

Será possível estimar a real taxa de infecção em Portugal?

João Gomes, FCUL, CMAFCIO

I. Introdução

Quando ouvir que “hoje” temos “**n**” novos infetados, será que esses números são reais? Será que refletem a situação real? Ou será apenas um indicador sobre o verdadeiro número de casos de pessoas infetadas alguns dias antes?

Saber o verdadeiro número de infetados permitiria corrigir e até, talvez, alterar as medidas de saúde pública que estão a ser tomadas (1). Por exemplo ao afirmar-se que a taxa de mortalidade por COVID-19 é de 3.3%, é, na verdade, um número assustador! Mas se se perceber que o número de infetados é muito superior ao que é declarado, essa percentagem pode ser drasticamente menor e, muito, muito menos assustador!

O presente trabalho pretende propor uma reflexão que conduza gradualmente à resposta a algumas das questões apresentadas.

A motivação deste estudo surgiu depois de termos lido o pre-print (2) realizado a partir de um estudo serológico em Gangelt. Gangelt é um pequeno município alemão integrado num distrito já com alguma dimensão (Heinsberg) e onde é possível aceder ao número de casos reportados (3). Estes município/distrito são importantes porque se suspeita que um evento de carnaval possa ter provocado um efeito pandémico naquela zona antes das medidas de confinamento entretanto declaradas.

Outra motivação deve-se a que a Alemanha tem uma taxa de testes PCR¹ semelhante a Portugal (4) e por isso é natural que, para uma mesma taxa de infeções, a taxa de detetados na população seja semelhante a Portugal- ou, de outro modo, Portugal e Heinsberg são comparáveis!

¹ O método utilizado para a deteção do vírus é o PCR (Polymerase Chain Reaction), pela elevada sensibilidade e especificidade na identificação simultânea de vários genes do vírus.

Os testes até agora realizados ao COVID-19 permitem estimar a taxa de infeção? Na verdade, não! Porquê?

Usando os modelos provavelmente mais discutidos actualmente neste contexto, os modelos SIR², podemos considerar que a População está dividida em três grupos: Os que ainda não foram infetados (S), os que foram infetados (I) e ainda são contagiosos por estarem doentes ou assintomáticos e os que já foram removidos (R), porque morreram ou já recuperaram. Se P_S , P_I e P_R forem, respetivamente, as probabilidades de determinado indivíduo pertencer a S, I e R, calculemos a probabilidade de um determinado teste ser positivo (T+), caso a escolha da pessoa a testar seja aleatória

$$P(T+) = P(T+|S)P_S + P(T+|I)P_I + P(T+|R)P_R = P(T+|I)P_I$$

Ora se a sensibilidade³ do teste for de, digamos, 60%, o valor obtido seria de apenas $0.6 * P_I \ll 0.1!$

Se por outro lado o teste for dirigido a pessoas com alguns sintomas, ou por suspeita de terem estado em contacto com pessoas infetadas, esse valor, $P(T+|Suspeito)$, será bastante superior e tem registado um valor próximo dos 0.1 (a percentagem de testes positivos entre os suspeitos de COVID em Portugal) (5).

Qual é o problema desta metodologia? É que as pessoas assintomáticas nunca serão testadas a não ser que, em determinada altura, possam fazer parte do grupo de suspeitos. No entanto a janela de oportunidade para serem PCR positivos é relativamente apertada (cerca de 5 dias) e por isso é natural que muitos deles passem pelo "crivo" dos testes. Mas quantos?

Será esse o nosso desafio neste trabalho - estimar o número de pessoas que não são notificadas apesar de poderem ter sido já infetadas!

Vamos começar por fazer inferência para Heinsberg, com base no estudo serológico publicado na Universidade Alemã de Bonn (2), e a partir dessa inferência extrapolar para Portugal com base em alguns pressupostos que enumeraremos mais à frente (6) (7) (8).

² [Modelos SIR](#)

³ proporção de pessoas infetadas que dão positivo no teste

II. Estudo piloto em Gangelt, Heinsberg, Alemanha

A. Município-Gangelt



Figura 1: Município de Gangelt integrado no distrito de Heinsberg (9)

Gangelt é um pequeno município no distrito de Heinsberg

População: 12529 habitantes

i. Amostra

- **Amostra⁴:** 1000 habitantes de 400 famílias
- **Sub-amostra avaliada e já publicada:** 500 habitantes
- **Infetados ativos detetados:** 10
- **Sero-positivos:** 75

ii. Inferência

O trabalho refere uma especificidade⁵ muito elevada (>99%) mas não refere qual a sensibilidade do teste e por isso em todo este trabalho vamos partir de um pressuposto: a sensibilidade deste teste estará algures entre os 50 e os 100%.

