



FACULDADE DE MEDICINA
UNIVERSIDADE DE
COIMBRA

MESTRADO INTEGRADO EM MEDICINA – TRABALHO FINAL

ANA LUÍSA ARAGÃO ALEGRIA OLIVEIRA

A Pandemia COVID-19 e as suas implicações na era digital

ARTIGO DE REVISÃO NARRATIVA

ÁREA CIENTÍFICA DE SAÚDE PÚBLICA

Trabalho realizado sob a orientação de:

DOUTOR ANTÓNIO JORGE FERREIRA

Índice:

Lista de siglas	4
Resumo	5
Palavras-chave	5
Introdução	6
Métodos	9
Discussão	13
Considerações gerais.....	13
Modelos epidemiológicos	17
1) SIR e SEIR	17
2) Modelos de crescimento logísticos e exponenciais	20
3) Outro tipo de modelos	21
Estudo caso.....	22
Conclusão	26
Agradecimentos	28
Referências	29

Lista de siglas

SARS-CoV-2 – Síndrome Respiratório Agudo Corona Vírus 2

WHO – World Health Organization

ML – Machine Learning

DS – Data Science

JHU CSSE - Johns Hopkins University

SIR – Suscetível, Infetado ou Recuperado

SEIR - Suscetível, Exposto, Infetado ou Recuperado

SIRD - Suscetível, Infetado, Recuperado ou Morto

SIQR - Suscetível, Infetado, em Quarentena ou Recuperado

D - SEIQ – Suscetível, Exposto, Infetado ou em Quarentena modelo dinâmico

SEIQR - Suscetível, Exposto, Infetado, em Quarentena ou Recuperado

SAIR – Suscetível, Assintomático, Infetado ou Recuperado

ATLM - Apparent Time-Lag Model

MAPE - Média percentual absoluta do erro

RMM - Redes Neurais Recorrentes

LSTM – Long Short-Term Memory Networks

Resumo

Atualmente vivemos um clima de grande incerteza relativamente à evolução da pandemia COVID-19. Por causa deste acontecimento, inúmeros países por todo o mundo decretaram por diversas vezes, o confinamento geral da população de maneira a impedir o explosivo aumento de casos que tem como consequência a sobrecarga dos sistemas de saúde. Estas medidas acarretam impactos significativos a nível económico e social em cada país. Assim, de maneira a auxiliar a tomada de decisões governamentais, diversos modelos de previsão do SARS-CoV-2 têm surgido. No entanto, e tal como o famoso físico dinamarquês Niels Bohr uma vez disse: “é difícil prever, especialmente o futuro”.

Os modelos mais comuns são os modelos compartimentais SIR e SEIR e variações destes, modelos exponenciais e logísticos e modelos baseados em *deep learning*. Para que estes modelos sejam precisos é essencial conhecer-se as características intrínsecas e extrínsecas relativas a este vírus. Estes dados, aos dias de hoje, ainda não são totalmente conhecidos, o que resulta na imprecisão nas previsões dos modelos epidemiológicos. Para o desenvolvimento destes modelos, diversos parâmetros tiveram de ser estimados. Estes parâmetros não são semelhantes em todos os países, o que faz com que estes modelos de previsão não sejam válidos universalmente.

Torna-se, então, necessário fazer um trabalho de comparação de resultados para a previsão da pandemia - que se pretende fidedigna. Este trabalho de comparação é um processo difícil e complexo. No caso de outras epidemias, os modelos de previsão não foram capazes de prever com exatidão a evolução das mesmas a longo prazo.

No caso da dinâmica do SARS-CoV-2, as previsões a longo prazo ainda não são satisfatórias, uma vez que as variações do ambiente implicam grandes alterações na taxa de transmissibilidade do vírus. Assim, de acordo com os modelos revistos apenas é possível prever com exatidão a evolução a cerca de 20 dias, assumindo que as condições em que a previsão foi gerada se mantenham.

Palavras-chave

COVID-19, epidemiology, forecasting, mathematical models

Introdução

Uma ocorrência pandémica acarreta inúmeros desafios à população Mundial e aos seus sistemas de Saúde. Neste momento encontramos-nos a enfrentar a epidemia da Covid-19. Causada pelo vírus SARS CoV-2, os doentes afetados apresentam diversos sintomas tais como tosse seca, febre e fadiga, os quais podem ser confundidos com outras patologias com taxas de letalidade inferiores. A dispneia é uma das características mais particulares desta patologia e que está intimamente ligada com a sua elevada taxa de mortalidade, quando comparada com as outras patologias que mimica. No nosso país a taxa de mortalidade é próxima dos 2%, mas que na população acima dos 80 anos pode atingir valores de 14,8% (1). A principal causa de morte nesta doença deve-se a complicações de pneumonia tais como falência respiratória.

O vírus da Covid-19 foi detetado pela primeira vez na região de Hubei na China em dezembro de 2019. De Hubei alastrou rapidamente ao resto do mundo devido à sua elevada taxa de transmissão. A transmissão acontece principalmente no contacto pessoa-a-pessoa com pessoas infetadas pelo SARS-CoV-2, através da transmissão de gotículas que contêm partículas virais ou através do contacto com superfícies ou objetos contaminados (2). Uma vez que o vírus é relativamente recente, as medidas farmacológicas que têm sido utilizadas para lidar com a doença demonstram uma eficácia insuficiente. A forma mais segura de prevenção nesta pandemia tem-se materializado sobre a forma de diminuição do contacto entre a população e o isolamento dos infetados (3).

A 11 de março de 2020 a World Health Organization (WHO) anunciou que o vírus SARS-Cov-2 tinha atingido o estatuto de pandemia (4). Em Portugal, o estado de emergência foi declarado pela primeira vez a 18 de março de 2020 de maneira “a prevenir a doença, conter a pandemia, salvar vidas e assegurar que as cadeias de abastecimento fundamentais de bens e serviços essenciais continuavam a ser asseguradas” (5).

A partir desse momento, e um pouco por todo o mundo, foi imposto o dever de confinamento, de forma a evitar o alastramento da doença a uma área geográfica cada vez maior, bem como prevenir a sobreocupação dos serviços de saúde. Desta maneira, toda a economia mundial tem vindo a ser abalada, registando-se quebras acentuadas na atividade económica. Por exemplo, o mercado de ações assinalou quedas abruptas de cerca de 40% no início da pandemia e a taxa de desemprego aumentou em quase todos os países cerca de 2%. Todos estes indicadores levaram a uma diminuição acentuada do produto interno bruto a nível mundial, chegando alguns países a registarem quebras na ordem dos 15% (6). Dada a importância da atividade económica a nível mundial, torna-se imperioso criar condições para que possa ser feita uma extrapolação realista e fidedigna dos dados da

COVID-19, de forma a ser possível amenizar os riscos financeiros e agilizar e priorizar os cuidados de saúde.

Nos últimos meses tem-se assistido a um ressurgimento do aumento exponencial de casos em certos países como Bélgica, Irlanda e Portugal. Nesta última vaga, estes países, atingiram números de casos muito superiores aos que tinham alguma vez registado desde o início da pandemia. Para além disso, é de ressaltar a elevada heterogeneidade a nível mundial na distribuição dos casos. Alguns países situam-se no início da subida de casos, enquanto outros já a ultrapassaram, com mais ou menos danos, e, estão neste momento no rescaldo dessa escalada de casos. Estas diferenças tendem a surgir por diversos motivos, quer pelas medidas de confinamento mais restritas ou mais eficazes, quer pela própria distribuição da população a nível local.

Com toda esta avalanche de dados a surgir é natural que o excesso de notícias, falsas ou não, relativamente à progressão da epidemia, venha a ter um impacto significativo na descredibilização das fontes seguras de informação, ao promover uma falsa sensação de segurança e ao disseminar produtos e serviços enganosos (7).

