de indivíduos susceptíveis não é repostos, a saturação gera a clássica curva em forma de sino que descreve o número de infecções ao longo do tempo e que caracteriza a extinção da epidemia.

Pode-se demonstrar de forma relativamente simples, utilizando algum dos modelos apresentados acima (SIR, por exemplo) que a relação entre R_0 e a fração final da população que está infectada (z) é dada pela seguinte relação transcendental:

$$z = 1 - e^{-R_0 z} \quad (24)$$

A equação 24 é bastante robusta e geral, válida para uma ampla gama de considerações teóricas, como mostra os autores J. Ma & D. J. D. Earn, da Universidade de McMaster (Canadá) em um trabalho publicado em 2006 (*Bulletin of Mathematical Biology* n. 68, pp. 679, 8 de abril de 2006).

Como dito, a expressão 24 se trata de uma equação transcendental, ou seja, a variável para a qual se busca um valor (solução da equação) depende do valor dessa mesma variável (solução)! Uma forma de solucionar esse problema é utilizar uma rotina computacional simples baseada num protocolo de iteração. Geralmente se inicia “chutando” um valor para z e verificando se os dois lados da equação são iguais. Não sendo, a partir de um critério inteligente e do valor inicial testado, se escolhe outro valor para z e se repete o processo de comparação até atingir a convergência, ou seja, que os lados direito e esquerdo da equação 24 sejam iguais. Dessa forma, uma vez que se faz uma boa estimativa de R_0 para uma determinada epidemia, o cálculo da fração da população que terá sido infectada uma vez extinguida a epidemia é de fácil realização. Como análise rápida, vemos que, quanto maior o valor de R_0 mais se aproxima z da unidade, ou seja, quanto maior R_0 maior a fração da população que resultará infectada no fim da epidemia. Isso pode ser observado a partir de uma análise gráfica sugerida como exercício: coloque em um mesmo gráfico as funções $f(z) = z$ e $f(z) = 1 - e^{-R_0 z}$ para vários valores de R_0 . Os pontos onde as duas funções se cruzam é solução da equação transcendental 24.

A seguir veremos de que forma os conceitos explorados nas páginas anteriores são aplicados ao caso concreto de uma famosa pandemia mundial de Influenza, ocorrida no ano de 1918. Se estima que entre 20 e 40 milhões de pessoas perderam a vida por conta da epidemia, considerada hoje o pior cenário a se considerar para fins de planejamento. A proporção entre mortes e casos (CFP – *case fatality proportion*, em inglês) foi ao redor de 10 vezes maior que em outros casos de Influenza, sendo especialmente alta entre adultos jovens. O lugar da atual pandemia de Covid-19 se saberá uma vez extinta. Devido a importância do surto, há muitos trabalhos publicados dedicados a seu estudo. Os que selecionamos, além de ilustrar a aplicação da matemática a epidemias concretas, demonstram como a qualidade dos dados utilizados influenciam as conclusões e previsões.

10. Pandemia de Influenza de 1918

Primeiramente, em nenhum artigo consultado, cujas informações foram ou não usadas nesse texto, se referem a essa pandemia como “gripe espanhola”, tal como o

faz amplamente a mídia não especializada, apesar de uma simples busca no Google informar que o início dessa pandemia nem sequer se deu em território europeu.

C. E. Mills, J. M. Robins e M. Lipsitch, publicaram um artigo (*Transmissibility of 1918 pandemic influenza*) na revista *Nature* em 16 de dezembro de 2004, onde estimam o número de reprodução R , para a pandemia em questão, a partir de dados de 45 cidades estadunidenses (*Nature* n. 432, pp. 904, 16 de dezembro de 2004). Veremos como procederam.

Sendo os dados existentes sobre a pandemia de 1918 bastante limitados em qualidade e abundância, estimar a evolução do número de casos com o tempo é o primeiro desafio. Como os casos de morte são mais confiáveis, uma vez que sempre são registrados, eles tomaram como material para estudo a evolução temporal dos casos de morte em excesso (acima da mediana das mortes entre 1910 e 1916) em decorrência de P&I (pneumonia e Influenza) nas 45 cidades analisadas e assumiram, a partir de outros trabalhos, uma CFP de 2% - ver, por exemplo: W. H. Frost, *Statistics of influenza morbidity with special reference to certain factors in case incidence and case fatality. Public Health Reports* **35**, 584-597 (1920). A distribuição de tempo entre infecção e morte foi inferida a partir dos dados das autopsias, apresentando um valor médio de duas semanas. Os autores afirmam que (dados no material suplementar do artigo, também acesso aberto) a linha de base escolhida, para o cálculo das mortes em excesso, não afeta as conclusões. A CFP para a Influenza de 1918 foi tão maior que em outros surtos que qualquer linha de base escolhida para é suficientemente baixa de forma a ser desprezível. Os resultados obtidos pelos autores estão reproduzidos na Figura 3 abaixo (Figura 1 no artigo original). Detalhes sobre critérios de escolha do período usado no ajuste como dos pontos para o cálculo dos valores extremos de R estão descritos de forma precisa no material suplementar do artigo.