- **% de infetados na população:** Esta % depende da sensibilidade do teste e em função disso temos cenários entre os seguintes valores:
 - o Sensibilidade-100%: 15% de infetados
 - o Sensibilidade-50%: 30% de infetados

⁴ A amostra está certificada por [Matthias Schmid](#) (2)

⁵ A proporção de não infetados em que o teste dá corretamente negativo

B. Distrito-Heinsberg

Heinsberg é um distrito da parte ocidental da Alemanha da qual faz parte Gangelt.

População: 254322 habitantes

i. Inferência por analogia com Gangelt

Partindo do pressuposto de homogeneidade entre Gangelt e Heinsberg podemos inferir cenários semelhantes ao de Gangelt:

- **Infetados**
 - o Sensibilidade 100%: 38148 pessoas já estão imunes (15%)
 - o Sensibilidade 50%: 76297 pessoas já estão imunes (30%)
- **Óbitos**
 - o Óbitos entre os detetados: 52 entre 1619(3.2%)
 - o % Óbitos entre os infetados:
 - **Sensibilidade 100%:** 0.14%
 - **Sensibilidade 50%:** 0.07%

ii. Casos por escalões etários (17 abril)

Por estarmos a trabalhar com o distrito (só temos dados disponíveis ao nível do distrito) é possível ter dados agrupados por escalões etários e assim ter informação não só sobre o número de casos, mas também sobre a sua distribuição etária.

Tabela 1: Casos detetados e casos por 100000 habitantes, por escalões etários, em Heinsberg (3)

esc etar		Casos			Casos/100000		
escalão	média	M	F	Total	M	F	Total
0-4	2	5	5	10	87	83	170
5-14	10	6	8	14	71	51	122
15-34	25	172	215	387	806	582	1388
35-59	47.5	316	370	686	799	691	1490
60-79	70	192	173	365	601	714	1315
+80	87	72	84	156	868	1225	2093

Esta tabela permite avançar para uma comparação direta com Portugal não só com base no número de casos detetados, mas também pela distribuição etária. Mais à frente será fundamental corrigir a diferença da idade média na estimação do número de infetados.

Assim temos para Heinsberg, e em jeito de resumo:

- **% Casos detetados na população:** 0.6%
- **Idade média dos infetados detetados:** 50.5 anos

III. O que podemos dizer de Portugal a partir deste estudo?

A. Dados da População Portuguesa

População: 10^7 habitantes

Casos detetados em Portugal a 17/4

- Casos: 19022
- Óbitos: 657
- % de Casos na População: 0.19%
- % Óbitos/Casos: 3.5%
- Idade média nos casos: 53.5

B. Pressupostos para ser possível inferir a partir dos dados de Heinsberg

- A forma como se realizam os testes é semelhante: Ter como alvo principal os casos suspeitos e repetir sempre que um suspeito é sintomático e negativo.
- A quantidade de testes é semelhante. Isso é verdade quando se pensa em termos de país e por isso deduzimos que o seja em termos de distrito (4).
- A percentagem de casos na população é um indicador proporcional do número real de infetados.
- A idade é um factor potenciador (por cada ano que passa é mais provável ter-se sintomas da doença) para se detetar um novo caso.

C. Qual o posicionamento de Portugal (17/4) face a Heinsberg, à data do estudo?

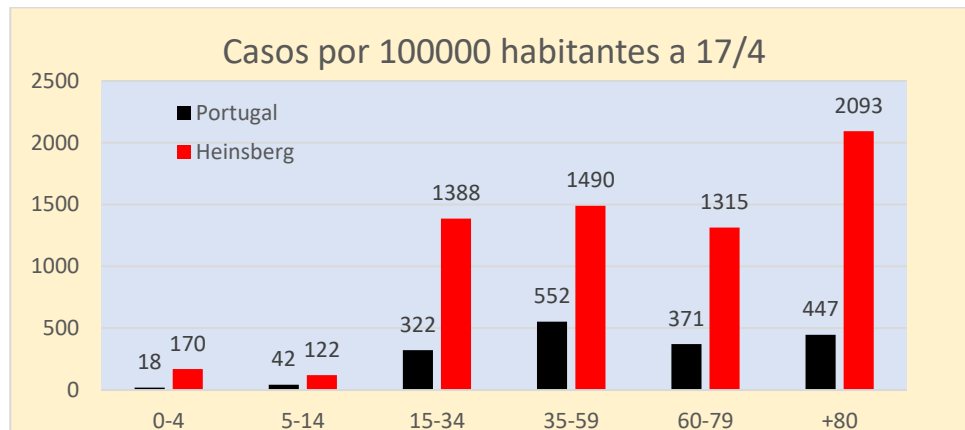


Figura 2: Comparação entre os casos (por 100000) em Portugal e Heinsberg (5) (3)