Assim sendo, ao longo deste último ano muitas têm sido as previsões sobre o comportamento da epidemia, recorrendo a algoritmos do tipo *machine learning* (ML), a *data science* (DS) e a outros modelos mais simplistas (2). Cada um destes tipos de modelos de previsões são ajustados aos países onde foram desenvolvidos, baseados nos dados disponíveis.

Apesar das suas limitações a nível tecnológico e de replicabilidade, estes modelos têm vindo a tentar colmatar as dificuldades implicadas no carácter imprevisível da pandemia de maneira a tentar amenizar as suas consequências a nível financeiro e também nas implicações relativas à gestão dos cuidados de saúde.

Uma vez que este vírus ainda é relativamente recente, torna-se imprescindível recorrer a modelos que têm de ser constantemente testados e melhorados para aumentar a fiabilidade das suas previsões.

A fiabilidade dos modelos de previsão tem ficado muito aquém das expectativas, devido a diversos fatores. O vírus SARS-CoV-2 apresenta características que nos são desconhecidas e por isso tem sido feito um esforço para tentar discernir diversos dos seus atributos. No entanto, atualmente, ainda não se conhece o real grau de infeção, ou a influência da meteorologia ou das diferentes características populacionais no comportamento do vírus. Adicionalmente, diversas novas variantes da COVID-19 têm surgido, gerando assim, uma maior complexidade à modelização do comportamento da pandemia. Estas variantes circulam já um pouco por todo o mundo e aparentam transmitir-se mais facilmente e mais rapidamente (8).

Mais recentemente, várias empresas da indústria farmacêutica desenvolveram vacinas que se encontram já a ser administradas em vários países. Esta parece ser neste momento a solução com maior taxa de sucesso no combate à progressão da pandemia.

No entanto, para ser possível atingir a imunidade de grupo, vacinar a maior parte da população irá demorar alguns meses - pelo menos em Portugal (9). Para além disso, é importante ressaltar que a vacina ainda não foi exaustivamente testada para as novas variantes do vírus, o que aumenta a probabilidade de uma menor eficácia na vacinação. Mesmo tendo em conta todas estas variáveis, em países com taxas de vacinação elevadas na sua população, como é o caso de Israel, vários dados sugerem que as suas taxas de transmissão tenham diminuído (10). Desta maneira, parece surgir mais um parâmetro a ter em conta na criação de modelos de previsão para esta pandemia que está longe de terminar.

O objetivo principal deste trabalho é analisar e comparar modelos de previsão da pandemia da COVID-19 que utilizem bases de dados fidedignas, bem como procurar tirar conclusões relativamente aos intervalos temporais que se podem considerar confiáveis ao interpretar os dados extrapolados dos modelos da evolução da pandemia existentes.

Métodos

Para a realização deste trabalho foi feita uma pesquisa na base de dados *PubMed* em dezembro de 2020 (término a 14 de dezembro 2020). A pesquisa foi iniciada cruzando as palavras-chave: “*covid-19*” e “*epidemiology*” associadas a “*forecasting*”, “*prediction*” e “*transmission*”. Desta, surgiram 98 artigos (cujas ocorrências das palavras-chave estão na Figura 1). Destes foram selecionados os publicados entre agosto e dezembro de 2020 e para além disso, aqueles que, de acordo com o seu resumo e título se enquadravam no tema deste trabalho. Desta forma a amostra inicial foi reduzida para 31 artigos.

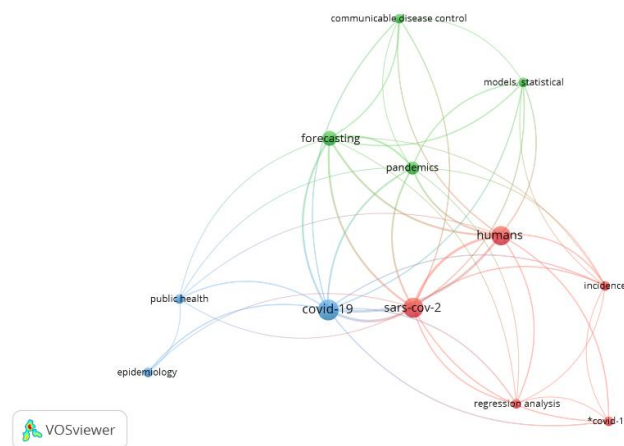


Figura 1- Mapa ocorrência de palavras-chave

Não havendo muita informação sobre o tema do trabalho, não foram colocadas quaisquer limitações quanto à tipologia dos artigos a considerar, incluindo-se, assim, revisões sistemáticas, meta-análises, entre outros.

Desta primeira seleção foram então escolhidos os artigos em que 1) os dados analisados envolvessem mais do que um país/região e 2) a análise de dados envolvesse de preferência mais do que um modelo matemático. Foram incluídos, também, os artigos que apesar de se centrarem em países específicos se referiam a situações especiais, como é o caso da situação de confinamento.

Por último, foram ainda selecionados 2 artigos que relacionam dados dos modelos de previsão com características da população.

Com base nestas especificações selecionaram-se, então, 17 artigos alinhados com a estratégia descrita. Dos 17 artigos selecionados, 4 são artigos de revisão de modelos de previsão, sendo que um deles é uma revisão sistemática e os restantes 12 artigos são artigos originais (tabela 1).

Não obstante ter-se selecionado a modalidade de artigo de revisão narrativa, a autora não quis deixar de dar o seu contributo pessoal, efetuando um estudo-caso adaptado à atual realidade portuguesa, aproveitando para este efeito o seu *background* e formação científica prévia.

No estudo-caso, criou-se em *MatLab* um modelo de previsão baseado no modelo SEIQR (definido adiante), de modo a verificar a confiança das previsões para diferentes horizontes temporais, relativamente à situação pandémica em Portugal. O modelo de previsão foi adaptado de (11). As previsões apresentadas, utilizam apenas os dados dos valores diários confirmados da COVID-19 fornecidos pela *Johns Hopkins University* (JHU CSSE).

Tabela 1 – Descrição dos artigos selecionados

Autores	Título	Tipo artigo	Modelo de previsão
I. Rahimi, F. Chen, A. Gandomi	A Review on COVID-19 Forecasting Models	Revisão de modelos de previsão	Modelos compartimentais (SIR, SEIR, SIRD), Modelos exponenciais, regressão linear e deep learning
P. Nadella, A. Swaminathan, S. Subramanian	Forecasting efforts from prior epidemics and COVID-19 predictions	Revisão de modelos de previsão em pandemias antigas	-
A. Gola, R. Arya, A. Animesh et al.	Review of Forecasting Models for Coronavirus (COVID-19) Pandemic in India during Country-wise Lockdown	Revisão de modelos na previsão	Modelos exponenciais, deep learning e modelos compartimentais (SEIR, SIR)
Y. Mohamadou, A. Halidou, T. Pascalin	A review of mathematical modeling, artificial intelligence and datasets used in the study, prediction and management of COVID-19	Revisão de modelos na previsão	Modelos compartimentais (SEIR, SIR, SIQR)

J. Guan, Y Wei, Y. Zhao et al.	Modeling the transmission dynamics of COVID-19 epidemic: a systematic review	Revisão sistemática de modelos de previsão dinâmicos	Modelos compartimentais (SEIR, SIR)
K. Law, K. Peariasamy, B. Gill et al.	Predicting the early depleting transmission dynamics of COVID-19: A time-varying SIR model	Artigo original	SIR
L. Taghizadeh, A. Karimi, C. Heitzinger	Uncertainty quantification in epidemiological models for COVID-19 pandemic	Artigo original	SIR e logístico
M. Utamura, M. Koizumi, S. Kirikami	An epidemiological model considering isolation to predict COVID-19 trends in Tokyo, Japan: Numerical analysis	Artigo original	ATLM
E. Curado, M. Curado	A discrete-time-evolution model to forecast progress of Covid-19 outbreak	Artigo original	SAIR
L. Kriston	Predictive accuracy of a hierarchical logistic model of cumulative SARS-CoV-2 case growth	Artigo original	Logístico
K. Pacheco-Barríos, A. Cardenas-Rojas, S. Giannoni-Luza, F. Fregni	COVID-19 pandemic and Farr's law: A global comparison and prediction of outbreak acceleration and deceleration rates	Artigo original	Logístico
Z. Zheng, K. Wu,	The prediction for development of	Artigo original	Matriz de estado