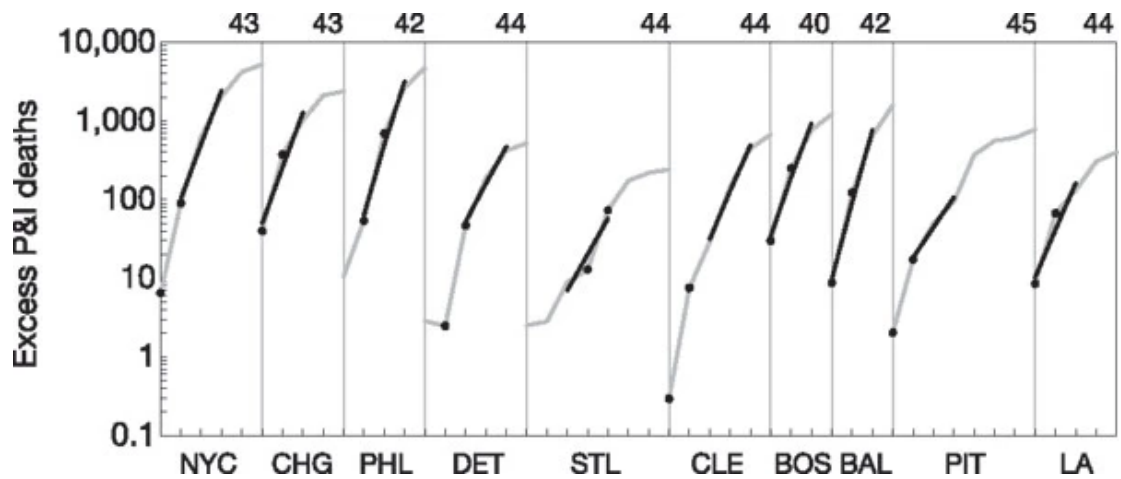


Figura 3. Curvas de mortes em excesso devido a P&I para as 10 cidades mais populosas dos E.U.A. O tamanho da população decresce da esquerda pra direita. O numero acima da barra cinza que separa as cidades indicam a semana do pico. Reparem que a escala vertical é logarítmica, logo, um crescimento exponencial é evidenciado por uma reta. As curvas cinzas representam os dados brutos e as pretas o ajuste exponencial (modelo SEIR – Susceptível/Exposto/Infectado/Recuperado - Não importa para o início da epidemia, que sempre apresentará um crescimento exponencial do numero de infectados). Os pontos negros indicam as semanas usadas para estimativas extremas de R , ou seja, utilizando apenas pontos isolados da curva e não um ajuste. Valores extremos de R são úteis como limites superiores pois são geralmente maiores que os obtidos via ajuste de um trecho considerável da curva. Cidades: New York City (NYC), Chicago (CHG), Philadelphia (PHL), Detroit (DET), St Louis (STL), Cleveland (CLE), Boston (BOS), Baltimore (BAL), Pittsburgh (PIT) and Los Angeles (LA).

Em posse do resultado dos ajustes exponenciais (fornece os valores para a taxa de crescimento r) e da distribuição de geração de tempo $\omega(\tau)$, obtida indiretamente pelos autores com a ajuda de outros trabalhos (devidamente citados no artigo), bastaria encontrar sua transformada de Laplace (computacionalmente) para obtermos os valores de R para cada r . Apesar, como já dito, do modelo SEIR vir com a forma geral de $\omega(\tau)$ embutida, os parâmetros que a particularizam (η_1 e η_2) devem ser obtidos de alguma forma. Os resultados obtidos pelos autores estão mostrados nos histogramas da Figura 4 abaixo:

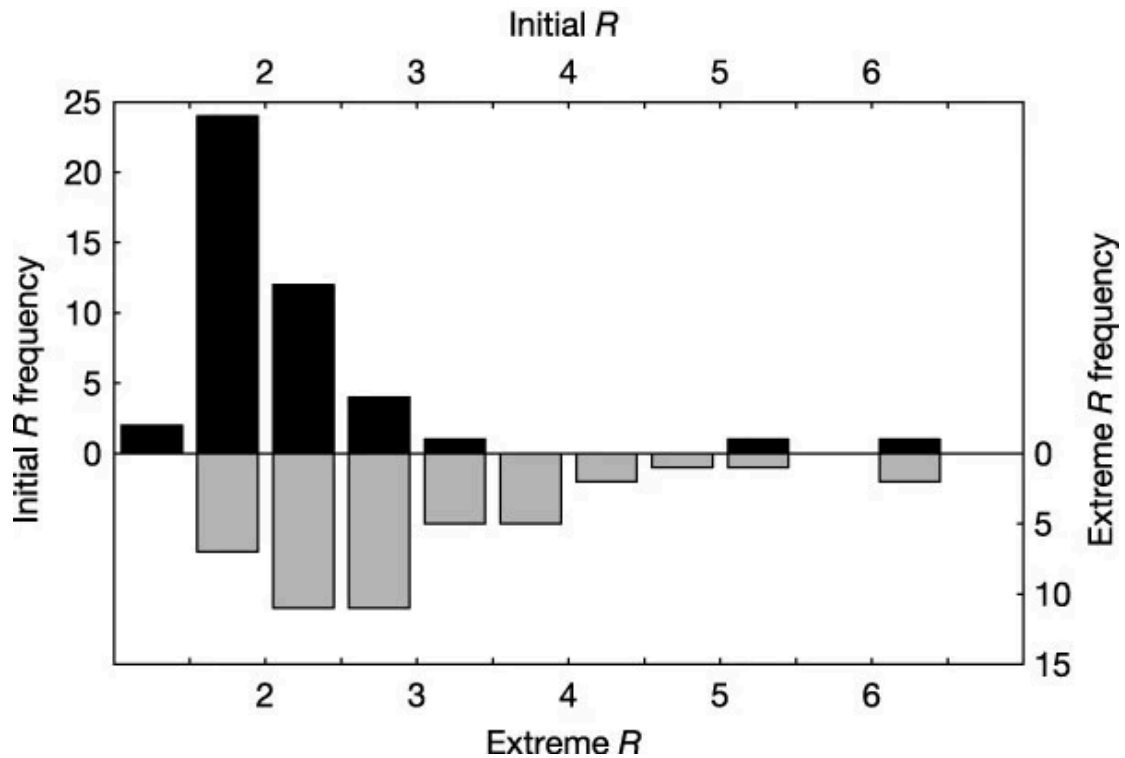


Figura 4. Histograma dos valores de R iniciais (barras negras) e extremos (barras cinzas) para as 45 cidades estadunidenses estudadas.

