THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

The importance of saturating density dependence for population-level predictions of SARS-CoV-2 resurgence compared with density-independent or linearly density-dependent models, England, 23 March to 31 July 2020

The importance of saturating density dependence for population-level predictions of SARS-CoV-2 resurgence compared with density-independent or linearly density-dependent models, England, 23 March to 31 July 2020

Emily S Nightingale[1,2] , Oliver J Brady[1,2] , CMMID Covid-19 working group[3] , Laith Yakob[2,4]
1. Department of Global Health and Development, Faculty of Public Health and Policy, London School of Hygiene & Tropical Medicine, London, United Kingdom
2. Centre of Mathematical Modelling for Infectious Diseases, London School of Hygiene & Tropical Medicine, London, United Kingdom
3. The members of the group are listed under Investigators
4. Department of Disease Control, Faculty of Infectious & Tropical Medicine, London School of Hygiene & Tropical Medicine, London, United Kingdom

Correspondence: Laith Yakob (laith.yakob@lshtm.ac.ukv)

Background: Population-level mathematical models of outbreaks typically assume that disease transmission is not impacted by population density ('frequency-dependent') or that it increases linearly with density ('density-dependent'). **Aim:** We sought evidence for the role of population density in SARS-CoV-2 transmission. **Methods:** Using COVID-19-associated mortality data from England, we fitted multiple functional forms linking density with transmission. We projected forwards beyond lockdown to ascertain the consequences of different functional forms on infection resurgence. **Results:** COVID-19-associated mortality data from England show evidence of increasing with population density until a saturating level, after adjusting for local age distribution, deprivation, proportion of ethnic minority population and proportion of key workers among the working population. Projections from a mathematical model that accounts for this observation deviate markedly from the current status quo for SARS-CoV-2 models which either assume linearity between density and transmission (30% of models) or no relationship at all (70%). Respectively, these classical model structures over- and underestimate the delay in infection resurgence following the release of lockdown. **Conclusion:** Identifying saturation points for given populations and including transmission terms that account for this feature will improve model accuracy and utility for the current and future pandemics.
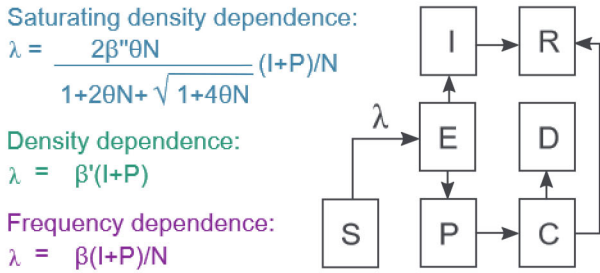
Introduction

Like many pathogens that cause respiratory diseases [1-3], severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) appears to be transmitted more effectively in densely populated areas [4-6]. The increased disease rates reported among high-density populations [4,5,7,8] may, however, be an artefact of confounders, such as a higher proportion of individuals of lower socioeconomic status or from minority ethnic groups in urban areas [9]. Using coronavirus disease (COVID-19)-associated mortality data from the Office for National Statistics, we aimed to assess the evidence for density dependence.

Standard transmission models that either do or do not account for this density dependence have been used interchangeably because their projections are generally equivalent when population density remains unperturbed or is homogeneous, e.g. at a national level. While the ca 1% infection fatality rate for COVID-19 [10] is insufficient to destabilise populations, the reaction of most countries' governments to curtail disease spread through lockdown and physical distancing has had unprecedented impacts on the density of mobile human populations. For example, the United Kingdom's lockdown, which came into effect on 23 March 2020, effectively reduced the freely moving population from 66.5 million to 10.6 million (key workers) [11]. This same intervention was employed by numerous countries, similarly impacting their mobile populations [12]. We evaluate the extent to which models built to inform the epidemiology of COVID-19 use an

SARS-CoV-2 transmission model compartments and alternative transmission assumptions, England, March–July 2020

Saturating density dependence:
$$\lambda = \frac{2\beta''\theta N}{1+2\theta N+\sqrt{1+4\theta N}}(I+P)/N$$

Density dependence:
$$\lambda = \beta'(I+P)$$

Frequency dependence:
$$\lambda = \beta(I+P)/N$$

SARS-CoV-2: severe acute respiratory syndrome coronavirus 2.