Z. Yao et al.	COVID-19 in global major epidemic areas through empirical trends in China by utilizing state transition matrix model		
J. Sun, X. Chen, Z. Zhang et al.	Forecasting the long-term trend of COVID-19 epidemic using a dynamic model	Artigo original	D-SEIQ
A. Attanayake, S. Perera, S. Jayasinghe	Phenomenological Modelling of COVID-19 epidemics in Sri Lanka, Italy and Hebei Province of China	Artigo original	Modelos exponencial e logístico
M. Castro, S. Ares, J Cuest et al.	The turning point and end of an expanding epidemic cannot be precisely forecast	Artigo original	SCIR
T. Zhou, J. Yuan	Semiparametric Bayesian inference for the transmission dynamics of COVID-19 with a state-space model	Artigo original	SIR
B. Rader, S. Scarpino, A. Nande et al.	Crowding and the shape of COVID-19 epidemics	Artigo original	

Discussão

Considerações gerais

A pandemia da COVID-19 tem colocado os sistemas de saúde a nível mundial em grande pressão, devido à elevada taxa de transmissão do vírus e à ineficácia das medidas farmacêuticas. Atualmente, as medidas mais consensualmente aceites para evitar a propagação do vírus são as de natureza não farmacêuticas (12).

A propagação epidémica é caracterizada por um crescimento exponencial, que é intrinsecamente imprevisível, sendo difícil ou impossível definir o momento em que a curva epidemiológica começa a diminuir ou quando irá ocorrer o momento final da epidemia (13).

Desde o início da pandemia foram desenvolvidos inúmeros modelos preditivos, os quais têm tido um impacto significativo na gestão de toda a pandemia. Estes modelos auxiliam os governos a definir melhor estratégias e a tomar decisões sensíveis à realidade local de cada país. A partir da análise de dados e padrões passados, os modelos de previsão traçam cenários em que é prevista a evolução futura do comportamento epidemiológico do vírus. Por isso, é de incontestável importância que estes sejam confiáveis e robustos, pois desempenham um importante papel na preparação para situações futuras (2).

No início, a disseminação do SARS-Cov-2 deveu-se principalmente à mobilidade humana (12) que se registou entre a região de Hubei e o resto da China e que a partir daí se propagou para o resto do mundo. Foi esta característica do vírus que veio obrigar a associação de tomada de medidas de distanciamento e de diminuição da mobilidade aos modelos matemáticos e estatísticos para melhor afinar as estratégias de antecipação da disseminação da COVID-19 (14).

O maior desafio é desenvolver um modelo que seja significativamente preciso e que, simultaneamente, permita fazer previsões com confiança a longo prazo. É de salientar que, para outras epidemias, modelos similares foram criados de maneira a ajudar a clarificar a sua progressão, como é o caso do vírus H5N1, da doença de Creutzfeldt-Jakob, do ébola e do SARS. As análises das previsões dos modelos para estas patologias mostram que as suas previsões não coincidiram com os valores reais, acabando por estar, na sua vasta maioria, significativamente sobre ou subestimadas. Por exemplo, no caso da doença de Creutzfeldt-Jakob os modelos previam cerca de 50 000 mortes a nível mundial e atualmente apenas há registo de 177 (7).

Os modelos de modelação ou de previsão criados para esta pandemia baseiam-se, sobretudo, nos dados publicados a nível mundial nomeadamente na plataforma da JHU CSSE (15–20).

Os dados avaliados pelos modelos dividem-se em dois grupos principais: 1) características intrínsecas do vírus (tempo de incubação, virulência, etc.) e 2) características extrínsecas ao vírus (população suscetível, cumprimento de medidas de distanciamento, etc.) (2).

De notar que os detalhes específicos dos mecanismos inerentes ao vírus e da epidemia são ainda desconhecidos - por exemplo, como se atinge a imunidade, a sua duração ou mesmo se existe; se é possível a ocorrência de reinfeções (18); a taxa de transmissão do vírus em assintomáticos ou a percentagem de infetados que não cumpre a quarentena (21).

A partir dos dados e através de uma rigorosa abordagem que envolve a análise da sensibilidade dos modelos, validação e adequação de dados (3) surgem os modelos que são apresentados atualmente, dos quais se destacam como dominantes, os modelos determinísticos Suscetível, Infetado e Recuperado (SIR) e Suscetível, Exposto, Infetado e Recuperado (SEIR) (21). Pode-se argumentar que estes modelos são demasiado simplistas e que por isso não são exemplificativos da complexidade da evolução pandémica (21). No entanto, estes modelos menos complexos apresentam bons graus de coincidência com a realidade para extrapolações de cerca de 2 semanas (2). É importante referir ainda, que o conjunto de dados relativos à COVID-19 é bastante extenso e que estes são atualizados diariamente, nomeadamente o número de casos confirmados, de mortes e de pessoas recuperadas.

A vantagem dos modelos SIR e SEIR está exatamente na sua simplicidade e no valor mais reduzido de parâmetros. Nos modelos com elevado número de parâmetros, não só a sua interpretação se torna mais complexa, como a sensibilidade destes à história do evento pandémico é mais reduzida. Os dados ficam sobre-ajustados, ou seja, os modelos tornam-se demasiado treinados para conjuntos de dados anteriormente observados, tornando-se ineficazes para a previsão de novos resultados. Daqui resulta com que haja uma certa preferência por modelos simples e com poucas variáveis, que podem ser implementados e atualizados com elevada simplicidade, facilitando, assim, o acompanhamento da propagação da pandemia e, mais importante ainda, a avaliação da eficácia das decisões políticas para controlo do aumento dos casos (18).

Para além dos modelos determinísticos atrás referidos, várias têm sido as tentativas de construir modelos dinâmicos – e.g. modelos logísticos, matriz de estado, SIR dependente do tempo - que apresentem parâmetros voláteis e que melhor se ajustem à variabilidade dos dados, estimando picos pandémicos e permitindo, assim, uma melhor avaliação da eficácia das medidas governamentais aplicadas (21). Outros modelos consideram também

parâmetros sociodemográficos como a densidade populacional, o índice urbano, os dados de rendimento ou a percentagem de população idosa, entre outros (16).

Na maioria dos modelos assume-se que os indivíduos se classificam quase exclusivamente a partir do seu estatuto epidemiológico, ou seja, pela sua capacidade de transmissão do vírus (3).

Diferentes modelos utilizam variáveis distintas; este fator deve ser, portanto, tomado em consideração quando se analisam as previsões produzidas por cada um deles. Por exemplo, um modelo pode não ter em consideração que todos os indivíduos infetados estejam confinados, mas pode considerar que a taxa de recuperação é um dado dinâmico uma vez que a qualquer momento pode surgir um tratamento farmacológico eficaz para o vírus. Outros modelos têm em conta as diferentes taxas de transmissão relativas a cada faixa etária, diferenças de cada região em termos de densidade populacional e ainda a facilidade de acesso a recursos de saúde de qualidade (13).

Apesar das suas limitações, os modelos parecem devolver previsões relativamente fiáveis para futuros próximos da evolução da pandemia, desde que as condições de propagação (i.e., medidas de confinamento ou os hábitos de uma dada população) permaneçam inalteradas. De um modo geral, as previsões parecem ser eficazes para intervalos temporais que vão desde as duas semanas até cerca de quarenta dias (2,20).

Pelo que foi dito, podemos concluir que a comparação de diferentes modelos de previsão é uma tarefa complexa, não só pela sua enorme variedade e adjacente fundamentação teórica, mas também devido às suas diferentes metodologias, nem sempre transparentes ao nível da recolha de dados, dos pressupostos dos modelos ou dos diferentes fatores epidemiológicos analisados.

