

Evolução do número de casos de COVID-19 em Portugal

relatório de *nowcasting*



Dia 17-07-2020

Sumário

Resumo

Curva epidémica e
parâmetros de
transmissibilidade
 R_0 e $R(t)$

Nacional

Norte

Centro

Lisboa e Vale do
Tejo

Alentejo

Algarve

Métodos

Dados

Imputação

Nowcasting

Cálculo de
parâmetros
epidemiológicos

Referências bibliográficas

Dados atualizados a
16-07-2020

Resumo

Contexto

O objetivo deste relatório de situação é apresentar as estimativas da curva epidémica da infeção por SARS-CoV-2 por data de início de sintomas e as estimativas dos parâmetros de transmissibilidade R_0 (número básico de reprodução) e $R(t)$ (número de reprodução efetivo em função do tempo).

Métodos

- 1) Correção da curva epidémica de SARS-CoV-2/COVID-19 em Portugal para o atraso de notificação de casos confirmados (*Nowcasting*);
- 2) Estimativa dos parâmetros epidemiológicos número básico de reprodução (R_0) e número efetivo de reprodução ($R(t)$).

Resultados e interpretação

Após a correção da curva epidémica para o atraso de notificação, com base nos dados recolhidos, **estima-se que até 14-07-2020 tenham ocorrido 48817 casos.**

O **número médio de casos secundários** resultantes de um caso infetado (R_0) para o período de 21-02-2020 a 16-03-2020 **foi de 2,07**. Estes valores são superiores a 1 indicando, claramente, um crescimento epidémico, na fase inicial da epidemia.

O **número médio de casos secundários** resultantes de um caso infetado, medido em função do tempo ($R(t)$) variou entre 0,81 e 2,4.

O valor médio do $R(t)$ nos últimos 5 dias de análise (10-14 de julho) **foi de 0,96 ao nível nacional**. Este valor indica que o número de novos casos a cada geração é aproximadamente constante. Foram ainda estimados os seguintes valores para as regiões com mais casos - **1,04 na região Norte, 0,99 na região Centro, 0,94 na região LVT, 0,89 na região do Alentejo e 1,17 na região do Algarve.**

Uma nota especial para a região Norte que apresenta o valor do $R(t)$ acima de 1 desde 13 de junho, tendo variado entre 1 e 1,17.

Para as regiões autónomas da Madeira e dos Açores, os valores dos parâmetros de transmissibilidade R_0 e $R(t)$ não são incluídos neste relatório devido ao reduzido número de casos reportado.

Contribuições: Baltazar Nunes, Constantino Caetano, Liliana Antunes, Ana Rodrigues, Carlos Dias (Departamento de Epidemiologia, Instituto Nacional de Saúde Doutor Ricardo Jorge); Manuel Carmo Gomes (Faculdade de Ciências, Universidade de Lisboa); Luísa Morgado (Departamento de Matemática, Universidade de Trás-os-Montes e Alto Douro).

Preparação dos dados: Rita Sá Machado, Cristina Costa, Heloise Lucaccioni, Mariana Perez Duque (Divisão de Serviços de Informação e Análise, Direção-Geral da Saúde).

Agradecimentos: Marília Antunes (Faculdade de Ciências, Universidade de Lisboa); Susana Silva (Departamento de Epidemiologia, Instituto Nacional de Saúde Doutor Ricardo Jorge).

O número básico de reprodução é um indicador da transmissibilidade da infeção. Deve ser calculado na fase inicial da epidemia, ainda sem todas as medidas de contenção e atraso implementadas. Corresponde ao número médio de casos secundários a que cada caso dá origem, numa população completamente suscetível. **Neste caso o R_0 foi calculado com base na curva epidémica até ao dia 16.03.2020.**

A estimativa obtida para o R_0 foi de **2,07** podendo o verdadeiro valor estar **entre 1,96 a 2,18**, com uma **confiança 95%**.

O número médio de casos secundários resultantes de um caso infetado, medido em função do tempo ($R(t)$) deve ser calculado ao longo da epidemia e mede a transmissão ao longo do tempo. Pode ser usado para medir a efetividade das medidas de contenção e atraso.

A **estimativa do $R(t)$ variou entre 0,81 e 2,4**, observando-se uma tendência de decréscimo desde o dia 12.03 (anúncio fecho das escolas), com quebras mais acentuadas em 16.03 (fecho das escolas) e 18.03 (anúncio do estado de emergência).

A partir do final do mês de abril, tem-se observado um aumento do valor do $R(t)$, mantendo-se este abaixo ou muito perto de 1. No entanto, **com base nos dados atuais, desde meio de maio que se mantém acima ou muito próximo de 1.**

Na Região Lisboa e Vale do Tejo, o $R(t)$ esteve acima de 1 desde 29 de abril até 6 de junho, tendo variado nesta fase entre 1,01 e 1,14. Atualmente, os últimos 8 dias analisados (07 a 14 de julho) apresentam valores do $R(t)$ próximos ou inferiores a 1.

