SARS-CoV-2 (COVID-19) em números

Yinon M. Bar-On¹, Avi Flamholz², Rob Phillips^{3,4}, e Ron Milo ^{1*}

¹ Weizmann Intitute of Science, Rehovot 7610001, Israel ² University of California, Berkeley, CA 94720, USA

https://elifesciences.org/articles/57309

³ California Institute of Technology, Pasadena, CA 91125, USA ⁴ Chan Zuckerberg Biohub, San Francisco, CA 94158, USA

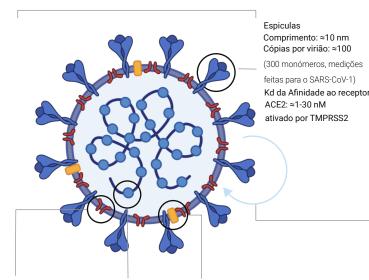
* Autor correspondente: ron.milo@weizmann.ac.il.

Aceitamos sugestões: esse artigo está sendo atualizado frequentemente em:

https://bit.ly/2W0eN64

Tamanho e Conteúdo

Diâmetro: ≈100 nm Volume: $\sim 10^6 \text{ nm}^3 = 10^{-3} \text{ fL}$ Massa: ~10³ MDa ≈ 1 fg



Proteína de membrana Nucleoproteínas ≈2000 cópias (medições feitas para

o SARS-CoV-1)

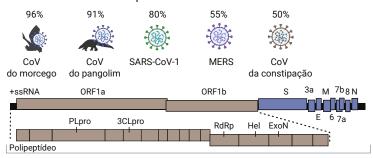
≈1000 cópias

Proteína de envelope ≈20 cópias

(medições feitas (100 monómeros, medições para o SARS-CoV-1) feitas para o TGEV coronavírus)

Genoma

Identidade Nucleotídica para com o SARS-CoV-2



Tamanho do genoma: ≈30kb; β-coronavirus com 10-14 ORF (24-27 proteínas)

 $\sim 10^{-3} \text{ nt}^{-1} \text{ ano}^{-1}$ (medições feitas para o SARS-CoV-1) Taxa evolutiva: Taxa de mutação: ~10⁻⁶ nt⁻¹ ciclo⁻¹ (medições feitas para o MHV coronavírus)

Escala de tempo da replicação

Cálculos feitos em cultura celular

Entrada do virião na célula: ~10 min (medições feitas para o SARS-CoV-1) Período de eclípse: ~10 horas (tempo para fazer viriões intracelulares) Liberação do vírus: ~10³ viriões (medições feitas para o MHV coronavírus)

Células hospedeiras

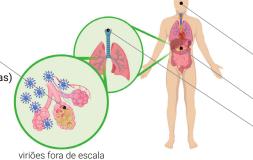
(lista provisória, número de células por pessoa)

Pneumócitos de tipo I e tipo II (~10¹¹ células)

Macrófagos alveolares (~10¹⁰ células)

Células calciformes na cavidade nasal (~109 células)

Volume da célula hospedeira: $\sim 10^3 \, \mu m^3 = 10^3 \, fL$



Concentração

Valores máximos observados após diagnóstico (Woelfel et al. 2020; Kim et al. 2020; Pan et al. 2020)

Nasofaringe: 10⁶-10⁹ ARNs/cotonete Garganta: 104-108 ARNs/cotonete

Fezes: 104-108 ARNs/g

Expectoração: 10⁶-10¹¹ ARNs/mL A contagem de ARN pode superestimar a presença de viriões infecciosos

Resposta do Anticorpo - Seroconversão

Os anticorpos aparecem no sangue depois de: ≈10-20 dias

Manutenção da resposta do anticorpo: **≈2-3 anos** (medições feitas para o SARS-CoV-1)

Estabilidade ambiental do virus

Relevância para a segurança pessoal

tempo para decair meia-vida 1000 vezes ≈1 hr ≈4-24 hr Aerossois: Superficies: ≈4-96 hr E.g. plásticos. (van Doremalen et al. 2020) Cartao (papel) e metais

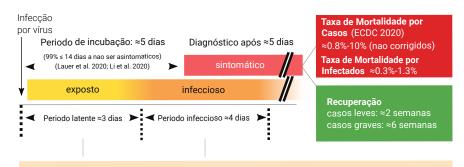
Baseado em quantificações de viriões infecciosos. Testado a 21-23°C e 40-65% umidade relativa. Números vão variar entre condições e tipo de superfícies (Otter et al. 2016). RNA viral observado em superfícies mesmo após algumas semanas (Moriarty et al. 2020).

Padrão "Típico" de Progressão da Infecção em um Único Paciente

Número básico de reprodução, Ro: geralmente 2-4

mas com variação temporal e espacial (Li et al. 2020; Park et al. 2020)

(número de novos casos originados a partir de um unico caso)



Existe uma variação substancial entre indivíduos que ainda não está caracterizada. Essas estimativas são de acordo com a mediana da população da China e não descreve essa variabilidade (Li et al. 2020; He et al. 2020).

"Notar a diferença entre os símbolos ≈, que denota "aproximadamente" com variação por um fator de 2, e o símbolo ~, que indica "ordem de magnitude" or precisão por um fator de 10.