O momento em que os dados são recolhidos, nomeadamente o ponto em que se situam na curva epidemiológica também confere um fator de heterogeneidade entre os diferentes modelos epidemiológicos. Até mesmo quando se opta por modelos que assentam em modelos matemáticos similares, os resultados gerados podem ser diferentes. Por esta razão, torna-se imprescindível a existência de guias de parametrização para os modelos preditivos (22). Por exemplo, os modelos que foram criados em países como a China num dado momento da pandemia, apenas são replicáveis em países que adquiriram conhecimento e recursos similares aos da China, tal como a Coreia do Sul e a Itália. Em países com capacidade de diagnóstico limitada, como por exemplo o Irão, o padrão apresentado é distinto dos países supramencionados (23).

Dados imperfeitos, pressupostos não verificáveis e/ou a imprevisibilidade do comportamento humano tornam a tarefa de idealização ou de implementação dos modelos de previsão ainda mais difícil (7).

Todas estas variáveis são, em geral, imprevisíveis e mutáveis, uma vez que dependem em grande medida das políticas de saúde pública, da sensibilização da população (18) e da evolução do nível de desenvolvimento económico de um dado país (22). Diferentes nações impuseram medidas não-farmacológicas distintas, de maneira a abrandar a progressão da epidemia, tais como quarentena obrigatória, utilização de máscara em locais públicos, fecho de mercados ou restrições à circulação nacional e internacional (2). Em países com uma significativa limitação dos recursos de saúde disponíveis, as variáveis a analisar serão necessariamente diferentes das que deverão ser escrutinadas em países em que não se verifiquem estas dificuldades. É o caso de países em que há uma testagem mais extensiva da população, onde é expectável que o número de doentes, nomeadamente assintomáticos, seja maior e, por consequência, o número de infetados seja mais pronunciado (18).

Outro fator a considerar em países com recursos limitados de prestação de cuidados de saúde (e por conseguinte menor capacidade de testagem da população), será a sua maior taxa de mortalidade, como se verifica em alguns países africanos, o que poderá ter como consequência um impacto socioeconómico e político devastador (16).

Outros fatores devem ser considerados como preponderantes na modelização da dinâmica da pandemia, como por exemplo a diferenciação em termos de faixa etária, a densidade populacional ou mesmo a influência das condições ambientais e climáticas.

Assim, a taxa de transmissão da COVID-19 aparenta ser superior em pessoas mais idosas em comparação com pessoas mais jovens (22). A densidade populacional está também intimamente relacionada com a ocorrência e intensidade dos picos da pandemia. Assim sendo, o número de casos num determinado período de tempo é proporcional ao grau de aglomeração de pessoas em locais fechados. De acordo com um estudo, (12) conclui-se que as cidades com maior densidade populacional apresentam picos da epidemia que duram mais tempo, bem como um número de casos superior relativamente a cidades com menor densidade populacional. No entanto, as zonas rurais, de menor densidade populacional, apresentam pico mais frequentes, embora com menor número de casos (12). Os estudos realizados até ao momento sobre a influência das condições ambientais e climáticas na COVID-19 parecem ser ainda insuficientes, uma vez que os estudos publicados apresentam resultados díspares entre si (19).

Apesar de todas estas incertezas e limitações, os modelos matemáticos e de simulação dinâmica de doenças infecciosas são essenciais ao entendimento e à previsão do desenvolvimento de qualquer epidemia (17).

Para que os modelos de previsão possam servir de auxílio à adoção de estratégias a nível local e/ou global, devem ser tidos em consideração os intervalos de confiança considerados admissíveis para cada um deles, bem como ter-se em conta a incerteza inerente a estas previsões (7).

A eficácia dos modelos é geralmente avaliada com base em métodos de medição do erro. O método mais utilizado é a média percentual absoluta do erro (MAPE) (2,3,17). O MAPE é um método de medição do erro em escala que compara os dados e é calculado através da normalização do erro em cada ponto, fornecendo os resultados em percentagem (2).

Em alguns países, na ajuda à tomada de decisões por parte das entidades governamentais são utilizados vários modelos para tentar diminuir e corrigir possíveis erros intrínsecos aos modelos. Por exemplo, nos Estados Unidos da América, o Centro para o Controlo de Doenças dos EUA (CDC) utiliza um conjunto de 32 modelos para conseguir uma maior precisão na previsão (2).

É importante referir que, à medida que se encontram disponíveis mais dados sobre o SARS-CoV-2, os parâmetros dos modelos de previsão atingem uma maior maturidade, no sentido em que tomam valores mais verosímeis e a sua capacidade de previsão é melhorada. O SARS-CoV-2 é bastante diferente de outras doenças infecciosas pois os seus sintomas únicos ainda não estão bem definidos e, por isso, infetados com sintomas sub-clínicos podem não ser totalmente identificados (24).

De forma a conseguir melhores previsões, é relevante que haja uma maior articulação entre as várias nações mundiais para se tentar chegar a um modelo padronizado e universal. Neste momento, tomam-se alguns países como casos-modelo (China, Paquistão, França, Itália, Estados Unidos da América, Brasil, Alemanha e Índia) na criação e desenvolvimento dos modelos de previsão (25).

Modelos epidemiológicos

1) SIR e SEIR

Os modelos epidemiológicos mais utilizados são os modelos SIR e SEIR (figura 2). Estes modelos são classificados como modelos compartimentais pois dividem toda a população em compartimentos. O modelo SIR considera três variáveis que representam o número de pessoas em cada um dos compartimentos num dado momento: os Suscetíveis (S), que representa os indivíduos que podem vir a ser infetados; os Infetados (I), que corresponde aos indivíduos infetados, e que, portanto, podem transmitir a doença; e os Recuperados (R) que contém os indivíduos que contraíram a doença. Dependendo da forma de implementação do modelo, os indivíduos em R podem ser ou não imunes e, por isso, ser contabilizados posteriormente em S.

O modelo SEIR tem em conta ainda os indivíduos Expostos (E), que se refere aos indivíduos expostos à doença, mas que ainda não estão infetados (2).

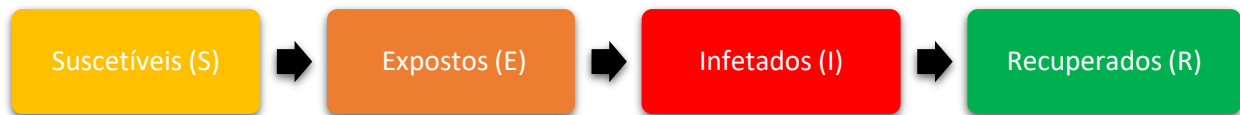


Figura 2 - Representação do modelo SEIR

Em ambos os modelos, a transição entre diferentes compartimentos é feita através de uma taxa que varia com a dinâmica da infeção e com as características da população. Estes modelos de transmissão epidémica tradicionais necessitam de estimativas precisas para essas taxas, nomeadamente o número básico de reprodução, o período de incubação e o período infeccioso (20).

Os modelos SIR e SEIR são muito utilizados para realizar as previsões epidemiológicas porque, apesar de terem um pequeno número de variáveis, são confiáveis a curto prazo e a sua implementação apresenta um baixo custo. Mais, por usarem poucas variáveis, a sua implementação é sensível a qualquer alteração das condições da epidemia e, portanto, podem aferir rapidamente qualquer alteração de políticas governamentais para tentar conter a propagação da COVID-19 (25). Estes modelos permitem, por exemplo, concluir se o confinamento da população é ou não eficaz num determinado período de tempo (19).

A capacidade de previsão destes modelos depende imenso da qualidade dos dados em que se baseiam. O atraso nas atualizações de dados, bem como eventuais demoras na realização de testes de diagnóstico, inconsistências nas taxas de rastreio e alterações substanciais às condições iniciais do modelo (3) podem comprometer significativamente a precisão da sua previsão. Assim, não é de admirar que os modelos baseados apenas nos dados disponíveis no início da pandemia, quando os períodos de transmissão ainda não estavam bem definidos, apresentem os maiores desvios nas previsões (21).