O valor médio do $R(t)$ nos últimos 5 dias de análise foi de **0,96 ao nível nacional**. Este valor indica que o número de novos casos a cada geração é aproximadamente constante. Foram ainda estimados os seguintes valores para as regiões com mais casos **1,04 na região Norte, 0,99 na região Centro, 0,94 na região LVT, 0,89 na região do Alentejo e 1,17 na região do Algarve.**

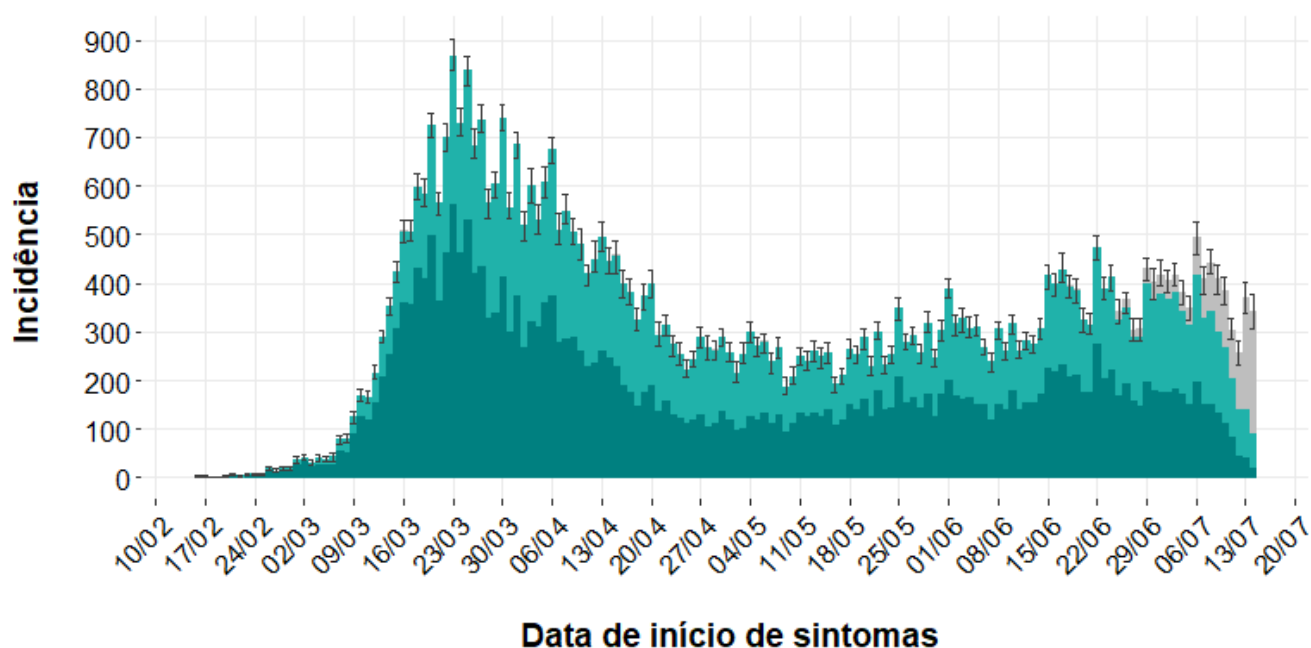


Figura 1: Curva epidémica dos casos de infeção por SARS-CoV-2 corrigida para o atraso da notificação, em Portugal. (Verde escuro - casos observados com data de início de sintomas; verde claro - casos observados com data de início de sintomas imputada; cinzento - estimativa dos casos ocorridos mas ainda não reportados).

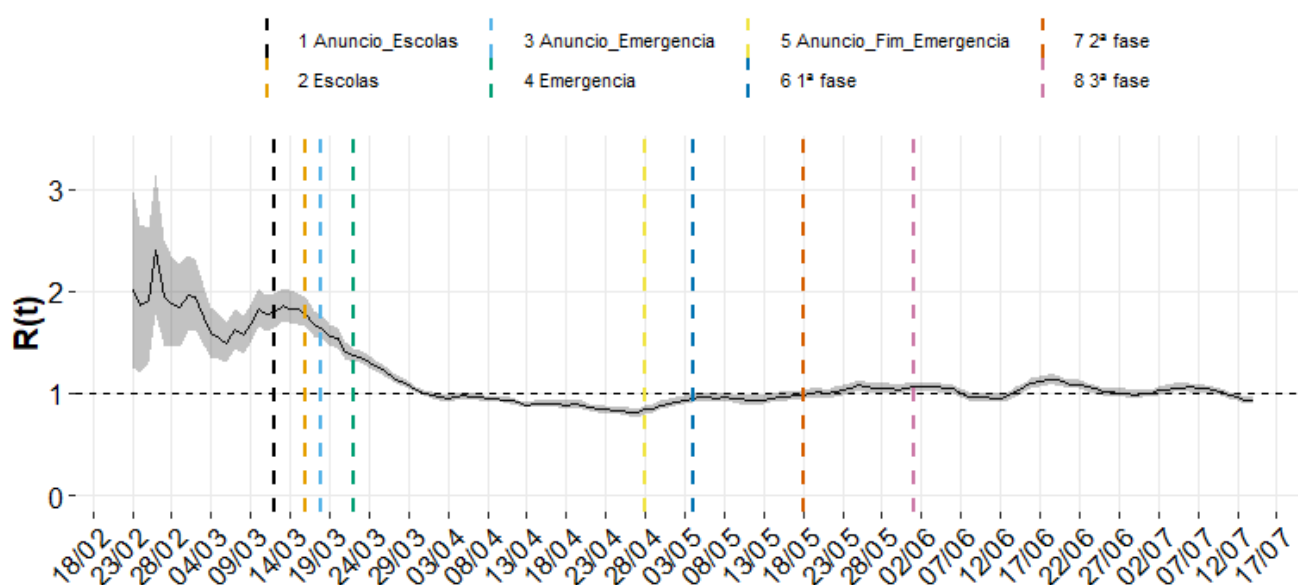


Figura 2: Evolução do valor do índice $R(t)$ em Portugal.