Translated by: Clarice de Azevedo Souza & Alicia Calvo-Villamañán

SARS-CoV-2 (COVID-19) em números

Yinon M. Bar-On¹, Avi Flamholz², Rob Phillips^{3,4}, e Ron Milo¹
¹Weizmann Institute of Science, Rehovot 7610001, Israel²University of California, Berkeley, CA 94720, USA
³California Institute of Technology, Pasadena, CA 91125, USA
⁴Chan Zuckerberg Biohub, 499 Illinois Street, SF CA 94158, USA

Resumo

A atual pandemia causada pelo SARS-CoV-2 é um lembrete severo do fato de que, seja à escala de um único hospedeiro humano ou de uma onda de infecção intercontinental, a dinâmica viral é frequentemente uma história sobre os números. Neste contexto, o nosso objetivo é fornecer uma fonte gráfica para os principais números que possam ajudar a perceber o vírus que está por trás da atual crise global. A discussão está enquadrada em torno a dois grandes temas: 1) a biologia do próprio vírus e 2) as características da infecção do mesmo num único hospedeiro humano. Este pequeno resumo apresenta os principais números relativos ao SARS-CoV-2, tendo como base principalmente artigos na literatura que foram submetidos a revisão por pares. Todos os números apresentados serão acompanhados por referências aos artigos de onde são originários. Recomenda-se ao leitor ter em mente que existe muita incerteza relativamente ao conhecimento desta pandemia, uma vez que o vírus que a causa está em rápida evolução. Nos seguintes parágrafos forneceremos cálculos rápidos que servem de exemplo para o tipo de discernimento que pode ser obtido a partir de alguns números-chave e o uso da lógica quantitativa. Estes cálculos podem servir para melhorar a nossa intuição acerca da pandemia, mas não podem nem devem substituir análises epidemiológicas detalhadas.

1. Quanto tempo demora para que uma única pessoa produza um milhão de pessoas infectadas?

Se todos continuássemos a nos comportar como de costume, quanto tempo demoraria a pandemia em se espalhar de uma única pessoa a um milhão de vítimas infectadas? O número básico de reprodução do vírus, Ro, sugere que cada infecção gera diretamente 2-4 novas infecções, isto na ausência de contramedidas tais como o distanciamento social. Uma vez que uma pessoa é infectada, existe um período conhecido como período latente antes de que ela possa transmitir o vírus a outras pessoas. A melhor estimativa atual do tempo mediano de latência é de ± 3 dias, aos que se seguem ± 4 dias de infecciosidade no seu expoente máximo (Li et al. 2020, He et al. 2020). A duração exata varia entre as pessoas, e algumas são infecciosas durante muitos mais tempo. Usando portanto um R₀≈4, sabemos que o número de casos quadruplica a cada 7 dias, ou seja, duplicada a cada 3 dias. Para crescer 1000 vezes (isto é, passar de 1 caso para 103) requer 10 duplicações, uma vez que 2¹⁰ ≈ 10³; 3 dias x 10 duplicações = 30 dias, ou cerca de um mês. Portanto, esperamos um crescimento de 1000x num mês, um milhão de vezes (106) em dois meses e um bilhão de vezes (109) em três meses. Embora estes cálculos sejam altamente simplificados, e não tenham em conta o efeito dos "super portadores" ou da imunidade de grupo ou de testes incompletos, enfatizam de por si a fato que este vírus pode espalhar-se a um ritmo desconcertante quando nenhuma medida é tomada para a sua disseminação. Isto ilustra o por quê de ser crucial limitar a propagação do vírus por medidas de distanciamento social. Para uma discussão mais ampla acerca do significado do R₀, os períodos de latência e infecção, assim como várias outras advertências, consulte a seção "Definições".

2. Qual é o efeito do distanciamento social?

O seguinte exemplo quantitativo altamente simplificado ajuda a esclarecer a necessidade do distanciamento social. Suponha que o leitor esteja infectado e encontre 50 pessoas ao longo de um dia de trabalho, descolamentos, socialização e execução de tarefas diárias. Para arredondar os números, suponhamos ainda que o leitor tem uma probabilidade de 2% de transmitir o vírus em cada um desses encontros, o que significa que o leitor pode infectar uma nova pessoa por dia. Se o leitor for infeccioso durante 4 dias, poderá infectar 4 novas pessoas em média. Isto está no limite superior dos valores de Ro para o SARS-CoV-2 na ausência de distenciamento social. Se o leitor reduzir o número de pessoas com quem está emcontacto por dia a 5 (preferencialmente menos) por causa do distanciamento social, pasará a infectar 0,1 pessoas por dia, ou 0,4 pessoas no total antes de se tornar menos infeccioso. O efeito desejado do distanciamento social é fazer com que cada infecção atual produza <1 novas infecções. Um número de reprodução efetivo (R_e) inferior a 1 garante que o número de infecções diminua eventualmente com o tempo. É extremamente importante alcançar o mais rapidamente possível um R_e<1, o que é substancialmente mais viável do que levar o R_e a quase zero por meio de medidas de saúde pública.

3. Por quê o período de quarentena dura duas semanas ?

O período de tempo entre a infecção e a aparição dos primeiros sintomas é denominado período de incubação. A estimação do período médio de incubação do SARS-CoV-2 é de aproximadamente 5 dias (Lauer et al. 2020). No entanto, existe muita variação de pessoa a pessoa. Aproximadamente 99% das pessoas que apresentam sintomas antes do dia 14, o que explica que o período de confinamento seja de duas semanas. É importante ressaltar que esta análise ignora as pessoas infectadas que nunca apresentam sintomas. Como as pessoas assintomáticas geralmente não são testadas, ainda não está claro quantos casos existem ou por quanto tempo as pessoas assintomáticas permanecem infecciosas.