Em geral, as previsões geradas por estes modelos são confiáveis para períodos de 15 a 20 dias. O modelo SEIR apresenta valores mais baixos de MAPE do que o modelo SIR. No entanto, ambos têm valores superiores quando comparados com outros modelos multi-paramétricos, nomeadamente o modelo ARIMA, o modelo LSTM e outros modelos que têm por base simplesmente processos de *deep learning* (2).

No contexto da pandemia COVID-19, vários foram os estudos que utilizaram os modelos SIR, SEIR e SIR, SEIR modificados para prever o futuro da propagação do vírus (2,3,18–21,24,25). A principal alteração verificada nos modelos modificados é a adição de um ou mais parâmetros que variam com o tempo - modelos dinâmicos (3,20,24,25). Estes parâmetros dinâmicos permitem uma melhor descrição da evolução da pandemia, adaptando-se o modelo mais facilmente aos dados que vão sendo atualizados. A análise da variação dos parâmetros dinâmicos fornece informação adicional sobre a evolução da pandemia, como por exemplo a medida da eficácia da adesão do público à diminuição de contactos físicos interpessoais ao longo do tempo e/ou da eficácia de prevenção de propagação da doença por indivíduos infetados (3).

Os modelos SEIR, que não têm parâmetros dinâmicos, substituem-nos pela realização de médias de dados sobre intervalos de tempo ou recorrendo a valores de parâmetros estimados através da evolução de outros vírus da mesma família (20).

Outros modelos optam por incluir compartimentos adicionais e/ou por substituir alguns dos compartimentos dos modelos originais. Grande parte dos modelos não contabiliza os doentes recuperados, mas sim os doentes infetados que estão em quarentena (Q) e que, portanto, não podem infetar, em princípio, outras pessoas – esta tipologia é representada pelos modelos SEIQ (20) e SIQR (19).

Um outro modelo analisado inclui um compartimento para descrever os indivíduos infetados assintomáticos (A) – modelo SAIR (18).

Alguns modelos contabilizam os doentes que faleceram (D). Este tipo de modelo é mais utilizado quando a população inicial é menor já que a contabilização dos doentes falecidos tem um impacto maior na população – representado pelo modelo SIRD (19).

Para uma determinação mais precisa dos parâmetros desconhecidos sobre a pandemia, três estudos (13,15,26) utilizam-se técnicas de inferência bayesiana sobre modelos do tipo SIR. Devido à realização destas inferências sobre parâmetros desconhecidos foi possível obter. Dos modelos analisados, os que apresentaram períodos mais longos de previsão com menor erro de confiança são o 1) modelo SAIR (18), que apresenta uma previsão confiável a 14 dias, e o 2) modelo D - SEIQ (20), que é capaz de prever com confiança até 30 dias. Estas previsões são válidas desde que algumas condições sejam verificadas, como a existência de recursos médicos adequados, coerência nas medidas de contenção e deteção e comunicação atempada dos casos. Uma limitação importante do modelo D-SEIQ é que este é apenas capaz de prever uma única vaga epidémica (20).

Deve ainda ser referido que para a definição dos períodos temporais atrás referidos os modelos foram testados em vários países com diferentes características, quer em termos populacionais, quer em termos de situação pandémica.

A introdução de parâmetros dinâmicos no modelo SIR levou à criação do modelo ATLM (24), que tem em conta o atraso na identificação e isolamento dos casos positivos. Este modelo foi aplicado apenas à cidade Tóquio e os resultados mostram que ele é confiável para um período de até 77 dias.

De uma maneira geral, nos estudos analisados, o desvio entre o pico previsto e o pico real varia entre 5 e 23 dias (21) .

2) Modelos de crescimento logístico e exponencial

Os modelos de crescimento logístico e exponencial são utilizados essencialmente para caracterizar a fase inicial de uma epidemia (14). No modelo de crescimento exponencial considera-se uma taxa de crescimento constante e, portanto, quanto maior for o número de infetados, maior será o número de novos casos. Desta maneira, o crescimento natural pode nunca terminar. Por outro lado, os modelos de crescimento logístico têm um comportamento do tipo exponencial na primeira fase, mas consideram que a taxa de crescimento é reduzida até ela atingir um valor mínimo (14). Se o crescimento exponencial é caracterizado por uma curva em forma de J, o crescimento logístico tem uma curva em forma de S (25), tal como é apresentado na figura 3.

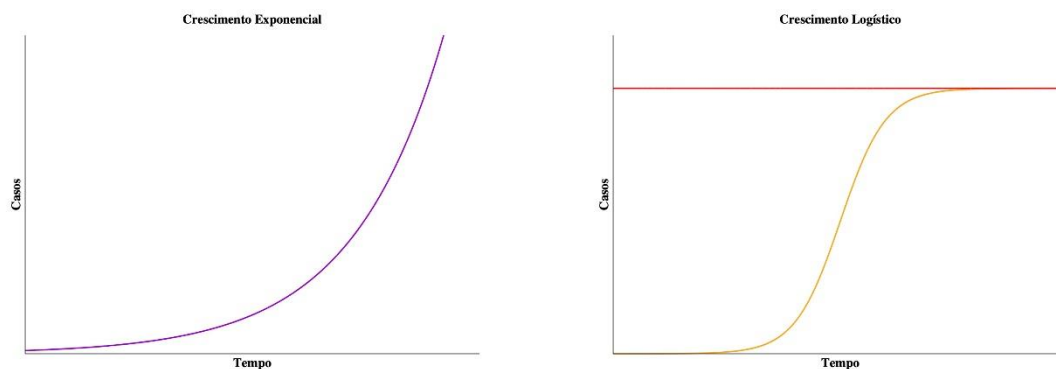


Figura 3 -Modelos de crescimento exponencial e logístico

Nestes modelos é importante que os dados sejam analisados isoladamente, ou seja, exclusivamente em tendências crescentes ou decrescentes. Por esta razão, os estudos que utilizam este género de modelos têm de dividir os dados em, pelo menos, dois conjuntos e analisá-los separadamente: uma primeira análise para os dados anteriores ao pico e uma segunda análise para os dados posteriores ao pico (2,14,17).

Estes modelos, tendo solução analítica conhecida, permitem realizar um ajuste das curvas aos dados existentes (17).

Os modelos logísticos e exponenciais foram testados em diferentes países, com diferentes índices de desenvolvimento e capacidade de resposta à pandemia, como o Sri Lanka, a Itália, a Bélgica, o Bangladesh, o Kuwait e a Índia (14,17).

Estes modelos apresentam previsões bastante fidedignas num horizonte temporal de cerca de duas semanas (17) e valores de MAPE equiparáveis aos dos modelos SIR para o mesmo espaço de tempo (25). Porém, para previsões com duração superior, a curva afasta-se significativamente de valores reais, atingindo diferenças de valor na ordem dos 19,5% para previsões feitas a um mês (25).

Uma forma de estes modelos serem capazes de fazer previsões para horizontes temporais superiores é a introdução de parâmetros dinâmicos dependentes do tempo e não fixos, como acontece na sua formulação. Esta alteração elimina algumas das limitações destes modelos, nomeadamente a capacidade de apenas preverem uma única vaga epidémica e apenas se não se verificarem alterações significativas das condições iniciais (como, por exemplo, o ritmo constante da testagem). Este tipo de modelos apresenta outra limitação relevante, pois em geral, são utilizadas as contagens cumulativas de casos e não as contagens de casos novos, o que pode levar a erros graves de modelação (17).