Os autores mencionam que os valores iniciais de R (a partir dos ajustes) não estão correlacionados com fatores tais como latitude, longitude, tamanho da população, densidade populacional, distribuição de idades ou de gênero. Já os valores extremos de R estariam estaria, fracamente, correlacionados positivamente com a densidade populacional (maior densidade maior R extremo).

Como estimativa única para R inicial (obtido do ajuste para as 3 primeiras semanas com mais de 1 morte para cada 100.000 habitantes, os autores concluem que estar entre 2 e 3, com mediana 2 e intervalo interquartil de 1,7 a 2,3 (conceitos úteis em estatística cuja bibliografia é de fácil acesso inclusive na internet – Wikipediável). Tais estimativas estão influenciadas por uma taxa de crescimento r com mediana estimada de $r_m = 0,20$.

A fração da população totalmente susceptível no início da epidemia determina a relação entre o número de reprodução R e o número de reprodução básico R_0 (número de casos secundários gerados por um indivíduo infectado em uma população 100% susceptível). A partir de observações ao longo da história, se formou um consenso de que o surto devastador no início do outono de 1918 foi precedido por uma espécie de anúncio que chegou na primavera com uma cepa de vírus parecido ao

do outono. A literatura chama esse efeito de onda anunciadora” (*herald wave*). Tal onda fez com que o surto do outono tenha encontrado ao redor de 30 % da população protegida. Dessa forma, assumindo que 70 % (ou 0,7) da população susceptível em setembro, as medianas para os valores iniciais e extremos para R_0 ficam em 2,9 e 3,9, respectivamente. Se considerarmos o valor superior estimado para R_0 (~ 4), teremos um ICC (expressão 22) de aproximadamente 0,75. Isso significa que uma intervenção de combate que evitasse 75% das possíveis infecções evitaria a epidemia (R seria menor que 1). Dessa forma, vacinando 75% da população seria mais que suficiente (medida sobreestimada) pois sabemos que o efeito de saturação limita a fração de uma população 100 % susceptível que seria infectada caso nada fosse feito. Calcular z (fração da população infectada na saturação) via equação transcendental 24 para melhor avaliar a porcentagem da população que deveria ser vacinada pode ser um exercício interessante para o estudante que está lendo esse texto agora.

Epidemias de SARS com R parecidos aos estimados aqui foram facilmente controladas com medidas de isolamento seletivo – tratamento a partir de casos identificados, ou seja, isolando e tratando os casos sintomáticos e acompanhando e isolando os contatos efetuados pelos mesmos. No caso da Influenza essa medida é pouco efetiva pois a Infectiosidade máxima se dá antes do desenvolvimento dos sintomas. Ou, como já foi dito, θ (expressão 23) para Influenza é bem mais elevado que para SARS.

Nesse artigo, as estimativas para R foram baseadas nos dados de excesso de mortes devido a P&I. No material suplementar do artigo eles também fazem as estimativas baseando-se em registros de casos sintomáticos para 3 das 45 cidades e concluem que os resultados são consistentes e comparáveis aos baseados na mortalidade. A confiabilidade dos resultados obtidos pelos autores é garantida mediante uma série de testes técnicos denominados “análises de sensibilidade”, que o quanto sensível são as estimativas a mudanças nas hipóteses consideradas. Dessa forma, esses testes de validação foram feitos para: escolha da linha base para a contagem do excesso de mortes, fonte de dados consultada, intervalo das curvas utilizado para o ajuste do modelo, tempo entre infecção e morte (considerando que a infecção ocorreu exatamente duas semanas antes da morte, R se vê diminuído em menos 0,1) etc.

J. Wallinga & M. Lipsitch revisitam esse trabalho de C. E. Mills, J. M. Robins e M. Lipsitch a luz com o intuito de demonstrar como, a partir de uma taxa de crescimento r dada (para Influenza A), hipóteses feitas a respeito da distribuição de

geração de tempo pode afetar o valor dos números de reprodução R estimados. Para isso, utilizam os resultados do trabalho de Hirotsu de 2004, no qual são analisados casos domésticos de Influenza A, eliminando as infecções co-primárias (duplo casos índices) e terciárias (infecções que não pelo caso índice da casa). Tais resultados já foram mencionados e estão na Figura 2, Figura original 2 (a) do trabalho de J. Wallinga & M. Lipsitch. Desses resultados pode-se deduzir um tempo médio entre infecção primária e secundária de $\tau_m = 2,85$ dias com desvio padrão de $\sigma = 0,93$ dias. Em posse de tais dados, somado a $r_m = 0,20$ obtido por C. E. Mills, J. M. Robins e M. Lipsitch para a epidemia de 1918, J. Wallinga & M. Lipsitch estimam $R = 1,57$ ao utilizar o modelo SIR; 1,65 usando o modelo SEIR; e 1,73 ao usar os dados empíricos e a equação 21, talvez a forma mais adequada.