A média do $R(t)$ para os dias 10-07-2020 a 14-07-2020 foi de 0,96, estando o seu verdadeiro valor compreendido entre 0,95 e 0,98 com 95% de confiança.

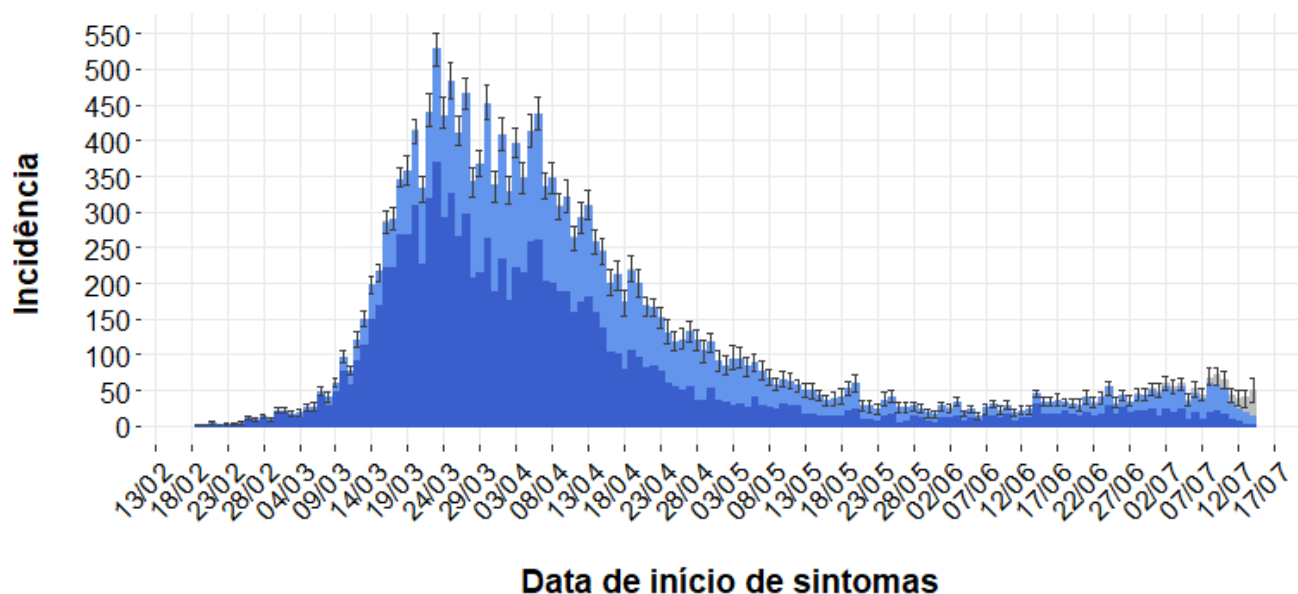


Figura 3: Curva epidémica original de casos de infeção por SARS-CoV-2, na região Norte, com data de início de sintomas reportada. (Azul escuro - casos observados com data de início de sintomas; azul claro - casos observados com data de início de sintomas imputada; cinzento - estimativa dos casos ocorridos mas ainda não reportados).

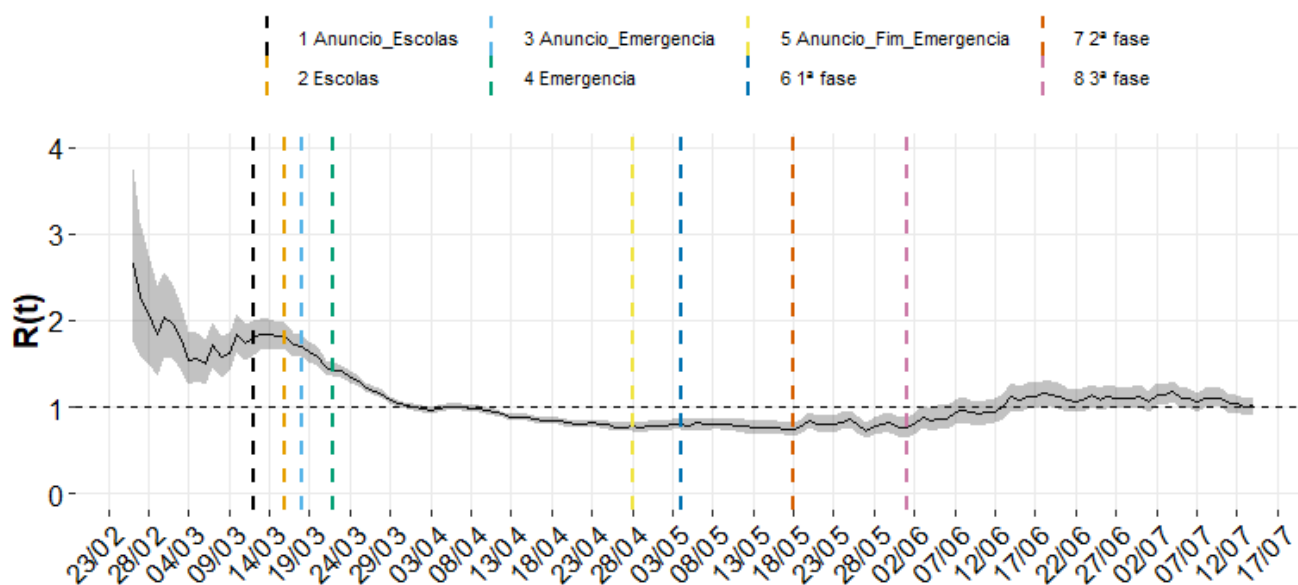


Figura 4: Evolução do $R(t)$ para a região Norte

A média do $R(t)$ para os dias 10-07-2020 a 14-07-2020 foi de 1,04, estando o seu verdadeiro valor compreendido entre 1 e 1,08 com 95% de confiança.