3) Outro tipo de modelos

Um outro modelo também divulgado tem por base as Leis de Farr (16). Estas leis representam a evolução no número de casos ao longo do tempo através de uma fórmula matemática que graficamente se assemelha a uma curva em forma de sino. Neste modelo, a dinâmica da epidemia é descrita através da relação de dois rácios aritméticos – um relativo às alterações dos números de casos ou de mortes ao longo do tempo (que representa a velocidade de propagação do vírus) e outro que expressa a aceleração da epidemia.

A análise de dados de 69 países em que se aplicou este modelo mostra que ele consegue prever com exatidão o comportamento qualitativo, mas não quantitativo da epidemia. O modelo identifica os países com maior mortalidade, com maior número de casos, bem como os que apresentam maior dinâmica em termos de avanço da pandemia.

A análise deste modelo mostra que ele pode ser aplicado a outras pandemias, pois descreve bem a evolução da epidemia e permite modelar rápida e eficazmente a sua evolução, mesmo com dados limitados e pressupostos simples. O estudo realizado em (8) mostra que para melhorar o desempenho deste modelo é importante incluir uma variável associada à migração de pessoas entre países.

Outro tipo de modelos, também muito utilizado para prever a evolução da COVID-19, baseia-se em técnicas de *deep learning* (25):

1) Um dos modelos desta família de modelos é o modelo das redes neuronais recorrentes (RNN). Este resulta de um algoritmo cuja característica diferenciadora é a sua capacidade de memorizar dados; ou seja, apresenta um comportamento dinâmico temporal. Este modelo apresenta melhores resultados para dados relativos a séries sequenciais.

Baseando-se nos dados anteriormente inseridos (memória), a rede gera novas informações sequenciais (2). Para a determinação destes novos dados, o modelo seleciona um comprimento de dados variável que, em teoria, pode ter tamanho quase infinito. Quanto maior for o número de dados, maior será o risco de haver um *vanishing gradient*. Quando este gradiente apresenta valores muito elevados, a aprendizagem feita pelo modelo deixa de ser realista. Para evitar este mecanismo foram criadas as redes *Long Short-Term Memory* (LSTM).

Os modelos RNN têm vindo a ser utilizados para explorar as relações entre a taxa de transmissão da COVID-19 e diferentes valores de temperatura e de humidade (25).

2) O modelo LSTM é um modelo similar ao modelo RNN, que permite não só processar dados pontuais como também sequências definidas de dados. Cada unidade LSTM é composta por uma célula que controla o período de tempo a ser analisado e três portas: uma de entrada, uma de saída e uma de esquecimento, que permitem regular o fluxo de informação. O *vanishing gradient* é evitado, considerando apenas sequências delimitadas de dados. Desta forma, o modelo tem uma insensibilidade relativa aos dados presentes e, com base na história passada, consegue antecipar possíveis aumentos de casos independentemente do estado atual da pandemia. Este modelo apresenta dados com MAPE de 63,35% para resultados a 25 dias, o que é um desempenho bom, mas pior quando comparado com alguns modelos SEIR (2).

A principal desvantagem dos modelos de *deep learning* é necessitarem de muitos dados e que estes tenham sido gerados em situações diferentes, para que possam prever com exatidão novos comportamentos. Será, provavelmente, por esta razão que os modelos de previsão baseados em *deep learning* não são, atualmente, os mais precisos. Além disso, o treino destes dados tem elevados custos monetários (25).

Estudo-caso

Um modelo dinâmico SEIQR (11) foi utilizado para melhor explicitar as diferenças de previsão em diferentes horizontes temporais. Para este modelo utilizaram-se estimativas para a taxa

de infecção, taxa relativa ao tempo que os infetados demoram para entrar em isolamento profilático, taxa de recuperação e taxa de morte.

Inicialmente, a previsão foi feita para corresponder a um período de quatro meses (figura 4). A partir da análise deste gráfico pode-se concluir que a previsão não é confiável, uma vez que os dados estimados não correspondem aos dados reportados. Aparentemente, os valores mais confiáveis são os relativos ao número de óbitos; contudo, a disparidade maior surge no que diz respeito aos valores dos infetados ativos.

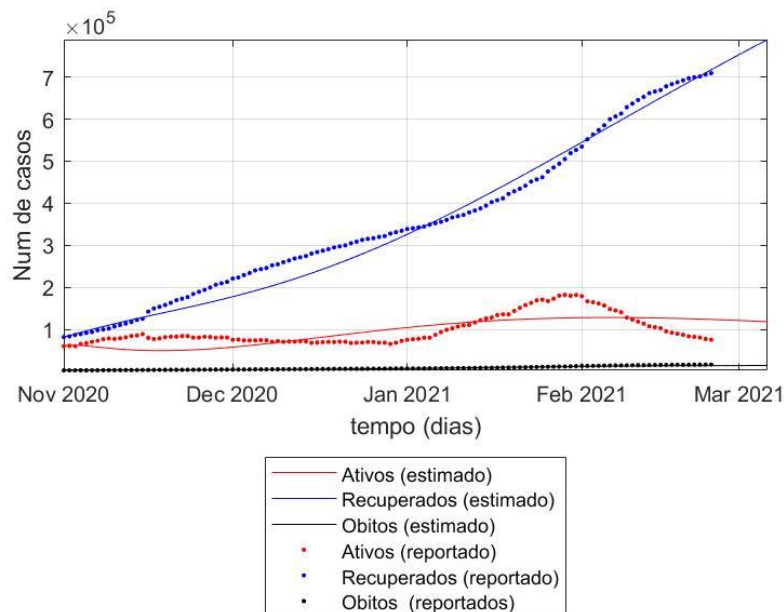


Figura 4 – Ajuste a 4 meses

Foram ainda feitas simulações que se aproximavam mais dos valores referidos como confiáveis por outros estudos feitos para este tipo de modelos. Assim, foi realizado outro teste para 45 dias (figura 5). Ao contrário do que se refere no cenário a 4 meses, a previsão apresenta já valores estimados muito similares aos reais: contudo, regista novamente maior semelhança ao nível do número de óbitos e de recuperados e maior disparidade relativamente ao número de ativos. A maior dificuldade do ajuste da curva surge no pico, que, como se esperaria, apresenta uma maior variação quando comparada com as condições iniciais da previsão.

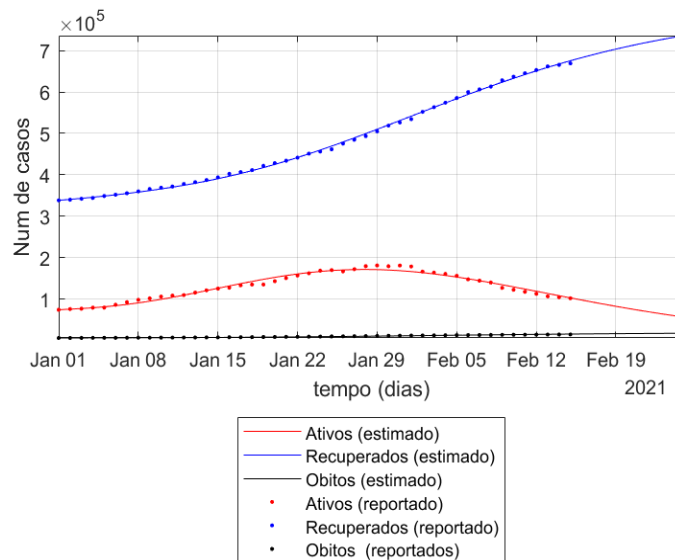


Figura 5 - Ajuste a 45 dias

Finalmente, foi realizado um teste para o horizonte temporal de 20 dias que corresponde ao período que a maioria dos estudos refere como confiável para a previsão (figura 6). Neste caso, a curva estimada ajusta-se quase sem variações aos valores reportados.