Tais resultados demonstram que a estimativa de R para um determinado r (taxa de crescimento exponencial das incidências no início do processo), um parâmetro extremamente importante para caracterizar a epidemia e avaliar medidas de contenção, depende bastante da escolha (modelo? Empírica?) da distribuição de geração de tempo $\omega(\tau)$, algo que é de difícil determinação precisa na maioria dos casos. J. Wallinga & M. Lipsitch evidenciam que a análise de dados empíricos domésticos para a determinação de $\omega(\tau)$ leva a uma estimativa para R algo menor que a obtida por C. E. Mills, J. M. Robins e M. Lipsitch.

Saber tudo isso sobre uma pandemia que mora há 100 anos atrás tem seu valor para os dias atuais pois, caso nos deparemos com uma epidemia minimamente similar, poderemos imediatamente discutir medidas eficientes de contenção.

11. Covid – 19 de 2019/2020

Um grupo de 9 pesquisadores de várias unidades (Instituto de *Big data*, Centro de ética e Centro de genética humana) da Universidade de Oxford, liderado por Christophe Fraser, publicaram um artigo na revista Science dia 31 de março de 2020. Com o título “*Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing*” (L. Ferretti et al., Science 10.1126/science.abb6936 (2020) – não há ainda numero para página ou volume), aplicam modelagem matemática para avaliar a intervenção de controle baseada no acompanhamento digital dos contatos realizados por indivíduos infectados. O artigo contribui também

com uma discussão sobre Possíveis requisitos éticos requeridos para a realização de uma intervenção desse tipo.

Algumas ponderações que compõe um preâmbulo para o artigo: Ao redor de metade dos casos são reportados; CFP (*case fatality proportion* – mortalidade com respeito ao total de casos) de 2%; 5% dos casos necessitam cuidados intensivos (UTI); Não há ainda tratamento nem vacina; mortalidade maior entre os mais velhos; únicas medidas de contenção conhecidas e aplicáveis são as clássicas: isolamento de infectados, acompanhamento e quarentena para contatos realizados por infectados, distanciamento físico e medidas de higiene; a transmissão pode dar-se via gotículas exaladas pelo infectado, aerossol, superfícies contaminadas e, possivelmente, via contato feco-bucal (fezes-boca). Para quem quer saber mais sobre dados relacionados a COVID 19 no mundo todo sugiro (eu, Kleber) a seguinte página da Organização Mundial da Saúde: <https://www.who.int/>.

Os autores exploram todos os conceitos expostos nesse texto e seguem uma rota teórica, diferente da utilizada nos modelos SIR ou SEIR, para construir um modelo matemático próprio para descrever diretamente a função Infeciosidade $\beta(\tau)$ como função do tempo. Para isso, consideram que a transmissão do SARS-CoV-2 (novo corona vírus) se dá baixo 4 diferentes circunstâncias: i) transmissão sintomática: via contato direto (infectado com sintomas – susceptível saudável); ii) transmissão pré-sintomática: via contato direto (infectado AINDA sem sintomas – susceptível saudável); iii) transmissão assintomática: via contato direto (infectado que NUNCA apresentam sintomas – susceptível saudável); e iv) transmissão ambiental: via contaminação mediada por superfícies.

Os autores defendem que, ao quantificar a frequência com que ocorre cada um dos 4 tipos de transmissão pode ajudar o estabelecimento de medidas de contenção da epidemia.

Para estimar alguns parâmetros importantes os autores utilizaram dados do início da epidemia na China. Como já discutimos nesse texto, suficientemente antes do pico na curva do número de casos em função do tempo, a transmissão ocorre exponencialmente. Com os dados utilizados nesse trabalho, os autores obtiveram uma taxa de crescimento exponencial $r = 0,14$ por dia (1/dia é a unidade de r). Com essa taxa pode-se calcular o intervalo de tempo necessário para que o número de infecções dobre: se o numero de infectados é proporcional a $e^{r\tau}$, o chamado “tempo de

duplicação” é naturalmente dado por $\tau_2 = \ln(2)/r = 5,0$ dias. Ou seja, a cada 5 dias o numero de casos se multiplica por dois.

Os autores fazem a seguinte ponderação: “Uma quantificação robusta e precisa da frequência relativa de ocorrência das diferentes formas de transmissão deveria ser objetivo de trabalhos minuciosos e detalhados, que requer seu tempo e esforço. Porém, a urgência planetária requer estimativas em tempo recorde, obrigando o uso de dados relativamente pobres e de fontes não muito confiáveis”. Nesse contexto, os autores reportam terem feito uma análise detalhada de dados obtidos em processos de transmissão bem definidos e, a partir dela, construíram a famosa função distribuição de geração de tempo. Foram capazes também de atribuir a frequência relativa com a qual ocorreram transmissões pré-sintomática. A partir do ajuste dos dados com o modelo matemático proposto, foram capazes de calcular o valor de R_0 e estimar as proporções (frequências relativas) com as quais ocorreram as diferentes formas de transmissão. Dentro desse contexto foram capazes de discutir sobre a eficácia de medidas como o isolamento de casos identificados ou o acompanhamento dos contatos feitos por infectados em evitar o espalhamento da epidemia.

Uma das informações importantes para o trabalho é o tempo de incubação, definido como o intervalo entre a infecção e o desenvolvimento dos sintomas. A partir de dados já publicados (ver artigo original para maiores detalhes), assumiram que o tempo de incubação segue uma distribuição log-normal com valor médio é de 5,5 dias, mediana de 5,2 dias e desvio padrão de 2,1 dias. Tal distribuição de tempo de incubação está na Figura 5 abaixo (reprodução da Figura 1 do artigo original).