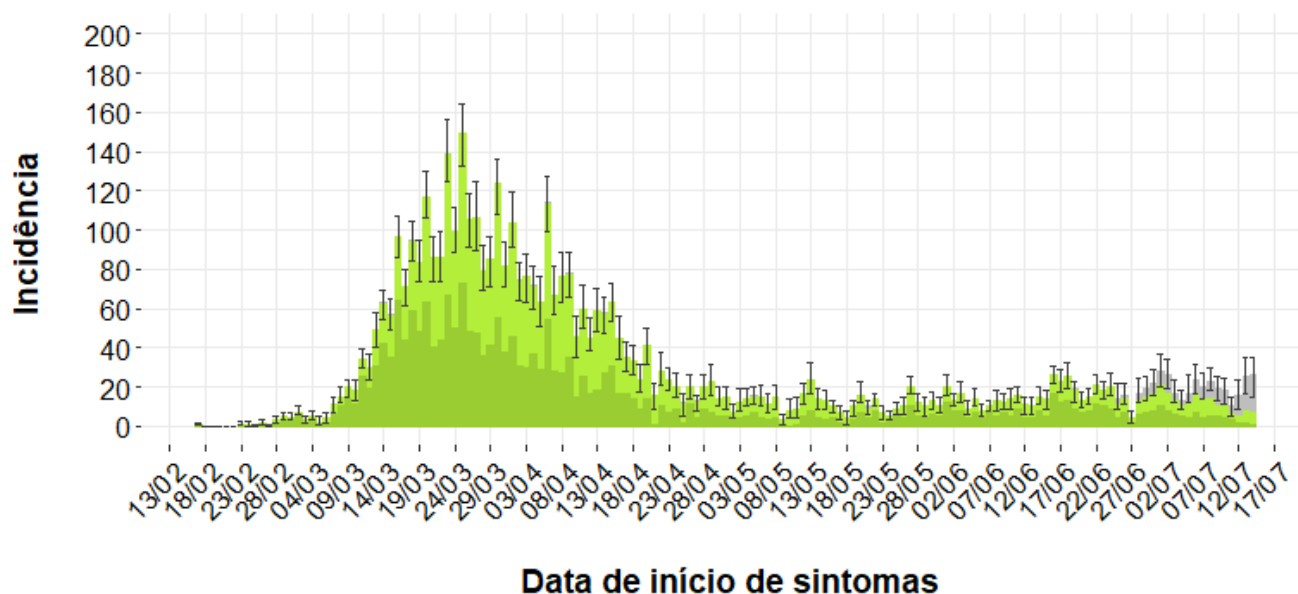


Figura 5: Curva epidémica original de casos de infeção por SARS-CoV-2, na região Centro, com data de início de sintomas reportada. (Verde escuro - casos observados com data de início de sintomas; verde claro - casos observados com data de início de sintomas imputada; cinzento - estimativa dos casos ocorridos mas ainda não reportados).