4. Como é que as máscaras N95 bloqueiam a entrada do SARS-CoV-2?

As máscaras N95 foram desenhadas para filtrar mais de 95% de todas as partículas com pelo menos 0,3 mícron (μm) de diâmetro (NIOSH 42 CFR Part 84). De fato, as medições da eficiência de filtragem de partículas por parte das máscaras N95 demonstram que elas são capazes de filtrar ≈99.8% das partículas com um diâmetro ~0,1μm (Regnasamy et al. 2017). O SARS-CoV-2 é um vírus rodeado de

um envelope viral com aproximadamente 0,1µm de diâmetro, pelo que as máscaras N95 são capazes de filtrar a maioria dos viriões livres. No entanto, estas máscaras fazem mais do que isso. Os vírus são frequentemente transmitidos através de gotículas respiratórias produzidas pela tosse o os espirros. Estas gotículas são geralmente divididas em gotas grandes (>5µm de diâmetro), que caem rapidamente para o chão e portanto são transmitidas apenas em curtas distâncias, e pequenas gotas (<5µm de diâmetro). Estas pequenas gotas podem evaporar-se em "núcleos de gotículas" e permanecer suspensas no ar por períodos significativos de tempo, e portanto podem ser inaladas. Alguns vírus, o do sarampo por exemplo, podem ser transmitidos por núcleos de gotículas (Tellier et al. 2019). Atualmente, não existe uma evidência direta de que o SARS-COV-2 possa ser transmitido por núcleos de gotículas. Em vez disso, acredita-se que as gotas grandes sejam o principal vetor da sua transmissão, geralmente instalando-se em superfícies, que são tocadas e transportadas pelas mãos até as membranas mucosas, como olhos, nariz e boca (CDC 2020). O diâmetro das gotículas grandes produzidas por espirros é de ~100µm (Han J. R. Soc. Interface 2013), enquanto que o diâmetro dos núcleos de gotículas produzidas pela tosse é da ordem de ~1µm (Yang et al 2007). Em conclusão, as máscaras N95 provavelmente protegem contra vários modos de transmissão viral.

5. O quão parecido é o SARS-CoV-2 aos vírus do resfriado comum e da gripe?

O SARS-CoV-2 é um beta-coronavirus cujo genoma é composto de uma única fita de RNA, de ≈30 kb. Já a gripe é causada por um vírus que pertence a uma família diferente de vírus de RNA, chamada influenza. Os virus que causam a gripe tem genomas menores (≈14 kb) codificados por 8 segmentos de RNA, e infectam células humanas de maneira diferente dos coronavirus. O resfriado comum é causado por uma variedade de vírus, incluindo alguns coronavirus e rinovírus. Os coronavirus que causam o resfriado comum (ex. cepas OC43 and 229E) são mais parecidos ao SARS-CoV-2 no tamanho do genoma (com variância de 10%) e conteúdo gênico, mas diferente do SARS-CoV-2 em seguência (≈50% de identidade nucleotídica) e severidade de infecção. Um aspecto interessante dos coronaviruses é que possuem o maior genoma dentre os vírus de RNA conhecidos (≈30 kb). Esses genomas grandes levam cientistas a acreditar na presença de um mecanismo de correção enzimática ("proofreading mechanism") que reduz a taxa de mutação e estabiliza o genoma. De fato, os coronavirus possuem uma enzima exonuclease de correção chamada ExoN, o que explica sua taxa de mutação reduzida (~10-6 por base por ciclo) quando comparada ao influenza (≈3×10-5 por base por ciclo (Sanjuan et al. 2010)). Essa taxa e mutação relativamente baixa será de interesse em futuras pesquisas para prever a velocidade em que coronavirus poderão evadir os esforços de vacinação.

6. O quanto se sabe sobre o genoma e proteoma do SARS-CoV-2?

SARS-CoV-2 possui um genoma de RNA de fita simples positiva que codifica 10 genes e produz 26 proteins, conforme anotação do NCBI (NC 045512). Como que 10 genes codificam >20 proteínas? Um gene comprido, orf1ab, codifica um polipeptídeo que é clivado em 6 proteínas por uma protease proveniente do mesmo polipeptídeo. Além de proteases, o polipeptídeo codifica uma RNA polimerase e outros fatores associados para copiar o genoma, uma exonuclease de correção de transcrição e várias outras proteínas não estruturais. Os genes restantes codificam predominantemente componentes estruturais do vírus: (i) as espículas virais que se ligam ao receptor celular específico na membrana na célula humana ou animal; (ii) uma nucleoproteína que recobre o genoma; e (iii) duas proteínas de membrana. Apesar de inúmeros estudos estarem focados em entender as funções das proteínas "acessórias" no ciclo de vida do vírus, estima-se que até então seja somente possível atribuir uma função bioquímica ou estrutural definida a metade dos produtos gênicos do SARS-CoV-2.

7. O que podemos aprender sobre a taxa de mutação do vírus?

No estudo de evolução de vírus, pesquisadores geralmente usam duas medidas que descrevem a taxa de variação genética. A primeira é a taxa evolutiva, que é definida pela média de substituições que se fixam por ano nas cepas dos vírus, dada em unidades de mutações por sítio de nucleotídeo por ano. A segunda é a taxa de mutação, que é o número de substituições por sítio por ciclo de replicação viral. Como que podemos relacionar esses dois valores? Considere um único sítio no final de um ano. A única medida de taxa de mutação em um β-coronavirus sugere que esse sítio irá acumular ~10⁻⁶ mutações em cada ciclo de replicação. Cada ciclo de replicação demora ~10 horas, portanto 103 ciclos/ano. Multiplicando a taxa de mutação pelo número de replicações, e ignorando os potenciais efeitos de seleção evolutiva e deriva genética, resulta em 10⁻³ mutações por sítio por ano, um número consistente com a taxa evolutiva inferida pelo sequenciamento de genomas de coronavírus. Como essa estimativa é consistente com a taxa medida, estima-se que o vírus está sujeito à replicação quase contínua no meio ambiente, gerando mutações constantemente, que se acumulam ao longo do ano. Usando o conhecimento da taxa de mutação, pode-se também inferir sobre uma única infecção. Por exemplo, como a taxa de mutação é de ~10-6 mutações/sítio/ciclo e um mL de escarro pode conter mais de 10⁷ RNAs virais, pode-se inferir que todos os sítios foram mutados mais de uma vez em tais amostras.