A função distribuição de geração de tempo (tempo entre infecção de um indivíduo e a primeira infecção secundária causada por esse indivíduo), $\omega(\tau)$, foi obtida com a ajuda das informações acima e com a observação de 40 pares de indivíduos fonte – indivíduos receptores especialmente selecionados, para os quais se sabiam com alto grau de certeza como se deu a transmissão direta e a data de surgimento dos sintomas para ambos fonte e receptor. Combinando essas informações com a distribuição de tempos de incubação, construíram a distribuição de geração de tempo $\omega(\tau)$ e, ao ajustar a construção empírica com uma função matemática obtiveram que a distribuição do tipo Weibull (Wikipediável), com tempo médio de 5,0 dias e desvio padrão de 1,9 dias foi a mais fidedigna (critérios matemáticos bem estabelecidos – ver artigo original e material suplementar) e

também está mostrada na Figura 5 . Para efeito de comparação apenas, os autores também mostram, na Figura 5, outras duas funções distribuição de geração de tempo obtidas em outros dois trabalhos, como descrito na legenda.

Para cada um dos 40 pares de indivíduos, os autores estimaram posteriormente a probabilidade da transmissão haver sido pré-sintomática e obtiveram uma média de 37%. Esse número pode ser interpretado como a fração de transmissões pré-sintomáticas entre todas as transmissões. O histograma completo está mostrado na parte direita da Figura 5. As incertezas associadas aos resultados e os testes estatísticos de confiança estão todos detalhados no material suplementar do artigo.

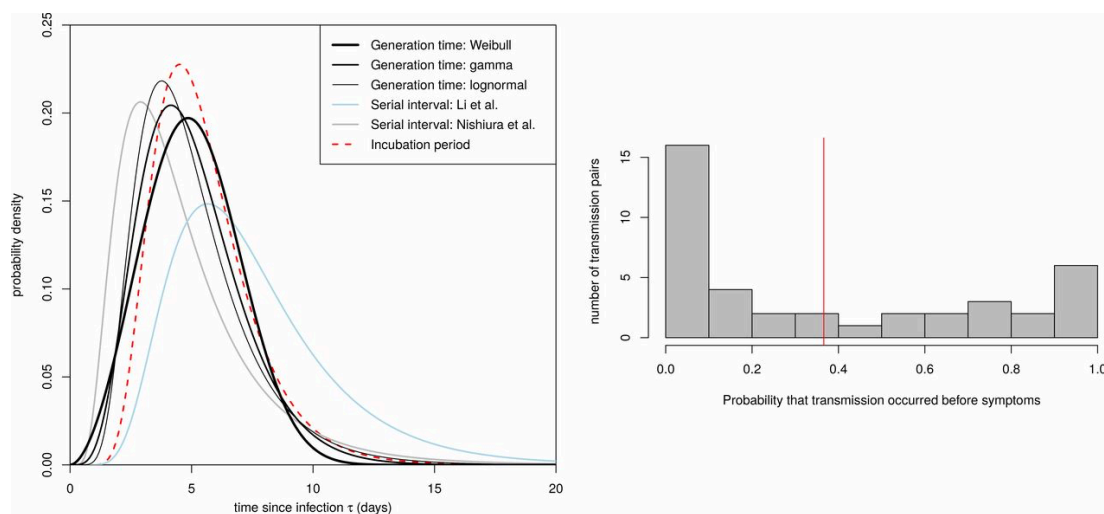


Figura 5. Painel da esquerda: funções distribuição de geração de tempo estão representadas pelas linhas pretas. A mais grossa é a que melhor ajustou os dados empíricos (distribuição Weibull). As linhas pretas de menor grossura representam distribuição Gamma e Lognormal (Wikipediável). Também estão representados no gráfico a distribuição de tempos de incubação (tracejada vermelha) e outros dois ajustes para distribuições de tempo de geração obtidas em outros trabalhos recentes (ver as referencias no artigo original).

Como já sabemos, em posse de todas essas informações os autores podem estimar o valor de R_0 e fazer todas as análises já descrita nesse texto. Porém, para dar mais riqueza as informações obtidas, introduzem antes um modelo matemático próprio, como já dito, que considera as 4 formas mencionadas de transmissão (I a IV) que escreve a seguinte expressão para a Infeciosidade como função do tempo desde

o momento da infecção. Da definição da função $\beta(\tau)$ da sessão 5 sabemos que a área debaixo da curva é numericamente igual ao numero de reprodução R0. A expressão seguinte para $\beta(\tau)$ representa uma média que considera heterogeneidade de indivíduos em vários aspectos (artigo original):

$$\beta(\tau) = P_a \chi_a \beta_s(\tau) + (1 - P_a) [1 - s(\tau)] \beta_s(\tau) + (1 - P_a) [s(\tau)] \beta_s(\tau) + \int_{l=0}^{\tau} \beta_s(\tau - l) E(l) dl$$

(25)

Na expressão 25 vemos a Infeciosidade sendo dada pela soma de 4 termos. Do primeiro ao quarto eles representam, respectivamente: assintomática, pré-sintomática, sintomática e ambiental. Dessa forma, $\beta_s(\tau)$ é a Infeciosidade (sintomática ou pré-sintomática) para o tempo depois da infecção τ ; χ_a é a razão entre as transmissões assintomáticas e sintomáticas; P_a é a fração que representa as transmissões assintomáticas; $s(\tau)$ se refere a distribuição dos tempos de incubação e $E(l)$ representa a taxa de infecções causada por elementos do ambiente (superfícies, por exemplo), contaminadas um tempo l antes do momento τ considerado. Para o leitor que, a duras penas, chegou até aqui e que tem conhecimentos básicos sobre cálculo pode enxergar justificativa matemática para cada termo da expressão 25.