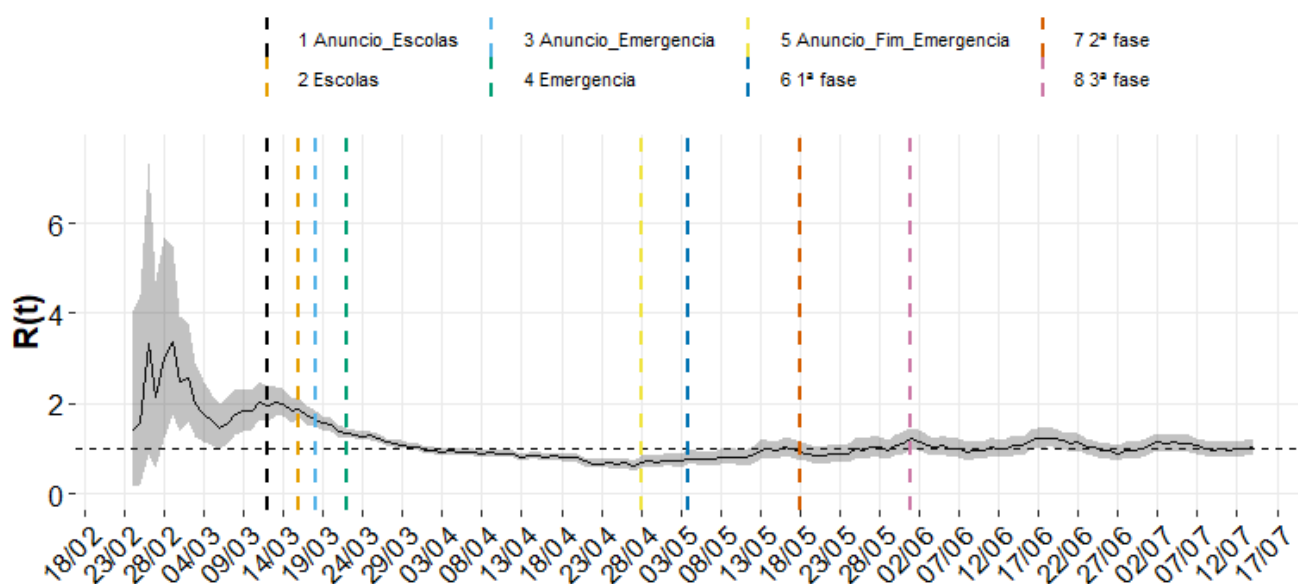


Figura 6: Evolução do $R(t)$ para a região Centro

A média do $R(t)$ para os dias 10-07-2020 a 14-07-2020 foi de 0,99, estando o seu verdadeiro valor compreendido entre 0,94 e 1,06 com 95% de confiança.

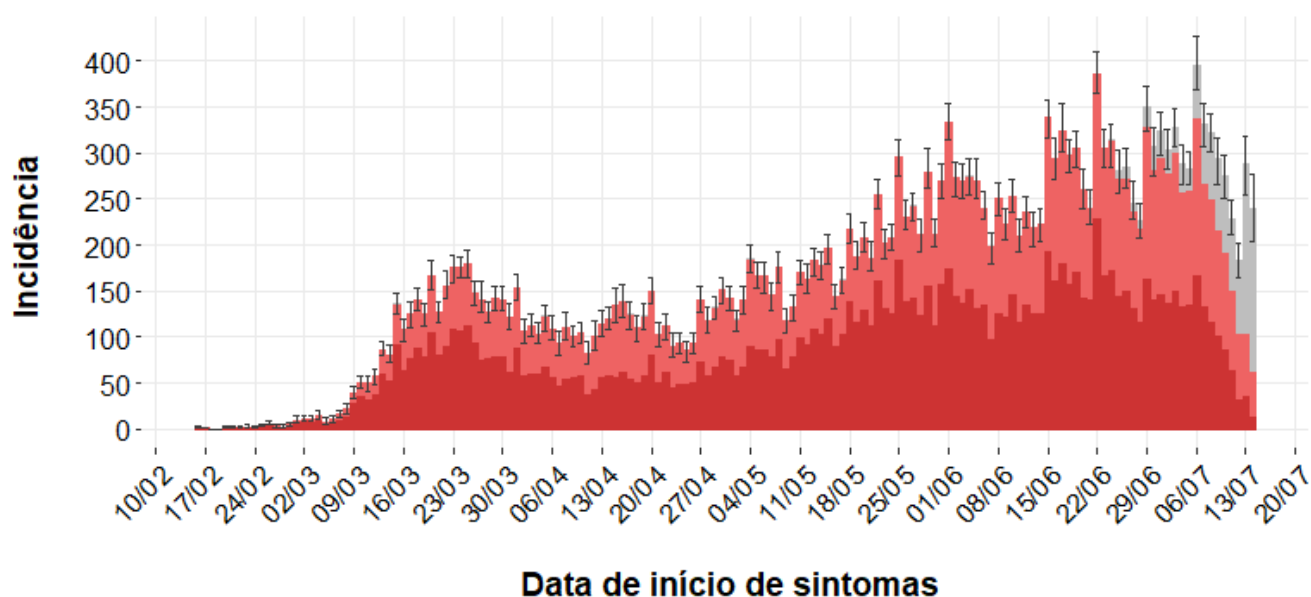


Figura 7: Curva epidémica original de casos de infeção por SARS-CoV-2, na região LVT, com data de início de sintomas reportada. (Vermelho escuro - casos observados com data de início de sintomas; vermelho claro - casos observados com data de início de sintomas imputada; cinzento - estimativa dos casos ocorridos, mas ainda não reportados).