8. Qual é a estabilidade e capacidade de transmissão do vírus em superfícies?

O RNA do SARS-CoV-2 tem sido detectado em varias superficies muitas semanas depois de serem tocadas (Moriarty et al. 2020). Nas definições, nos esclarecemos a diferença em se detectar RNA viral e vírus ativos. A probabilidade de infecção humana provinda de tal exposição ainda não está caracterizada, haja visto que experimentos para fazer essa determinação são muito difíceis. Logo, medidas de precaução devem ser tomadas. Nós estimamos que durante o período infeccioso, uma pessoa com infecção ainda não diagnosticada encoste em superfícies diversas inúmeras vezes. Em seguida, essas superfícies serão tocadas por centenas de outras pessoas. De acordo com o número básico de reprodução $R_{\rm 0}$ $\approx 2-4$ pode-se inferir que nem todas as pessoas que encostem nessas superfícies serão infectadas. Logo, precisa-se com urgência estudos relativos ao risco de infecção por contato com superfícies.

Glossário

Medidas clinicas

Periodo de incubação : tempo entre a exposição e os sintomas.

Seroconversão : tempo entre a exposição ao vírus e a detecção de

Seroconversão : tempo entre a exposição ao vírus e a detecção de resposta dos anticorpos.

Inferencias epidemiologicas

 $\mathbf{R}_{\!\scriptscriptstyle{0}}\!\!:$ número médio de casos directamente gerados por uma infecção individual.

Periodo latente: tempo entre a exposição e a infecção.

Periodo infeccioso: tempo para o qual o indivíduo é infectado.

Interval of half-maximum infectiousness: o intervalo de tempo durante o qual a probabilidade de transmissão viral é maior que metade do pico de infecciosidade. Este intervalo é semelhante ao período infeccioso, mas aplica-se também em casos onde a probabilidade de infecção não é uniforme no tempo.

Espécies virais

SARS-CoV-2: Síndrome respiratorio agudo grave coronavirus 2. Um β -coronavírus está a causar o presente surto de COVID-19.

SARS-CoV-1: β-coronavirus que causou em 2002 o surto SARS na China..

MERS: um β -coronavírus que causou um surto de Síndrome Respiratória do Médio Oriente na Jordânia em 2012..

MHV: o vírus do herpes murino, um modelo de β -coronavírus no qual foram realizadas muitas pesquisas laboratoriais.

TGEV: virus de gastroenterite transmissivel, um modelo α -coronavírus que infecta porcos.

229E and OC43: duas cepas de coronavírus (α - e β - respectivamente) que causam uma fracção da gripe comum.

Ciclo-Vida do Virus

Periodo de eclipse: tempo entre a entrada do vírus e a presença de viriões intracelulares.

Periodo latente (nível celular): tempo entre a entrada do vírus e a presença de viriões extracelulares. Não confundir com o epidemiológico período latente descrito em baixo.

Tamanho de ruptura: número de viriões produzidos a partir da infecção de uma única célula. Mais apropiadamente, nos referimos ao "número de vírus liberados por celular infectada" para virus não-litico como o SARS-CoV-2.

Vírion ou virião: uma particula viral.

Poliproteina: uma proteína longa que é proteoliticamente clivada em proteínas distintas. Diferente de um polipéptido, que é uma cadeia linear de aminoácidos que formam uma proteína.

Biologia Humana

Macrofago alveolar: células imunes encontradas nos pulmoes que envolvem materiais estranhos como pós e micróbios ("fagocitos profissionais")

Pneumócitos: células não imunes no pulmão.

ACE2: Angiotensin-converting enzyme 2, superficie receptora da célula de mamifero à qual se liga o SARS-CoV-2.

TMPRSS2: Protease transmembranar, serina 2, uma serina protease ligada a membrana de mamíferos que cliva a espícula viral após ligação com ACE2, revelando um peptídeo de fusão que participa na fusão membranar que permite a injeção subsequente de DNA viral no citoplasma hospedeiro.

Nasofaringe: o espaço acima do palato mole atrás do nariz que liga o nariz à boca.

Vota

Repare na diferença de anotação entre o símbolo \approx , que indica "aproximadamente" e tem uma precisão dentro de um factor 2, e o símbolo \sim , que indica "ordem de magnitude" ou precisão dentro de um factor 10.

Mais sobre definições e métodos de medição Qual é o significado do R₀, do "período latente" e do "período infeccioso"?