Considerações detalhadas sobre a expressão e sobre os parâmetros utilizados (que não reproduzirei aqui) com as respectivas incertezas (algo de extrema importância) estão claramente expostos no artigo original, tanto no corpo principal quanto no material suplementar. O sucesso dessa abordagem está em poder identificar e quantificar a contribuição de cada termo da função $\beta(\tau)$. Dessa forma, a área debaixo da curva de cada contribuição fornecerá o valioso número médio de infecções ocorridas para cada tipo de transmissão considerada (denominado no artigo por RA (assintomática), RP (pré-sintomática), RS (sintomática), e RE (ambiental). Os resultados estão na Figura 6 (reprodução da Figura 2 do artigo original).

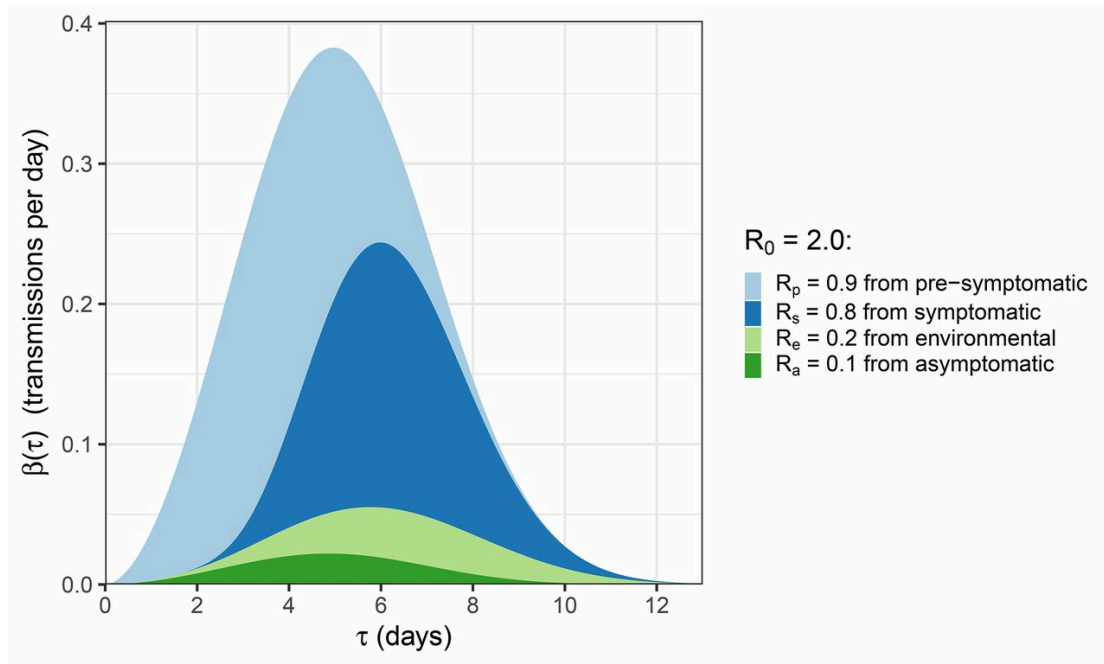


Figura 6. Infeciosidade como função do tempo a partir da infecção (média individual). De acordo com a definição, a soma de todas as áreas coloridas entre dois valores de τ fornece o número total de transmissões nessa janela de tempo, ou seja, o número de reprodução básico R_0 . A área debaixo da curva de uma determinada cor fornece o número de infecções devido a respectiva rota de transmissão (no gráfico as cores estão representadas uma encima da outra).

A partir do gráfico da Figura 6 acima pode-se estimar o valor da razão entre o número de infecções pré-sintomáticas e a soma das pré com as sintomáticas: $R_p / (R_p + R_s)$. O valor obtido de 0,55 (0,37 a 0,72) é substancialmente maior que o avaliado a partir dos 40 pares de indivíduos (0,37 (0,28 a 0,45)) porém, as incertezas associadas a ambas estimativas nos fazem concluir que os resultados concordam entre si (importância das incertezas!!!!).

Na sessão 7, definimos θ , uma quantidade bastante importante para avaliar medidas de intervenção, que consiste na razão entre o número de infecções não sintomáticas pelo número total de infecções. No caso da expressão 23 da sessão 7 só foram consideradas transmissões sintomáticas e pré-sintomáticas. No caso desse artigo devemos considerar transmissões não sintomáticas as assintomáticas e as ambientais. Nossa expressão para θ fica então $\theta = 1 - R_s / R_p$. O valor encontrado foi de 0,62 (0,50 a 0,92).

Os autores se dedicam agora a avaliar (via modelagem) o impacto de duas medidas clássicas de contenção: i) isolamento dos indivíduos sintomáticos e ii) identificação posta em quarentena dos contatos efetuados pelos casos sintomáticos. O controle efetivo da epidemia se dá, como já sabemos, quando as intervenções são capazes de reduzir o número de reprodução R para um valor menor que 1, o que poderia ser conseguido a partir da combinação das duas formas de intervenção aplicadas simultaneamente. O resultados obtidos estão mostrados na Figura 7 (reprodução da Figura 3 do artigo original).