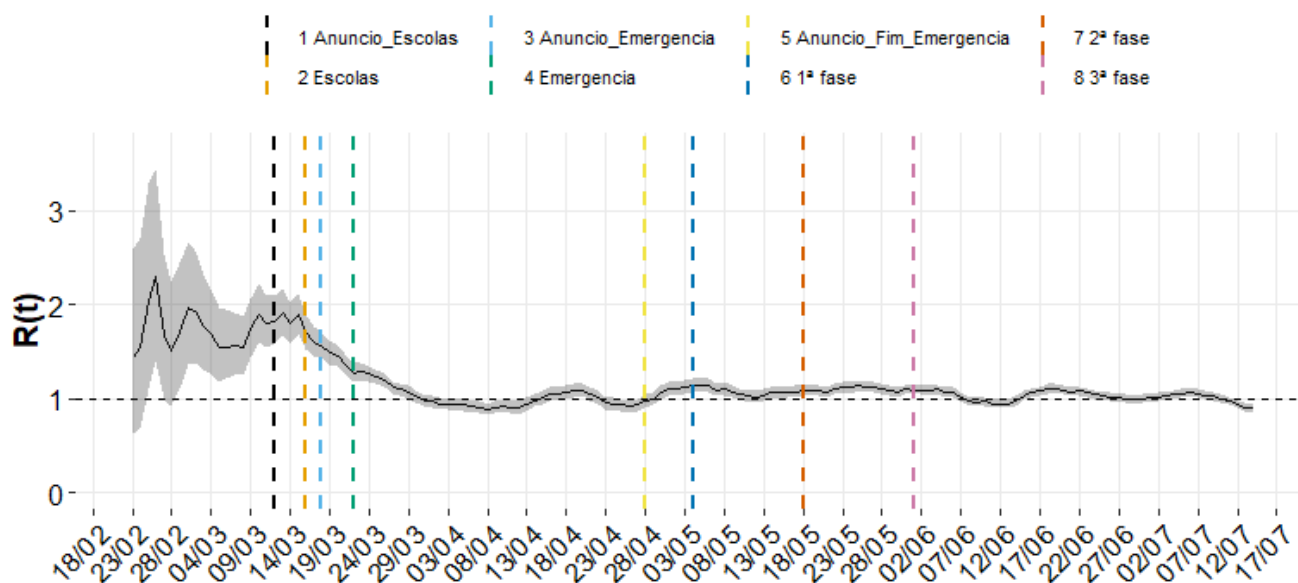


Figura 8: Evolução do $R(t)$ para a região LVT.

A média do $R(t)$ para os dias 10-07-2020 a 14-07-2020 foi 0,94, estando o seu verdadeiro valor compreendido entre 0,93 e 0,96 com 95% de confiança.

Alentejo

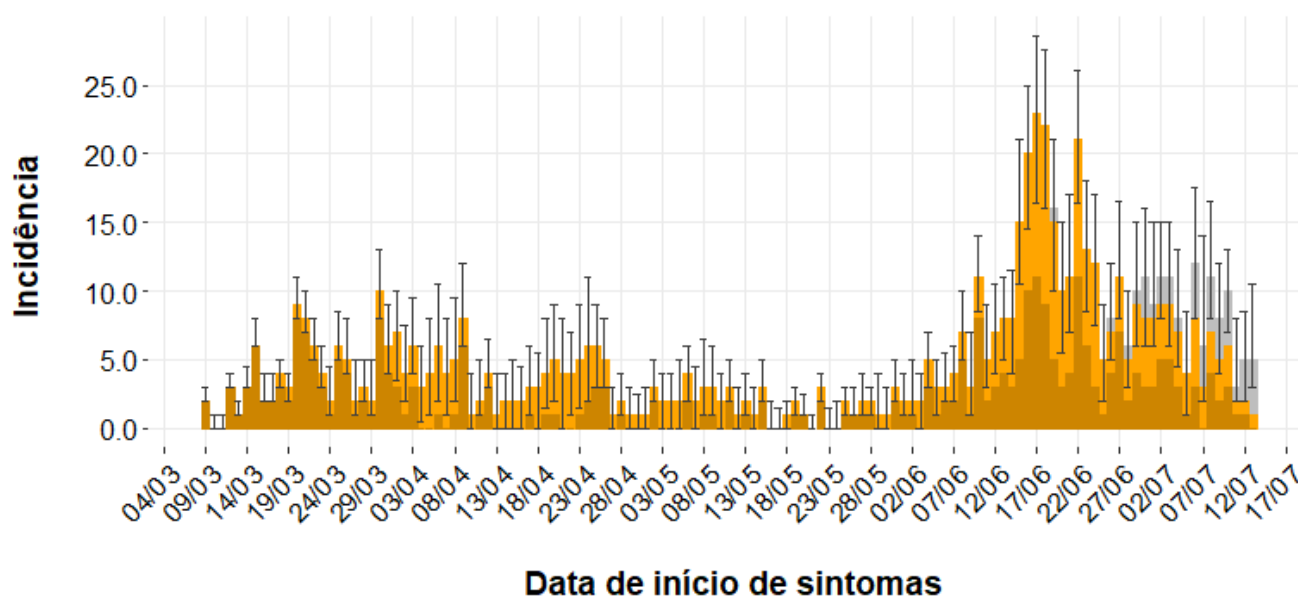


Figura 9: Curva epidémica dos casos de infeção por SARS-CoV-2 corrigida para o atraso de notificação, na região do Alentejo. (Laranja escuro - casos observados com data de início de sintomas; Laranja claro - casos observados com data de início de sintomas imputada; cinzento - estimativa dos casos ocorridos mas ainda não reportados).