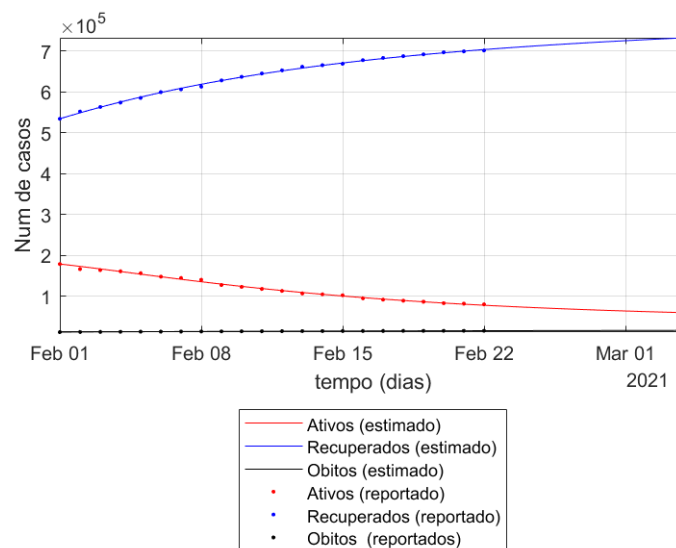


Figura 6 - Ajuste a 20 dias

Por último, considerou-se pertinente fazer uma previsão para a evolução do número de casos a partir do dia 23 março e para um período de 20 dias (figura 7). De acordo com a previsão, a tendência é que o número de casos ativos continue a diminuir até dia 7 de abril e que após essa data ocorra um crescimento no número de casos ativos; a taxa relativa ao número de óbitos deverá manter-se constante e o número de recuperados deve aumentar à mesma taxa

de variação. Assim, este modelo prevê que no dia 13 abril tenhamos um total de cerca de 27885 casos ativos. Estes dados só poderão ser confrontados com os dados reais posteriormente à data de entrega deste trabalho.

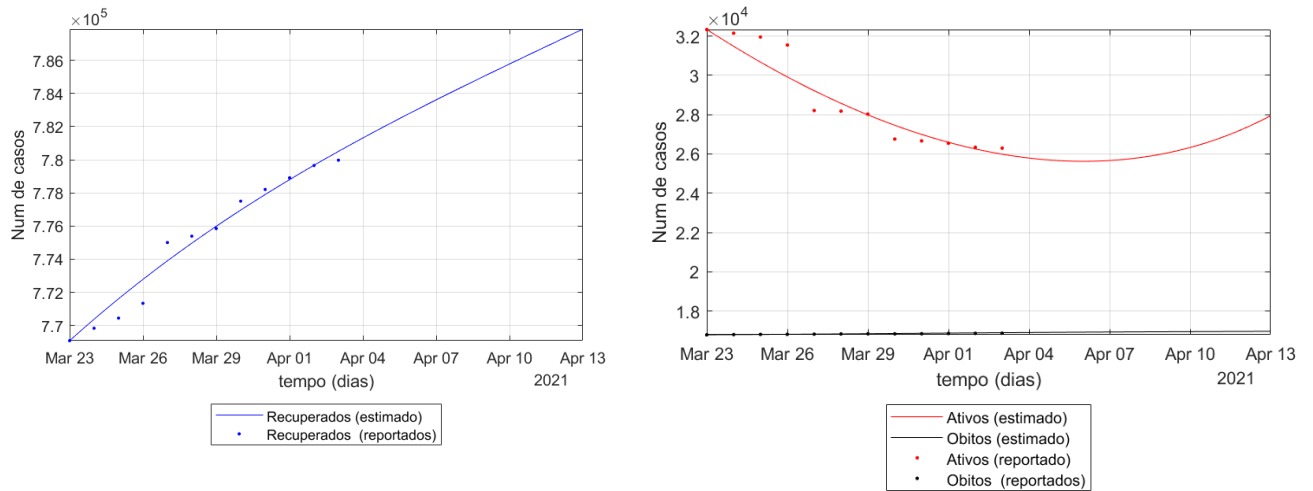


Figura 7 - Previsão para 20 dias

Conclusão

Neste trabalho pretendeu-se estudar os modelos de previsão utilizados para estimar a evolução da pandemia. A previsão é um auxiliar precioso no apoio à tomada de decisão dos governos e na definição de estratégias para o controlo de uma epidemia. O objetivo é sempre impedir o rápido crescimento do número de casos e prevenir a sobrecarga dos sistemas de saúde.

No caso da atual pandemia COVID-19, tem-se optado por medidas de contenção, como é o caso do confinamento total de um país, medidas essas que têm custos económicos e sociais substanciais e que, por isso, precisam de ser constantemente avaliadas e aferidas.

O maior desafio na modelização de uma pandemia é definir a precisão dos modelos e avaliar o horizonte temporal em que as suas previsões são fiáveis. A precisão dos modelos é condicionada pelo conhecimento atual do vírus. A indeterminação relativa à existência de imunidade, ao valor real da taxa de transmissão e às condições de transmissão limitam no tempo a validade dos resultados obtidos sobre a previsão do futuro da pandemia COVID-19.

A esta incerteza intrínseca à construção de modelos junta-se a complexidade da própria pandemia que, por vezes, obriga a estimar alguns dos parâmetros que definem os modelos, o que, mais uma vez, dificulta a avaliação da fiabilidade dos modelos. A fiabilidade dos modelos preditivos requer uma comparação empírica das previsões da modelação com observações reais, o que serve, também, para avaliar e afinar os modelos epidemiológicos.

A análise dos vários modelos de previsão mostra que mesmo modelos com um maior número de parâmetros nem sempre estão associados a melhores resultados. A troca de modelos mais simplistas por modelos mais complexos nem sempre é benéfica do ponto de vista da acuidade preditiva. A introdução de um grande número de parâmetros para modelizar a dinâmica da pandemia tem um elevado impacto ao nível do aumento dos custos e dificulta a possibilidade de utilizar esses modelos em países ou locais que apresentem características pandémicas diferentes daquelas para as quais foram inicialmente concebidos.

O principal fator que condiciona, de momento, o sucesso dos modelos de previsão prende-se com a manutenção das condições sobre as quais os modelos foram criados. O recurso a uma maior variabilidade e volume de dados aumenta a fiabilidade dos modelos, mas, por exemplo, em modelos de previsão que utilizaram os dados relativos ao início da pandemia, apresentam desvios significativos quando comparados com modelos que usam dados mais recentes. A quantidade e a qualidade dos dados usados na construção do modelo é, pois, muito importante para garantir fiabilidade a longo prazo.

A maioria dos modelos de previsão usados na descrição da atual pandemia são modelos determinísticos. Estes modelos dividem-se em grandes grupos: modelos SEIR e SIR, modelos

exponenciais e logísticos, modelos que se baseiam em *deep learning* entre outros. Os modelos que apresentam melhores resultados são aqueles que estimam os valores desconhecidos através de técnicas de modelagem, como é o caso de alguns modelos SEIR que recorrem ao uso de inferências bayesianas.

A análise de todos os modelos mostra que, em termos gerais, os que apresentam melhores resultados são os modelos tipo SEIR com parâmetros dinâmicos, que fornecem resultados a menos de 20 dias melhores que quaisquer outros modelos.

É importante lembrar, tal como foi dito anteriormente, que todos os modelos têm as suas limitações e que, por isso, não é prudente usar um único tipo de modelo. Além disso, há sempre que analisar, com cuidado, os pressupostos e os casos base utilizados de maneira a poder avaliar o grau de qualidade das previsões.

Medidas como confinamento social para conter a pandemia podem ser avaliadas com recurso aos modelos de previsão que parecem corroborar este tipo de decisões levando a uma diminuição da velocidade de propagação da doença.

Uma vez que a evolução da pandemia está ainda a decorrer, vários modelos surgiram ao longo do processo de elaboração deste trabalho. Desta maneira, é necessário que a renovação da análise dos modelos de previsão seja efetuada regularmente. Só quando a pandemia COVID-19 passar é que poderemos concluir com certeza se os modelos de previsão foram precisos, ao contrário do que aconteceu com outras pandemias. Até lá, é essencial que os investigadores continuem a elaborar e a atualizar os modelos de previsão de maneira a que estes se tornem o mais precisos possível e que possam informar melhor o público/entidades.