Os gráficos da Figura 7 contêm, de forma compacta, uma quantidade enorme de informações. Eles mostram a eficiência das duas medidas de controle em termos da taxa de crescimento exponencial do número de casos (r). Como um está relacionado ao outro, se tratou apenas de uma opção dos autores, certamente para priorizar a comunicação. A contenção da epidemia significa r menor que 0, ou seja, um crescimento exponencial com taxa negativa (diminuição exponencial com o tempo). Naturalmente, essa condição leva a R menor que 1. Os gráficos devem ser lidos de acordo com o código de cores primeiramente. Quanto mais rosa, maior é o valor de r . No eixo y está o sucesso em colocar em quarentena os contatos efetuados pelos indivíduos sintomáticos e, no eixo x , o sucesso em isolar os casos sintomáticos. A linha preta cheia representa $r = 0$, ou seja, a fronteira entre controle e não controle da epidemia, e as linhas tracejadas representam a incerteza na linha limite $r = 0$ devido a propagação da incerteza no cálculo de R_0 . Ou seja, a linha $r = 0$ é a preta cheia na média porém, pode estar entre as duas linhas tracejadas. Dessa forma, se a combinação dos relativos sucessos nas duas formas de intervenção levar a um ponto acima da linha preta cheia teremos a contenção da epidemia. Caso levem a um ponto abaixo da linha preta cheia, a combinação particular dos sucessos não conterão o espalhamento das transmissões. Tudo isso claro, tendo consciência dos limites das incertezas dadas pelas linhas tracejadas. Vemos também a efetividade das intervenções caem drasticamente com o tempo de atraso em aplicá-las. O gráfico da esquerda é para aplicação das medidas de controle imediatamente após o surgimento dos sintomas enquanto o primeiro gráfico da direita é para 3 dias de atraso.

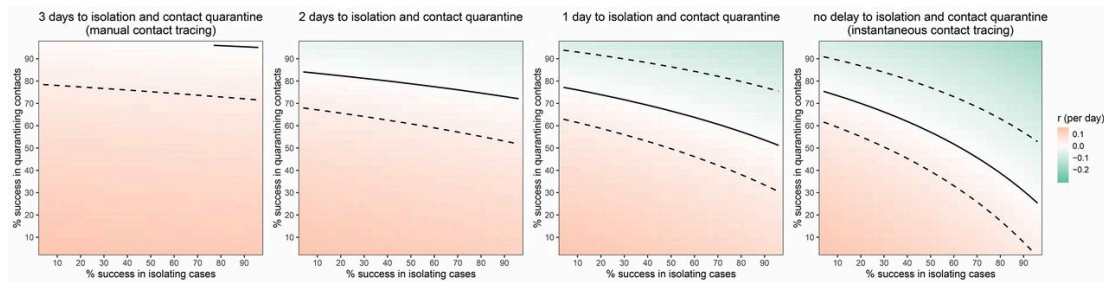


Figura 7. Mapa de calor mostrando a eficácia das intervenções de controle (isolamento dos casos sintomáticos e identificação e quarentena dos contatos feitos por casos sintomáticos) em função dos sucessos relativos de ambas intervenções e do tempo de atraso em aplicá-las (com relação ao momento de surgimento dos sintomas).

Os autores ponderam que atrasos na aplicação das medidas de contenção são inevitáveis porém, podem ser minimizados com a ajuda de um aplicativo de celular, por exemplo, que informa instantaneamente os contatos de um indivíduo que acaba ser identificado como infectado. O aplicativo conservaria gravado os eventos de proximidade entre indivíduos (telefones) de forma a alertar e sugerir quarentena aos contatos mais próximos de forma imediata. Nesse sentido, China implantou um aplicativo com de uso necessário para a utilização de transporte ou acesso a locais públicos. O aplicativo permitia o uso de dados de movimento e, com ajuda de inteligência artificial, cruzá-los com diagnósticos e indicar reforço ou relaxamento de medidas de isolamento. Aplicativo similar foi utilizado pela Coreia do Sul que, assim como China, obtiveram êxito no controle local da epidemia atribuindo-o, em parte, ao uso dos aplicativos.

Após todas as considerações acima e levando em consideração que o estudo aqui apresentado mostra a importância das medidas serem aplicadas o quanto antes, os autores propõe um algoritmo simples para o desenvolvimento de um aplicativo que envolve leitura de códigos QR, uso coordenadas GPS e dados de diagnóstico para alertar indivíduos e indicar, de forma instantânea, procedimentos de quarentena ou isolamento social a possíveis infectados preservando o anonimato do caso identificado. O algoritmo está detalhado no artigo original.

Como o uso de tais aplicativos despertam discussões relacionadas a preservação da privacidade individual, os autores dedicam uma sessão à considerações éticas relacionadas a utilização de tais aplicativos.

Por fim, algumas considerações finais sobre os resultados do trabalho. Primeiramente, ao utilizar dados da própria epidemia SARS-CoV-2 os autores estimam um valor para R_0 menor que os encontrados em outros trabalhos. Menor R_0 implica que uma menor fração de casos deve ser evitada para se conter a epidemia ($R < 1$). Porém, como o valor de θ (0, 62) é bastante alto, significando que uma fração alta das transmissões não são via contato direto com indivíduos sintomáticos, não é tarefa fácil a aplicação de medidas de intervenção que não a vacinação (que ainda não temos). Mais especificamente, os resultados do trabalho, ver Figura 6, apontam que a contribuição para R_0 das transmissões pré-sintomáticas é de 0,9 (0,2 a 1,1) o que poderia sustentar por si só a epidemia (R maior ou igual a 1). Para a SARS essa contribuição é próxima de zero indicando que medidas de intervenção que funcionaram para SARS não funcionarão para a COVID – 19. O alto valor de θ também leva a concluir que o isolamento apenas de indivíduos sintomáticos não é eficiente para a contenção da epidemia.

As conclusões do artigo se referem a dados da epidemia na China. Há evidências de que a transmissão na Europa ocorreram com maior velocidade o que levaria a um valor de R_0 maior para aquele continente.

Os autores disponibilizaram na internet uma interface bastante interessante que permite ao leitor introduzir diferentes parâmetros e obter os resultados aqui analisados (<https://bdi-pathogens.shinyapps.io/covid-19-transmission-routes/>).