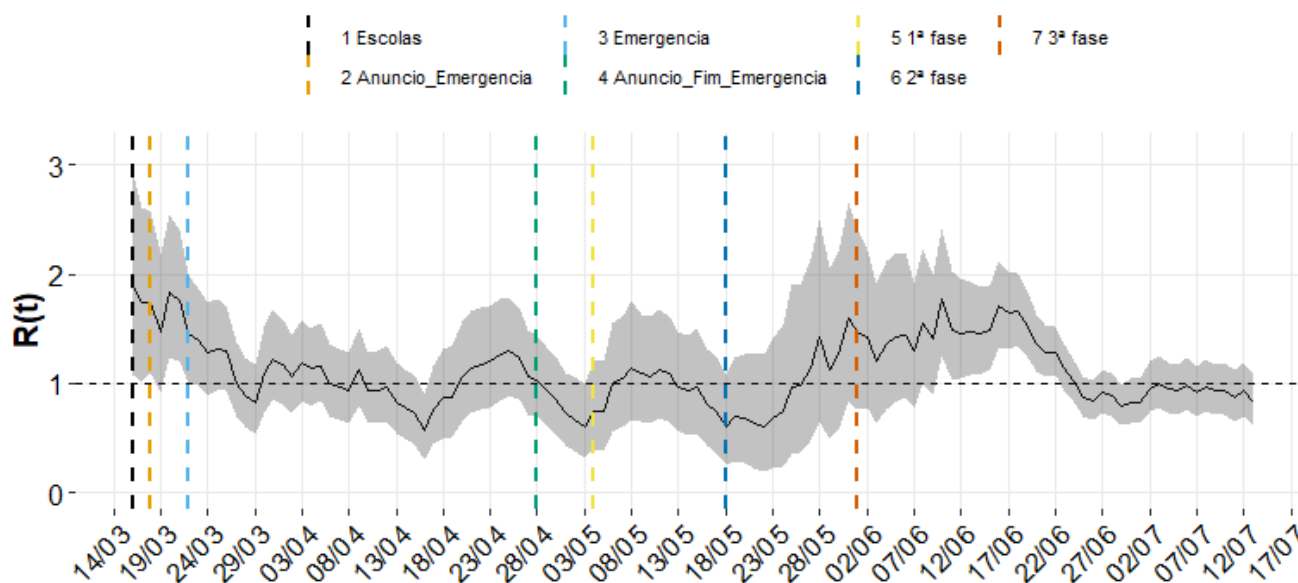


Figura 10: Evolução do $R(t)$ para a região do Alentejo

A média do $R(t)$ para os dias 09-07-2020 a 13-07-2020 foi de 0,89, estando o seu verdadeiro valor compreendido entre 0,8 e 0,99 com 95% de confiança.

Algarve

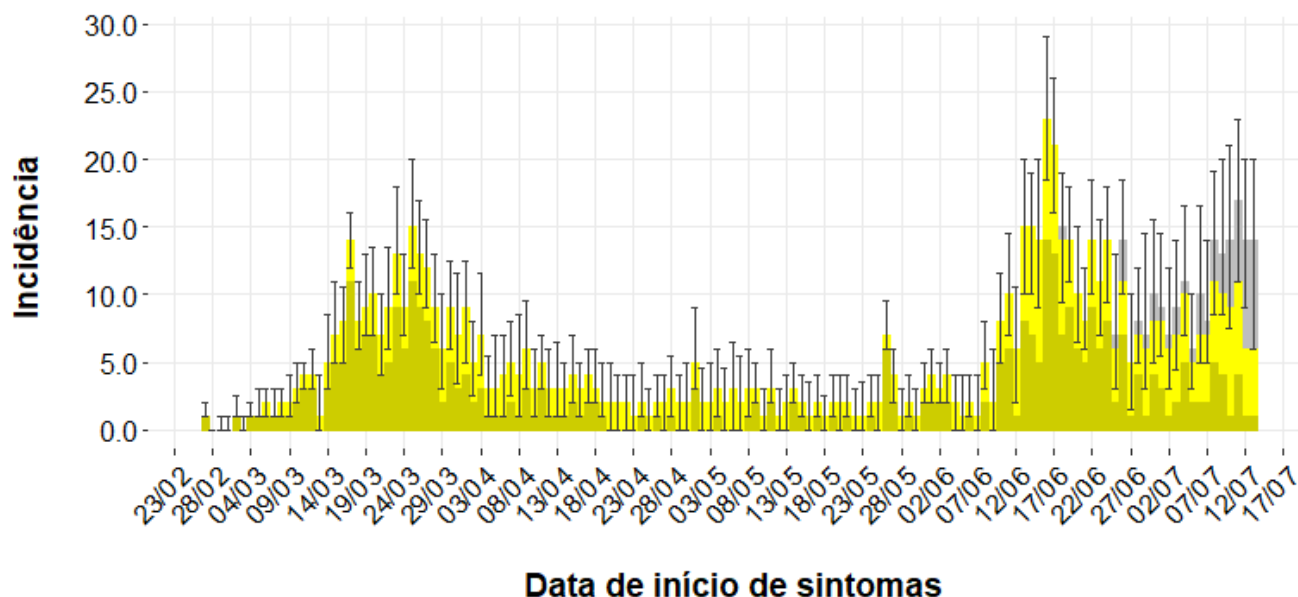


Figura 11: Curva epidémica dos casos de infeção por SARS-CoV-2 corrigida para o atraso de notificação, na região do Algarve. (Amarelo escuro - casos observados com data de início de sintomas; Amarelo claro - casos observados com data de início de sintomas imputada; cinzento - estimativa dos casos ocorridos mas ainda não reportados)