O número básico de reprodução, $R_{\scriptscriptstyle 0}$, é a estimação de o número médio de novas infecções que são geradas directamente por uma única pessoa infecciosa. O índice O indica que é em referência aos estágios iniciais de uma epidemia, quando todos os indivíduos na região são suscetíveis (isto é, não existe ainda imunidade) e ainda nenhuma medida de prevenção foi tomada. Como a geografía e a cultura afetam aspectos tais como quantas pessoas encontramos diariamente, o quanto as tocamos e compartilhamos comida com elas, as estimativas de Ro podem variar entre os locais. Além disso, como o Ro é definido na ausência de medidas de prevenção e imunidade, geralmente só podemos avaliar o R efetivo (Re). NO inicio de uma epidemia, antes de qualquer medida de prevenção, $R_E \approx R_0$. Vários dias tem de passar para que uma pessoa recém-infectada se torne infecciosa. Esse período é conhecido como "período latente" e é normalmente seguido por vários dias de infecciosidade chamados "período infeccioso". É importante perceber que os valores oficiais para todos estes parâmetros são médias populacionais inferidas a partir de modelos epidemiológicos adequados às contagens de pacientes infectados, sintomáticos e moribundos. Como estes testes são sempre incompletos, e o ajuste do modelo imperfeito, os dados variam entre locais diferentes, isto é, há uma incerteza substancial associada aos valores relatados. Além disso, estes valores médios não descrevem a variação de pessoa para pessoa. Por exemplo, o RNA viral foi detectável em pacientes com sintomas moderados por >1 semana a partir do início dos sintomas e mais de 2 semanas após em pacientes com sintomas graves (ECDC 2020). Embora o RNA detectável não seja o mesmo que o vírus ativo, essa evidência exige cautela no uso de parâmetros médios incertos para descrever uma pandemia. Por que não há distribuições oficiais detalhadas desses parâmetros entre pessoas? A medição direta de períodos latentes e infecciosos a nível individual é extremamente desafiante, uma vez que identificar com precisão o tempo exato da infecção é geralmente muito difícil.

Qual é a diferença entre as medições de RNA viral e a presença de vírus infecciosos?

Para fazer o diagnóstico e a quantificação do vírus podemos utilizar várias metodologias diferentes. Uma abordagem comum é quantificar a quantidade de RNA viral presente numa amostra ambiental (por exemplo, uma superfície) ou clínica (por exemplo, um escarro) por meio de uma técnica chamada reação quantitativa em cadeia da polimerase com transcrição reversa (RT-qPCR). Este método mede o número de cópias do RNA viral presente numa amostra. A presença de RNA viral não implica necessariamente a presença de viriões infecciosos. Os viriões podem ser defectuosos (por exemplo, devido a mutações) ou podem ter sido inativados por condições ambientais. Para avaliar a concentração de vírus capazes de provocar infecção, os investigadores geralmente medem a "dose capaz de infectar 50% da cultura de tecidos" (TCID₅₀). Medir o TCID₅₀ envolve infectar culturas de células suscetíveis com diluições do vírus e observar a diluição na qual metade das células são infectadas. As contagens virais relatadas pelo TCID₅₀ tendem a ser muito mais baixas do que as medições por RT-qPCR, o que pode ser uma das razões pelas quais os estudos baseados em RNA (Moriarty et al. 2020) relatam a persistência do RNA viral nas superfícies por muito mais tempo do que os estudos baseados no TCID₅₀ (van Doremalen et al. 2020). É importante ter isso em conta ao interpretar os dados sobre cargas viráis, por exemplo, em casos como um relatório que mede o RNA viral em amostras de fezes de pacientes durante vários dias após a sua recuperação (We et al. 2020). No entanto, no caso de muitos vírus, mesmo uma pequena dose de vírus pode levar à infecção. Para a constipação comum, por exemplo, -0,1 de $TCID_{50}$ é suficiente para infectar metade das pessoas expostas (Couch et al. 1966).

Qual é a diferença entre a taxa de mortalidade por casos e a taxa de mortalidade por infecções?

Estatísticas globais em infecções e mortalidade estão a aparecer de muitos países, fornecendo diferente visões na gravidade e progressão da pandemia. A avaliação da gravidade da pandemia é crítica para decisões políticas e muito esforço tem sido colocado na quantificação. A medida mais comum para a gravidade da doença é a taxa de mortalidade. Uma das medidas mais comuns apresentada é a taxa de mortalidade por casos (TMC), que é a proporção de mortes no total de caso diagnosticados. Os valores da TMC em diferentes países variam significativamente , de 0.3% até 10%. Vários factores-chave afectam a TMC. Primeiro, parâmetros demográficos e prácticas associadas ao aumento ou diminuição de risco, diferem bastante entre sociedades. Por exemplo, a prevalência de fumar, a idade média da população, e a capacidade do sistema de saúde. De facto, a maioria das pessoas

que morrem com o SARS-CoV-2 tem uma condição preexistente como por exemplo doenças vasculares ou histórico de fumar (China CDC 2020). Existe ainda o potencial das estimativas de TMC serem influenciadas. Por exemplo, a tendência de se identificar os casos mais severos (viés de seleção) tende a sobrestimar a TMC. Do outro lado, existe uma demora entre o aparecimento dos primeiros dos sintomas e a morte, o que pode subestimar a TMC no início da evolução de uma epidemia. Mesmo corrigindo para esses fatores, a TMC não reflete a total realidade já que muitos casos leves ou assintomáticos não são testados. Logo, a TMC tende a sobrestimar a taxa de mortalidade por pessoa infectada, chamada de taxa de mortalidade por infecções (TMI). Para se estimar o número total de pessoas infectadas, precisa-se testar uma amostra aleatória para a presença de anticorpos virais, cuja presença indica exposição prévia ao vírus. Até quando isso foi escrito, esse testes ainda não estão amplamente disponíveis. Logo, pesquisadores se utilizam de banco de dados de estrangeiros provindos de países com alto número de infecções (Verity et al. 2020), ou modelos epidemiológicos onde se estima o número de casos não documentados (Li et al. 2020). Esses métodos resultam em um quadro aproximado da real gravidade da doença.

Qual é o número de vírus liberados por uma célula infectada e tempo de replicação viral?