Por fim, reproduzo abaixo as considerações finais feitas pelos autores:

“The economic and social impact caused by widespread lockdowns is severe. Individuals on low incomes may have limited capacity to remain at home, and support for people in quarantine requires resources. Businesses will lose confidence, causing negative feedback cycles in the economy. Psychological impacts may be lasting. Digital contact tracing could play a critical role in avoiding or leaving lockdown. We have quantified its expected success and laid out a series of requirements for its ethical implementation. The App we propose offers benefits for both society and individuals, reducing the number of cases and also enabling people to continue their lives in an informed, safe, and socially responsible way. It offers the potential to achieve important public benefits while maximising autonomy. Specific issues exist for groups within the population that may not be amenable to such an approach, and

these could be rapidly refined in policy. Essential workers, such as health care workers, may need separate arrangements”.

Traduzindo, ficaria mais ou menos assim:

“O impacto econômico e social causado por bloqueios generalizados é grave. Indivíduos com baixa renda podem ter capacidade limitada para permanecer em casa, e o apoio a pessoas em quarentena requer recursos. As empresas perderão segurança, causando ciclos negativos na economia. Os impactos psicológicos podem ser duradouros. O rastreamento digital de contatos pode desempenhar um papel crítico para evitar ou deixar o bloqueio. Quantificamos o sucesso esperado e estabelecemos uma série de requisitos para sua implementação ética. O aplicativo que propomos oferece benefícios para a sociedade e para os indivíduos, reduzindo o número de casos e também permitindo que as pessoas continuem suas vidas de maneira informada, segura e socialmente responsável. Oferece o potencial de obter importantes benefícios públicos, maximizando a autonomia. Questões específicas podem existir para determinados grupos dentro da população e devem ser analisados separadamente. Trabalhadores essenciais, como profissionais de saúde, podem precisar soluções diferenciadas”.

12. Conclusões

Esse texto teve por objetivo final (quando comecei escrever não imaginaria que ficasse desse tamanho) apresentar de forma didática os conceitos matemáticos básicos utilizados na análise de epidemias. Com isso, pretende também mostrar as dificuldades em determinar parâmetros importantes para uma boa modelagem e, ao mesmo tempo, evidenciar o quão sensíveis podem ser as previsões (ou avaliações sobre medidas de intervenção) em relação a qualidade da fonte de dados (e dos dados em si) tomadas para alimentar os modelos. Mais uma vez sinto a necessidade de afirmar minha ignorância sobre o tema. As especificidades das aplicações da matemática aqui apresentadas foram aprendidas a medida que lia os artigos consultados. Isso faz com que esse texto, após críticas e sugestões, seja direcionado, no máximo, a estudantes em início de graduação em ciências. Gostaria que o texto também fosse útil para um público mais amplo além do familiarizado com a

matemática envolvida. Embora creia não haver conseguido de forma plena, espero que o leitor que resolva passar por cima das expressões matemáticas tenha, igualmente, tirado alguma informação interessante. É importante dizer que há tantos estudos semelhantes ao apresentando quantos grupos de pesquisa sobre o tema existem no mundo. Dessa forma, considerar um ou outro no qual se basear ao adotar medidas públicas de enfrentamento a epidemia é outro desafio a ser levado em conta pelas autoridades. Sem dúvida, saber ler criticamente um artigo científico ou um informe técnico é de suma importância pois os resultados apresentados, apesar do “valor científico”, apresentam confiabilidade segundo as hipóteses e considerações certamente colocadas de forma clara no texto. A origem e qualidade da fonte de dados influenciam diretamente na quantificação das incertezas nas conclusões finais, informação tão importante quanto o resultado em si. Em uma possível segunda parte desse artigo podemos discutir outros trabalhos interessantes que chamaram atenção do público em geral, como o artigo publicado na revista *Science* em 14 de abril desse ano, que prevê isolamento social intermitente até 2022, explorado pela mídia sem sequer ser citado de forma apropriada (S. M. Kissler et al. *Science* 14 Apr 2020: eabb5793. DOI: 10.1126/science.abb5793). Outro trabalho que demanda comentários é o famoso “estudo do Imperial College” de Londres, publicado dia 26 de março desse ano (<https://www.imperial.ac.uk/media/imperial-college/medicine/sph/ide/gida-fellowships/Imperial-College-COVID19-Global-Impact-26-03-2020.pdf>), cuja previsão de mais de meio milhão mortos (levando em conta isolamento social) esperamos nunca ser concretizada.

13. Principais trabalhos consultados, todos disponíveis gratuitamente a qualquer pessoa com acesso a internet.

1. Grassly & Fraser, *Mathematical models of infectious disease transmission*, *Nature reviews microbiology* n. 6, pp. 477, 13 de maio de 2008.
2. Wallinga & Lipsitch, *How generation intervals shape the relationship between growth rates and reproductive numbers*, *Proceedings of the Royal Society* n. 274, pp. 599, 28 de novembro de 2008.
3. Ma & Earn, *Generality of the final size formula for an epidemic of a newly invading infectious disease*, *Bulletin of Mathematical Biology* n. 68, pp. 679, 8 de abril de 2006.

4. Mills, Robins & Lipsitch, *Transmissibility of 1918 pandemic influenza*, *Nature* n. 432, pp. 904, 16 de dezembro de 2004.
5. L. Ferretti et al., *Science* 10.1126/science.abb6936 (2020). Publicado *on line* dia 31 de março, ainda não dispõe de página e volume porém, pode ser lido por qualquer pessoa na página da revista.