Model compartments are: 'S'usceptible, 'E'xposed, 'I'nfectious, 'P're-critical infectious, 'C'ritically ill, 'D'ead and 'R'ecovered. The three functional forms are shown for the force of infection, λ. Each is a product of the infectious proportion of the population and the transmission coefficient, β. The transmission coefficients are differentiated between the three functional forms using primes. The transmission coefficient is itself the product of the transmission probability per contact with an infectious individual and the contact rate. Under a linearly density-dependent assumption, the contact rate increases linearly with the total population, N. This numerator N cancels out the N denominator which is why β` is multiplied by the number of infectious individuals instead of the proportion. The saturating density-dependent formulation assumes the force of infection is a product of the transmission coefficient and a function that increases contact rate nonlinearly with population density as informed by parameter θ derived from analysing England's regional mortality data.

underlying structure that can accommodate the drastic changes and variation in densities experienced by most global populations.

As lockdowns were gradually released over the latter part of 2020, global populations were expected to re-equilibrate to a 'new normal' whereby densities of mobile people were increased but in which contact patterns were expected to remain reduced through physical distancing interventions [13]. Using a suite of mathematical models, we illustrate the impact that the different, routinely ignored, assumptions underlying transmission and density may have in projecting infection dynamics and measuring intervention effectiveness.

Methods

Data

Reported COVID-19-related deaths between 1 March and 31 July 2020 were obtained in anonymised linelist form from Public Health England and were filtered to include all deaths which occurred within 28 days of positive COVID-19 test (n = 36,311). We aggregated individual records to lower-tier local authority (LTLA), and nationally by 10-year age bands in order to calculate age-standardised expected counts.

Local authority shapefiles and single-age population estimates were obtained from the Office for National Statistics [14]. Four sub-regions of Buckinghamshire (Aylesbury Vale, Chiltern, South Bucks, Wycombe) were aggregated in order to match most recent population estimates. The City of London was aggregated with Westminster because of its very small resident population, and the Isles of Scilly were excluded since no COVID-19-related deaths had been reported there during the study period. Index of multiple deprivation (IMD) [15], percentage of minority ethnic population [16] and percentage of key workers among the working population [17] are characteristics of the LTLA population potentially associated with both COVID-19 mortality and population density, therefore we included them as covariates in all models. Percentage of key workers was missing for Westminster and Cornwall; these were imputed by the median value across all neighbouring LTLA.

Statistical analysis

Negative binomial regression models were fitted to the number of deaths (n) per LTLA, adjusting initially for the three covariates (IMD, % minority population, % key workers) and subsequently adding a fourth covariate, namely the lag in weeks behind the first death nationally. We adjusted for age distribution within the LTLA via inclusion of age-adjusted expected deaths (E) as an offset; these were calculated according to national age-specific rates (deaths per 100,000 per age band) applied to local population estimates in 10-year age bands. We accounted for population density in one of four functional forms: (i) constant/independent of population density, (ii) linear, (iii) log-linear and (iv) saturating.

For observed number of deaths (D), age-adjusted expected deaths (E) and defining x_i^{IMD}, x_i^{mino}, x_i^{KW}, x_i^{lag} and x_i^{dens}, respectively, as the deprivation score, % minority population, % key workers, lag in weeks behind first death nationally and population density of LTLA this yields the following model specification:

$$D_i \sim NegBin(\mu_i E_i, \phi)$$

$$log\left(\mu_i\right) = \beta_0 + \beta_1 x_i^{IMD} + \beta_2 x_i^{mino} + \beta_3 x_i^{KW} + \beta_4 x_i^{lag} + \beta_5 f_k\left(x_i^{dens}\right)$$

where μ_i is the rate relative to the expected, ϕ the size parameter (1/overdispersion) of the negative binomial distribution and β_j the regression coefficients. The functional form of population density is defined as

$$f_k\left(x\right) = \begin{cases} 0, & k = A \\ x, & k = B \\ log(x), & k = C \\ sat(x, \theta), & k = D \end{cases}$$

where

$$sat(x, \theta) = \frac{2\theta x}{1+2\theta x+\sqrt{(1+4\theta x)}}$$

Check for updates