Duas características importantes do ciclo de vida viral é o tempo para gerar novos vírus e o número de novos vírus que cada célula hospedeiro produz. A quantidade de novos vírus produzidos por cada célula infectada é mais facilmente definida em vírus líticos, como os que infectam bacterias (bacteriofagos), dado que os vírus se multiplicam dentro da célula hospedeira, a qual rompe para liberar os novos vírus. No entanto, o SARS-CoV-2 não libera seus descendentes por ruptura da célula hospedeira, mas por brotamento (Park et al. 2020). Apesar de não haver lise celular, ainda pode-se estimar o número médio de partículas virais produzidas por uma única célula. É extremamente difícil medir o tempo de duração de um ciclo de replicação viral, ou medir o número de vírus liberados in vivo, logo pesquisadores estão limitados a fazer essas medições em cultura de tecidos. A maneira mais simples e utilizada para medir essa dinâmica de crescimento viral é garantir que apenas um ciclo reprodutivo ocorra. Isso é obtido infectando células com um grande número de partículas virais, para que todas as células sejam infectadas, e nenhuma célula seja infectada por infecções secundárias. Supondo que os vírus entre nas células rapidamente (estima-se 10 minutos para o SARS-CoV-2), o tempo que demora para que novos vírus sejam gerados pode ser estimado quantificando o período entre a inoculação inicial e o aparecimento de novos vírions intracelulares. Esse período também é chamado de "período de eclipse". O período de eclipse não leva em conta o tempo que demora até que as novas partículas virais sejam liberadas. O período entre a entrada de um vírus na célula hospedeira e o aparecimento de novos vírus extracelulares, chamado de "período latente" (que difere do período latente epidemiológico, veja o Glossário), estima a total duração do ciclo de replicação viral. O número de vírus liberados por uma célula hospedeira pode ser estimado após saturação na produção de

novos vírus, dividindo o número de partículas virais pelo número de células infectadas. Apesar de que tanto o número de partículas virais geradas por célula infectada quanto a duração da replicação viral poderem variar significativamente dependendo do animal hospedeiro, de acordo com o tipo de célula alvo e ação do sistema imunológico, esses números servem para se ter uma ideia aproximada do ciclo de vida do vírus a nível celular.

References and excerpts

Note that for about 10 out of 45 parameters, the literature values are from other coronaviruses. We await corresponding measurements for SARS-CoV-2.

Size & Content

<u>Diameter</u>: (Zhu et al. 2020) - "Electron micrographs of negative-stained 2019-nCoV particles were generally spherical with some pleomorphism (<u>Figure 3</u>). Diameter varied from about 60 to 140 nm." Volume- Using diameter and assuming the virus is a sphere

Mass: Using the volume and a density of ~ 1 g per mL

Number of surface spikes trimers: (Neuman et al. 2011) - "Our model predicts ~90 spikes per particle."

<u>Length of surface spikes trimers</u>: (Zhu et al. 2020) - "Virus particles had quite distinctive spikes, about 9 to 12 nm, and gave virions the appearance of a solar corona."

Receptor binding affinity (K_a): (Walls et al. 2020) - Walls et al. reports K_d of ≈ 1 nM for the binding domain in Table 1 using Biolayer interferometry with k_{on} of $\approx 1.5 \times 10^5 \, M^3 \, s^3$ and k_{off} of $\approx 1.6 \times 10^4 \, s^3$. (Wrapp et al. 2020) - Wrapp et al. reports K_d of ≈ 1.5 nM for the spike (Fig.3) and ≈ 35 nM for the binding domain (Fig.4) using surface plasmon resonance with k_{on} of $\approx 1.9 \times 10^5 \, M^3 \, s^3$ and k_{off} of $\approx 2.8 \times 10^3 \, s^3$ for the spike and k_{on} of $\approx 1.4 \times 10^5 \, M^3 \, s^3$ and k_{off} of $\approx 4.7 \times 10^3 \, s^3$ for the binding domain. The main disagreement between the studies seems to be on the k_{off} .

Membrane (M: 222 aa): (Neuman et al. 2011) - "Using the M spacing data for each virus (Fig.6C), this would give -1100 M2 molecules per average SARS-CoV, MHV and FCoV particle" Envelope (F: 75 aa): (Godet et al. 1992) - "Based on the estimated molar ratio and assuming that coronavirions bear 100 (Roseto et al., 1982) to 200 spikes, each composed of 3 S molecules (Delmas and Laude, 1990) it can be inferred that approximately 15-30 copies of ORF4 protein are incorporated into TGEV virions (Purdue strain)."

Nucleoprotein (364 aa): (Neuman et al. 2011) - "Estimated ratios of M to N protein in purified coronaviruses range from about 3M:1N (Cavanagh. 1983, Escors et al. 2001b) to 1M:1N (Hogue and Brian. 1986, Liu and Inglis. 1991), giving 730–2200 N molecules per virion."

Genome

<u>Type</u>: (<u>ViralZone</u>) +ssRNA "Monopartite, linear <u>ssRNA(+) genome</u>"

Genome length: (Wu et al. 2020) - Figure 2

Number of genes: (Wu et al. 2020) - "SARS-CoV-2 genome has 10 open reading frames (Fig. 2A)." or (Wu et al. 2020) - "The 2019-nCoV genome was annotated to possess 14 ORFs encoding 27

Number of proteins: (Wu et al. 2020). "By aligning with the amino acid sequence of SARS PP1ab and analyzing the characteristics of restriction cleavage sites recognized by 3CLpro and PLpro, we

speculated 14 specific proteolytic sites of 3CLpro and PLpro in SARS-CoV-2 PP1ab (Fig. 2B). PLpro cleaves three sites at 181–182, 818–819, and 2763–2764 at the N-terminus and 3CLpro cuts at the other 11 sites at the C-terminus, and forming 15 non-structural proteins."