Esta figura permite perceber que o número de casos em Heinsberg é bastante superior aos casos detetados em Portugal embora em termos de idade média, Portugal tenha um valor superior. Em jeito de resumo:

- Casos por 100000: Portugal tem 1743 contra 6576 de Heinsberg (**em Portugal são cerca de 4 vezes menos casos**)
- A idade média dos casos em Portugal é de 53.5 anos contra 50.5 anos dos habitantes de Heinsberg

D. Inferência

A primeira conclusão é que estamos muito "atrasados", quando comparados com Heinsberg, relativamente ao número de casos/infetados.

Podemos avançar com alguma inferência para Portugal sobre o número de pessoas já infetadas?

Sim, desde que os nossos pressupostos sejam aceitáveis:

1. Os casos detetados são um bom indicador do número de infetados na população e, portanto, podemos inferir que o número de infetados em Portugal é de cerca de 4 vezes menos do que em Heinsberg;
2. A média de idades nos casos de Heinsberg (50.5) é 3 anos menos do que a média de idade dos casos detetados em Portugal (53.5 anos) e como a idade é um factor potenciador para a deteção, é natural que a percentagem de casos detetados em Portugal seja maior, quando comparado com o mesmo número de infetados.
3. Quanto e de que modo se corrige?

- a. Vamos utilizar como "proxy" o crescimento anual de letalidade avaliada num outro trabalho ainda não publicado (10% por ano)
- b. Ajustar a % de infetados estimados para Portugal corrigido do "efeito de confundimento" pela idade

IV. Conclusões do estudo (17/4/2020)⁶

Pelo exposto e respondendo à questão de partida, tendo como base o estudo feito em Gangelt, extrapolado para Heinsberg (distrito do qual faz parte Gangelt), e a partir dos quais se inferiu para Portugal, quando ouvimos dizer que o número de casos infetados hoje é "n", esse, não será certamente um valor real porque lá, tal como deverá ser por cá, a percentagem de realmente infetados (15 a 30%) é muito superior aos 0.6% detetados!

Vejamos novamente um pequeno resumo dos dados portugueses a 17 de Abril de 2020

Tabela 2: Portugal - infetados detetados e óbitos declarados

Portugal	
infetados detetados	19022
óbitos na população	657
% infetados detetados na População	0.19%
% mortos/detetados	3.5%
idade média dos detetados	53.5

E com base nos valores encontrados em Gangelt e extrapolados para Heinsberg poderemos inferir para Portugal, em função da sensibilidade do teste serológico utilizado em Gangelt, o seguinte número de infetados:

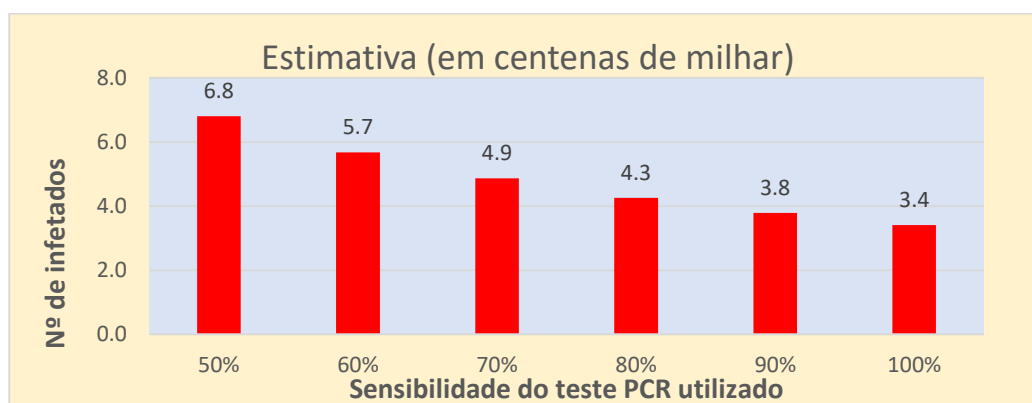


Figura 3: Número de infetados ($\times 10^5$) estimados para Portugal a 17 de Abril
Em jeito de síntese:

Se acreditarmos que a taxa de deteção em Portugal é semelhante à de Heinsberg (essa hipótese não será de excluir pois o número de testes

⁶ "They think if people avoid getting big doses of the virus—which can happen in hospitals or via close contact with someone infected—fewer people will become severely ill, "while at the same time developing immunity" that can help finally end the outbreak." (7)

por milhão em Portugal é semelhante aos valores da Alemanha (4)) poderíamos inferir que a relação entre o número de casos detetados e a realidade teria uma taxa semelhante e, nesse caso, teríamos cerca de $\frac{1}{4}$ de infetados relativamente a Heinsberg o que nos colocaria entre os 340000 e os 680000 de infetados. Este valor é demasiado elevado? Será mesmo?

Ora vejamos: Para termos uma imunidade de grupo sem vacina - isto é, contando apenas com a imunidade provocada pelo contacto com o vírus - necessitaríamos de cerca de 7 milhões de pessoas (10) e, se quisermos ser menos exigentes, precisaríamos de pelo menos metade (3.5 milhões) para que o risco de velocidade pandémica fosse relativamente baixa.

Nessa perspetiva dizermos que temos 500000 pessoas que já entraram em contacto com o vírus e que este valor é excessivamente alto é provavelmente um erro! Talvez fosse melhor dizer que este número é preocupante por ser demasiado pequeno quando se fala em entrar na chamada 2ª fase - a 1ª fase após as medidas de confinamento.

Seria talvez preferível pensar em como separar os grupos de risco (>65 e doentes com deficiências imunitárias) antes dessa fase.

Porque não utilizar os hotéis como ponto de partida para permitir que as pessoas pudessem ser realmente afastadas dos possíveis contágios dos familiares mais jovens? Porque não incentivar as pessoas >65 a fazerem "férias"? O que é preferível - apoiar a hotelaria para que esta pudesse praticar preços comportáveis pela nossa população mais idosa ou subsidiar os hospitais para os tratar? Os testes em massa seriam uma fonte fundamental para evitar o "efeito cruzeiro" nos hotéis. Talvez se diminuíssemos o número de testes na população de baixo risco e nos focássemos nos locais com mais potencial pandémico fôssemos muito mais eficazes: Lares, hotéis, Serviços de ambulatório, etc!

Em jeito de conclusão, este estudo tem apenas como objectivos - por um lado a reflexão e por outro sugerir que o número de infetados em Portugal é muito baixo, mas pode ser muito mais elevado do que se pensa actualmente!

Charneca da Caparica, 18 de Abril de 2020

V. Referências

- 1.DGKH. [Online]https://www.krankenhaushygiene.de/ccUpload/upload/files/2020_03_31_DGKH_Einladug_Lageeinschaetzung.pdf.
- 2.Hendrik Streeck, Gunther Hartmann, Martin Exner and Matthias Schmid. Vorläufiges Ergebnis und Schlussfolgerungen der COVID-19 Case-Cluster-Study (Gemeinde Gangelt). Bonn: Universitätsklinikum Bonn, 2020.
- 3.COVID-19-Dashboard, Robert Koch-Institut:RKI. [Online]<https://experience.arcgis.com/experience/478220a4c454480e823b17327b2bf1d4>.
- 4.Licensing,Worldometer.Worldometers. [Online]<https://www.worldometers.info/licensing/what/>.
- 5.DGS. DGS: COVID-19. [Online] <https://covid19.min-saude.pt/relatorio-de-situacao/>.
- 6.Bandera County Courier. [Online] <https://www.bccourier.com/these-are-the-first-lessons-of-the-heinsberg-study/>.
- 7.technologyreview.com/. [Online]<https://www.technologyreview.com/2020/04/09/999015/blood-tests-show-15-of-people-are-now-immune-to-covid-19-in-one-town-in-germany/?fbclid=IwAR0RiF5taXf0vYjMpQNXVOopUvM-FuisrjqRdkSXGKmVwGHqVYnJPBT74Mw>.
- 8.theguardian.com. [Online]<https://www.theguardian.com/world/2020/apr/09/the-cluster-effect-how-social-gatherings-were-rocket-fuel-for-coronavirus>.
- 9.pt.wikipedia.org/wiki/Gangelt. [Online]<https://pt.wikipedia.org/wiki/Gangelt>.
- 10.TAE HYONG KIM, JENNIE JOHNSTONE, MARK LOEB. Vaccine herd effect. s.l. : Scandinavian Journal of Infectious Diseases, 2011; 43: 683-689, 2011. pp. 683-689.