De acordo com o estudo-caso realizado, tal como anteriormente referido, o modelo não consegue reproduzir os dados para intervalos de tempo longos. Portanto para conseguir realizar previsões confiáveis devem ser utilizados intervalos de tempo inferiores a 20 dias. Uma das principais dificuldades vividas pela epidemiologia moderna é prever com confiança, nomeadamente ao nível do número de infetados – particularmente quando se dão alterações ligeiras em momentos de mudança da dinâmica da epidemia. De referir que, de acordo com o modelo de previsão, o qual apenas poderá ser confirmado posteriormente à realização deste trabalho, Portugal deverá contar com 27885 casos ativos infetados no dia 13 de abril de 2021.

Em conclusão, “tal como referido por um dos epidemiologistas mais proeminentes da pandemia da COVID-19, Neil Ferguson, os modelos "não são bolas de cristal" (17) e, por isso, deve ser desenvolvida uma perspetiva global sobre a pandemia que passa inicialmente pela utilização de técnicas de modelização simples, como forma de poder retratar quadros mais amplos sobre a doença.

Agradecimentos

Agradeço a todos os que me acompanharam nesta longa, mas alegre jornada. Os quais me auxiliaram a não deixar desvanecer a determinação e vontade, que tanto é necessária nesta nossa nobre profissão.

Agradeço em especial ao Doutor António Jorge Ferreira pela orientação e disponibilidade.

Referências

1. Mortality Risk of COVID-19 - Statistics and Research - Our World in Data [Internet]. Our World in Data. 2021 [cited 2021 Mar 28]. Available from: <https://ourworldindata.org/mortality-risk-covid>
2. Gola A, Arya RK, Animesh A, Dugh R. Review of Forecasting models for Coronavirus (COVID-19) pandemic in India during Country-wise Lockdowns [Internet]. medRxiv. 2020. p. 2020.08.03.20167254. Available from: <https://www.medrxiv.org/content/10.1101/2020.08.03.20167254v1%0Ahttps://www.medrxiv.org/content/10.1101/2020.08.03.20167254v1.abstract>
3. Law KB, Peariasamy K, Gill BS, Singh SSL, Sundram BM, Rajendran K, et al. Predicting the early depleting transmission dynamics of COVID-19: A time-varying SIR model. Sci Rep [Internet]. 2020;1–11. Available from: <https://doi.org/10.1038/s41598-020-78739-8>
4. Jebiril N. World Health Organization Declared a Pandemic Public Health Menace: A Systematic Review of the Coronavirus Disease 2019 “COVID-19.” SSRN Electron J. 2020;(May).
5. Decreto do Governo que regulamenta o estado de emergência - XXII Governo - República Portuguesa [Internet]. XXII Governo - República Portuguesa. 2020 [cited 2021 Mar 28]. Available from: <https://www.portugal.gov.pt/pt/gc22/comunicacao/documento?i=decreto-do-governo-que-regulamenta-o-estado-de-emergencia->
6. Lora Jones, Daniele Palumbo, David Brown. Coronavirus: How the pandemic has changed the world economy - BBC News [Internet]. BBC. 2021 [cited 2021 Mar 28]. Available from: <https://www.bbc.com/news/business-51706225>
7. COVID-19: Fake News | Europol [Internet]. Europol. 2020 [cited 2021 Mar 28]. Available from: <https://www.europol.europa.eu/covid-19/covid-19-fake-news>
8. About Variants of the Virus that Causes COVID-19 | CDC [Internet]. CDC. 2021 [cited 2021 Mar 28]. Available from: <https://www.cdc.gov/coronavirus/2019-ncov/transmission/variant.html>
9. Carolina Quaresma. “Expectativas.” Portugal pode conseguir imunidade de grupo em agosto [Internet]. TSF. 2021 [cited 2021 Mar 28]. Available from: <https://www.tsf.pt/portugal/sociedade/portugal-pode-conseguir-imunidade-de-grupo-em-agosto-13379442.html>
10. Wise J. Covid-19: Pfizer BioNTech vaccine reduced cases by 94% in Israel, shows peer reviewed study. BMJ [Internet]. 2021;372(February):n567. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/33632676>
11. E. Cheynet. Generalized SEIR Epidemic Model (fitting and computation) [Internet]. 2021 [cited 2021 Mar 30]. Available from: <https://www.mathworks.com/matlabcentral/fileexchange/74545-generalized-seir-epidemic-model-fitting-and-computation>
12. Rader B, Scarpino S V., Nande A, Hill AL, Adlam B, Reiner RC, et al. Crowding and

- the shape of COVID-19 epidemics. *Nat Med* [Internet]. 2020;26(12):1829–34. Available from: <http://dx.doi.org/10.1038/s41591-020-1104-0>
13. Castro M, Ares S, Cuest JA, Manrubia S. The turning point and end of an expanding epidemic cannot be precisely forecast. *Proc Natl Acad Sci U S A*. 2020;117(42):26190–6.
 14. Attanayake AMCH, Perera SSN, Jayasinghe S. Phenomenological Modelling of COVID-19 epidemics in Sri Lanka, Italy and Hebei Province of China. *medRxiv*. 2020;2020(March).
 15. Taghizadeh L, Karimi A, Heitzinger C. Uncertainty quantification in epidemiological models for COVID-19 pandemic. *medRxiv*. 2020.
 16. Pacheco-Barrios K, Cardenas-Rojas A, Giannoni-Luza S, Fregni F. COVID-19 pandemic and Farr's law: A global comparison and prediction of outbreak acceleration and deceleration rates. *PLoS One* [Internet]. 2020;15(9 September):1–25. Available from: <http://dx.doi.org/10.1371/journal.pone.0239175>
 17. Kriston L. Predictive accuracy of a hierarchical logistic model of cumulative SARS-CoV-2 case growth. *medRxiv*. 2020;2:1–9.
 18. Curado EMF, Curado MR. A discrete-time-evolution model to forecast progress of Covid-19 outbreak. *PLoS One* [Internet]. 2020;15(10 October):1–14. Available from: <http://dx.doi.org/10.1371/journal.pone.0241472>
 19. Mohamadou Y, Halidou A, Kapen PT. A review of mathematical modeling, artificial intelligence and datasets used in the study, prediction and management of COVID-19. *Appl Intell*. 2020;50(11):3913–25.
 20. Sun J, Chen X, Zhang Z, Lai S, Zhao B, Liu H, et al. Forecasting the long-term trend of COVID-19 epidemic using a dynamic model. *Sci Rep* [Internet]. 2020;1–10. Available from: <https://doi.org/10.1038/s41598-020-78084-w>
 21. Guan J, Wei Y, Zhao Y, Chen F. Modeling the transmission dynamics of COVID-19 epidemic: a systematic review. *J Biomed Res*. 2020;34(6):422.
 22. Nadella P, Swaminathan A, Subramanian S V. Forecasting efforts from prior epidemics and COVID-19 predictions. *Eur J Epidemiol* [Internet]. 2020;35(8):727–9. Available from: <https://doi.org/10.1007/s10654-020-00661-0>
 23. Zheng Z, Wu K, Yao Z, Zheng J, Chen J. The prediction for development of COVID-19 in global major epidemic areas through empirical trends in China by utilizing state transition matrix model. *medRxiv*. 2020;1–12.
 24. Utamura M, Koizumi M, Kirikami S. An epidemiological model considering isolation to predict COVID-19 trends in Tokyo, Japan: Numerical analysis. *JMIR Public Heal Surveill*. 2020;6(4):1–19.
 25. Rahimi I, Chen F, Gandomi AH. A Review on COVID-19 Forecasting Models. :1–26.
 26. Zhou T, Yuan JI. Semiparametric Bayesian inference for the transmission dynamics of COVID- 19 with a state-space model. Vol. 14(4), *arXiv*. 2020. p. 337–9.