Evolution rate: (Kovama et al. 2020) - "Mutation rates estimated for SARS, MERS, and OC43 show a large range, covering a span of 0.27 to 2.38 substitutions ×10-3 / site / year (10-16)." Recent unpublished evidence also suggest this rate is of the same order of magnitude in SARS-CoV-2. Mutation rate: (Sanjuan et al. 2010) - "Murine hepatitis virus ... Therefore, the corrected estimate of the mutation rate is $\mu_{s/n/c} = 1.9x10^6 / 0.55 = 3.5 \times 10^6$."

Genome similarity: For all species except pangolin: (Wu et al. 2020) - "After phylogenetic analysis and sequence alignment of 23 coronaviruses from various species. We found three coronaviruses from bat (96%, 88% and 88% for Bat-Cov RaTG13, bat-SL-CoVZXC12 and bat-SL-CoVZC45, respectively) have the highest genome sequence identity to SARS-CoV-2 (Fig. 1A). Moreover, as shown in Fig. 1B, Bat-Cov RaTG13 exhibited the closest linkage with SARS-CoV-2. These phylogenetic evidences suggest that SARS-CoV-2 may be evolved from bat CoVs, especially RaTG13. Among all coronaviruses from human, SARS-CoV (80%) exhibited the highest genome sequence identity to SARS-CoV-2. And MERS/isolate NL13845 also has 50% identity with SARS-CoV-2." For pangolin: (Zhang et al. 2020) - Figure 3

Replication Timescales

<u>Virion entry into cell</u>: (Schneider et al. 2012) - "Previous experiments had revealed that virus is internalized within 15 min" and (Ng et al. 2003) - "Within the first 10 min, some virus particles were internalised into vacuoles (arrow) that were just below the plasma membrane surface (Fig. 2, arrows). ... The observation at 15 min postinfection (p.i.), did not differ much from 10 min p.i. (Fig. 4a)"

Eclipse period: (Schneider et al. 2012) - "SARS-CoV replication cycle from adsorption to release of infectious progeny takes about 7 to 8 h (data not shown)." and (Harcourt et al. 2020) - Figure 4 shows virions are released after 12-36 hrs but because this is multi-step growth this represents an upper bound for the replication cycle.

<u>Burst size</u>: (<u>Hirano et al. 1976</u>): "The average per-cell yield of active virus was estimated to be about $6-7 \times 10^2$ plaque-forming units." This data is for MHV, more research is needed to verify these values for SARS-CoV-2.

Host Cells

Type: (Shieh et al. 2005) - "Immunohistochemical and in situ hybridization assays demonstrated evidence of SARS-associated coronavirus (SARS-COV) infection in various respiratory epithelial cells, predominantly type II pneumocytes, and in alveolar macrophages in the lung." and (Walls et al. 2020) - "SARS-COV-2 uses ACE2 to enter target cells" and (Rockx et al. 2020) - "In SARS-COV-2-infected macaques, virus was excreted from nose and throat in absence of clinical signs, and detected in type I and II pneumocytes in foci of diffuse alveolar damage and mucous glands of the nasal cavity....In the upper respiratory tract, there was focal 5 or locally extensive SARS-CoV-2 antigen expression in epithelial cells of mucous glands in the nasal cavity (septum or concha) of all four macaques, without any associated histological lesions (fig. 2I)."

Type I and Type II pneumocyte and alveolar macrophage cell number: (Grapo et al. 1982) - Table 4 and (Stope et al. 1992) - Table 5

Epithelial cells in mucous gland cell number and volume: (ICRP 1975) - surface area of nasal cavity, (Tos & Mogensen 1976) and (Tos & Mogensen 1977) - mucous gland density, (Widdicombe 2019) - mucous gland volume, (Ordoñez et al. 2001) and (Mercer et al. 1994) - mucous cell volume. We divide the mucous gland volume by the mucous cell volume to arrive at the total number of mucous cells in a mucous gland. We multiply the surface density of mucous glands by the surface area of the nasal cavity to arrive at the total number of mucous glands, and then multiply the total number of mucous glands by the number of mucous cells per mucous gland.

Type II pneumocyte volume: (Fehrenbach et al. 1995) - "Morphometry revealed that although inter-individual variation due to some oedematous swelling was present, the cells were in a normal size range as indicated by an estimated mean volume of 763 ± 64 µm^{3**}

Alveolar macrophage volume: (Crapo et al. 1982) - "Alveolar macrophages were found to be the largest cell in the populations studied, having a mean volume of 2,491 μ m3"

Concentration

Nasopharynx, Throat. Stool. and Sputum: (Woelfel et al. 2020) - Figure 2. and (Kim et al. 2020) - Figure 1 and (Pan et al. 2020) - Figure. We took the maximal viral load for each patient in nasopharyngeal swabs, throat swabs, stool or in sputum.

Antibody Response - Seroconversion

Seroconversion time (time period until a specific antibody becomes detectable in the blood): (Zhao et al. 2020) - "The seroconversion sequentially appeared for Ab, IgM and then IgG, with a median time of 11, 12 and 14 days, respectively" and (To et al. 2020) - "For 16 patients with serum samples available 14 days or longer after symptom onset, rates of seropositivity were 94% for anti-NP IgG (n=15), 88% for anti-NP IgM (n=14), 100% for anti-RBD IgG (n=16), and 94% for anti-RBD IgM (n=15)"

Maintenance of antibody response to virus: (Wu et al. 2007) - "Among 176 patients who had had severe acute respiratory syndrome (SARS), SARS-specific antibodies were maintained for an average of 2 years, and significant reduction of immunoglobulin G-positive percentage and titers occurred in the third year."