Algarve

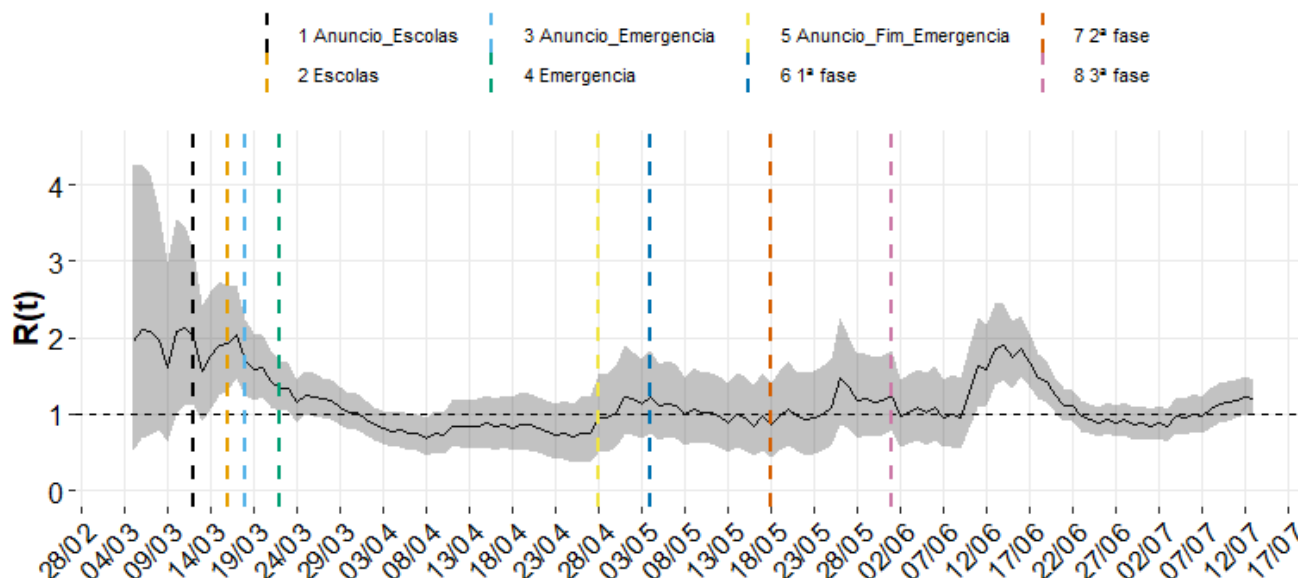


Figura 12: Evolução do $R(t)$ para a região do Algarve

A média do $R(t)$ para os dias 09-07-2020 a 13-07-2020 foi de 1,17, estando o seu verdadeiro valor compreendido entre 1,09 e 1,28 com 95% de confiança.

Dados

No dia 16-07-2020 o grupo de trabalho da DGS uma tabela com a data de início de sintomas e a data de confirmação de 47395 casos de COVID-19. As datas de início de sintomas variavam entre 20-02-2020 a 15-07-2020. A data de início de sintomas era omissa em 21283 casos (45%), a data de confirmação laboratorial era também omissa em 4205 casos (9%) e 4193 casos tinham ambas as datas em falta (9%).

Imputação

Nos casos com data de diagnóstico laboratorial em falta, é assumida a data da notificação do caso. A imputação da data de início de sintomas recorre à distribuição do tempo entre o início de sintomas e o diagnóstico, nos casos em que esta informação se encontra disponível. Quando a informação sobre a região de saúde que reportou o caso e o grupo etário está disponível, é calculada a distribuição empírica estratificada pelas variáveis supramencionadas e pelo dia da semana da data de diagnóstico, calculada com base nos dados observados (completos) para uma janela temporal de 15 dias, que se inicia 15 dias antes da data de diagnóstico do caso.

Esta decisão prende-se com uma possível alteração da distribuição do tempo entre a data de início de sintomas e a data do diagnóstico ao longo do tempo, por região e por grupo etário. A escolha da janela temporal foi baseada numa análise preliminar dos dados que apontava para que 90% dos casos relativos a um dado dia sejam diagnosticados até 15 dias após a data de início de sintomas.

A atribuição do atraso na notificação com base na distribuição é obtida aleatoriamente, gerando-se 100 possibilidades para a data de início de sintomas de cada caso.

Nowcasting

A estimação do número de casos de COVID-19 já ocorridos na população (início de sintomas) mas que ainda não foram diagnosticados, para esse efeito recorre ao método desenvolvido por Antunes L; Antunes M, Nunes B [1].

Cálculo de parâmetros epidemiológicos

As estimativas do R_0 foram obtidas pelo pacote R_0 do programa R. Este método foi desenvolvido com base no trabalho de White et al. (2009) [2]. Neste artigo, os autores utilizam métodos de máxima verosimilhança para estimar o valor R_0 , assumindo que a distribuição do intervalo de tempo entre o início de sintomas de um infetante e de um infetado (*serial interval*) é conhecido e que o número de infeções causadas por um infetado segue uma distribuição *Poisson*.

As estimativas do $R(t)$ em função do tempo (em janelas de 7 dias) foi obtida pelo método implementado no *package* EpiEstim [3]. Este método calcula o número efetivo de reprodução com base na curva dos novos casos e na distribuição do *serial interval*. Considerou-se para o *serial time* e uma distribuição Gama com os parâmetros (média de 3,96 e desvio-padrão de 4,75) de acordo com Du Z et al (2020) [4].



- [1] Antunes L.; Antunes M, Nunes B (orient.). Modelação do atraso na notificação dos óbitos no sistema de vigilância diária da mortalidade. Dissertação de mestrado, FCUL, 2019.
<http://hdl.handle.net/10400.18/2577>
- [2] White LF, Wallinga J, Finelli L, et al. Estimation of the reproductive number and the serial interval in early phase of the 2009 influenza A/H1N1 pandemic in the USA. Influenza Other Respir Viruses. 2009;3(6):267-276. <https://doi.org/10.1111/j.1750-2659.2009.00106.x>
- [3] Cori A, Ferguson NM, Fraser C, et al. A new framework and software to estimate time-varying reproduction numbers during epidemics. Am J Epidemiol. 2013;178(9):1505-1512.
<https://doi.org/10.1093/aje/kwt133>
- [4] Du Z, Xu X, Wu Y, et al. Serial Interval of COVID-19 among Publicly Reported Confirmed Cases. Emerg Infect Dis. 2020;26(6):1341-1343. <https://doi.org/10.3201/eid2606.200357>