Virus Environmental Stability

Half life on surfaces; (van Doremalen et al. 2020) - For half-lives we use Supplementary Table 1. For time to decay from $\sim 10^4$ to $\sim 10~{\rm TCID_{50}/L^{-1}}$ air or mL⁻¹ medium, we use the first time titer reached detection limit in Figure 1A for surfaces. For aerosols, we use ten half-life values (1000-fold decrease from 10^4 to 10, meaning 10 halvings of concentration) from Supplementary Table 1. More studies are urgently needed to clarify the implications of virion stability on the probability of infection from aerosols or surfaces.

RNA stability on surfaces: (Moriarty et al. 2020). "SARS-CoV-2 RNA was identified on a variety of surfaces in cabins of both symptomatic and asymptomatic infected passengers up to 17 days after cabins were vacated on the Diamond Princess but before disinfection procedures had been conducted (Takuya Yamagishi, National Institute of Infectious Diseases, personal communication, 2020)."

"Characteristic" Infection Progression in a Single Patient

Basic reproductive number. B_c: (Li et al. 2020) - "Our median estimate of the effective reproductive number, Re—equivalent to the basic reproductive number (R0) at the beginning of the epidemic—is 2.38 (95% CI: 2.04–2.77)" and (Park et al. 2020) - "Our estimated R0 from the pooled distribution has a median of 2.9 (95% CI: 2.1–4.5)."

Latent period (from infection to being able to transmit): (Li et al. 2020). - "In addition, the median estimates for the latent and infectious periods are approximately 3.69 and 3.48 days, respectively." and Table 1 and (He et al. 2020). - We use the time it takes the infectiousness to reach half its peak, which happens two days before symptom onset based on Figure 1b. As symptoms arise after 5 days (see incubation period), this means the latent period is about 3 days.

Incubation period (from infection to symptoms): (Lauer et al. 2020) - "The median incubation period was estimated to be 5.1 days (95% Cl, 4.5 to 5.8 days), and 97.5% of those who develop symptoms will do so within 11.5 days (Cl, 8.2 to 15.6 days) of infection. These estimates imply that, under conservative assumptions, 101 out of every 10 000 cases (99th percentile, 482) will develop symptoms after 14 days of active monitoring or quarantine." and (Li et al. 2020) - "The mean incubation period was 5.2 days (95% confidence interval [Cl], 4.1 to 7.0), with the 95th percentile of the distribution at 12.5 days."

Infectious period (partially overlaps latent period): (Li et al. 2020) - "In addition, the median estimates for the latent and infectious periods are approximately 3.69 and 3.48 days, respectively." and Table 1 and (He et al. 2020) - We quantify the interval between half the maximal infectiousness from the infectiousness profile in Figure 1b.

<u>Disease duration</u>: (WHO 2020) - "Using available preliminary data, the median time from onset to clinical recovery for mild cases is approximately 2 weeks and is 3-6 weeks for patients with severe or critical disease"

Time until diagnosis; (Xu et al. 2020) - We used data on cases with known symptom onset and case confirmation dates and calculated the median time delay between these two dates.

Case Fatality Rate: (ECDC geographic distribution of cases from 29/03/2020) - We use data from all countries with more than 50 death cases and calculate the uncorrected raw Case Fatality Rate for each country. The range represents the lowest and highest rates observed.

Infected Fatality Rate: (Verity et al. 2020) - "We obtain an overall IFR estimate for China of 0.66% (0.39%,1.33%)" and (Ferguson et al. 2020) - "The IFR estimates from Verity et al.12 have been adjusted to account for a non-uniform attack rate giving an overall IFR of 0.9% (95% credible interval 0.4%-1.4%)."

Acknowledgements

We thank Uri Alon, Niv Antonovsky, David Baltimore, Rachel Banks, Arren Bar Even, Naama Barkai, Molly Bassette, Menalu Berihoon, Biana Bernshtein, Pamela Bjorkman, Cecilia Blikstad, Julia Borden, Bill Burkholder, Griffin Chure, Lillian Cohn, Bernadeta Dadonaite, Emmie De wit, Ron Diskin, Ana Duarte, Tal Einav, Avigdor Eldar, Elizabeth Fischer, William Gelbart, Alon Gildoni, Britt Glausinger, Shmuel Gleizer, Dani Gluck, Soichi Hirokawa, Greg Huber, Christina Hueschen, Amit Huppert, Shalev Itzkovitz, Martin Jonikas, Leeat Keren, Gilmor Keshet, Marc Kirschner, Roy Kishony, Amy Kistler, Liad Levi, Sergei Maslov, Adi Millman, Amir Milo, Elad Noor, Gal Ofir, Alan Perelson, Steve Quake, Itai Raveh, Andrew Rennekamp, Tom Roeschinger, Daniel Rokhsar, Alex Rubinsteyn, Gabriel Salmon, Maya Schuldiner, Eran Segal, Ron Sender, Alex Sigal, Maya Shamir, Arik Shams, Mike Springer, Adi Stern, Noam Stern-Ginossar, Lubert Stryer, Dan Tawfik, Boris Veytsman, Aryeh Wides, Tali Wiesel, Anat Yarden, Yossi Yovel, Dudi Zeevi, Mushon Zer Aviv, and Alexander Zlokapa for productive feedback on this manuscript. Figure created using Biorender.

Translated by:

- Clarice de Azevedo Souza, PhD (claricesouza@berkeley.edu), Project Manager, Innovative Genomics
- Alicia Calvo-Villamañán (alicia.calvo@pasteur.fr, PhD student in the Synthetic Biology Lab @ Institut Pasteur and CRI-Paris)