

# Analysis of temporal trends in potential COVID-19 cases reported through NHS Pathways

Quentin J. Leclerc      Emily S. Nightingale      Samuel Abbott  
CMMID COVID-19 Working Group      Thibaut Jombart

Thursday 14 May 2020

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Methods</b>	<b>2</b>
2.1	Data extraction . . . . .	2
2.2	Temporal analysis . . . . .	2
2.3	Correlation with reported deaths . . . . .	2
<b>3</b>	<b>Results</b>	<b>2</b>
<b>4</b>	<b>References</b>	<b>6</b>

## 1 Introduction

NHS Pathways is a triage system for public calls and online reports for medical care [5]. This system is currently being used throughout England to assist individuals reporting potential COVID-19 symptoms. Since the 18th of March 2020, data on daily phone calls and completed online assessments which have received a potential COVID-19 final disposition are openly available. These assessments are either completed via calls to 111 and 999 (which are respectively for non-urgent, and urgent medical problems), or through 111-online self-completed reports. The fraction of assessments corresponding to actual COVID-19 cases is unknown, but in the absence of wide-scale testing, the NHS Pathways dataset may be one of the best available proxies for COVID-19 incidence in the community. While prone to self-reporting biases, it is likely to better reflect milder cases and be less biased by different severity profiles than hospital admission data, which by definition reflect the most acute cases.

Here, we analyse NHS Pathways data until 12/05/2020 to assess the temporal dynamics of COVID-19 in England. Specifically, we investigated potential changes in the growth rate of the epidemic over time, and compared the observed patterns across NHS regions. We derived time-varying estimates of the growth rates, halving time and effective reproduction numbers for the different regions. We also assessed the potential correlation between NHS Pathways with COVID-19 daily deaths in England, to gain an initial understanding of its possible value within an early detection system.

## 2 Methods

### 2.1 Data extraction

We extracted the NHS Pathways data up to 2020-05-12 through the NHS Digital website [REF], where they are updated daily, every weekday. The number of reports are stratified by Clinical Commissioning Group (CCG), gender and age group (0-18, 19-69 and 70-120 years old) of the patients. We mapped the CCGs to their corresponding NHS regions using publicly available CCG data [REF], and used this geographic resolution for our analysis. All dates indicated refer to the date of reporting.

### 2.2 Temporal analysis

Total numbers of reports (including all three data sources: 111 calls, 999 calls, and 111-online reports) were modelled using quasi-Poisson generalised linear models (GLM) with log links, to account for exponential trends as well as over-dispersion of the data [8]. Predictors included time (in days since the first data point (18/03/2020)) with interaction terms for varying slopes and intercepts between NHS regions, and day of the week (weekend, monday, or rest of the week) to account for potential differences in reporting over the weekend and at the start of the week. Growth rates ( $r$ ) for each NHS region and their 95% confidence intervals were directly deduced from the corresponding coefficients. All models were fitted using maximum-likelihood.

To assess potential changes of the growth rate over time, analyses were performed over rolling windows of 14 days from the earliest available date (Wednesday 18 March 2020) to the latest available one (Tuesday 12 May 2020) (see Figure S1 in Supplementary Material for a sensitivity analysis). Growth rates and associated confidence intervals were calculated for each time window. Whenever the upper bound of  $r$  was negative, corresponding halving times were calculated as  $\log(0.5)/r$ . Growth rates were converted to effective reproduction numbers  $R_e$  using the approach described in Wallinga and Lipsitch [9] and implemented in the *epitrix* package [10], with a serial interval modelled as a gamma distribution with mean 4.7 days and standard deviation 2.9 days [11] (see Figure S2 in Supplementary Material for further details on the choice of distribution).

### 2.3 Correlation with reported deaths

To test the validity of the NHS Pathways dataset as an early detection system, we compared daily total counts of reports (including all three data sources: 111 calls, 999 calls, and 111-online reports) with publicly available NHS data on COVID-19 daily deaths [12]. This dataset includes daily counts of COVID-19 deaths in hospitals in England NHS regions. All dates refer to the date of death. However, the data are subject to bias from reporting delays, with more recent counts excluding a proportion of deaths which have not yet been reported. To account for this, we excluded data from the last 3 weeks from this analysis.

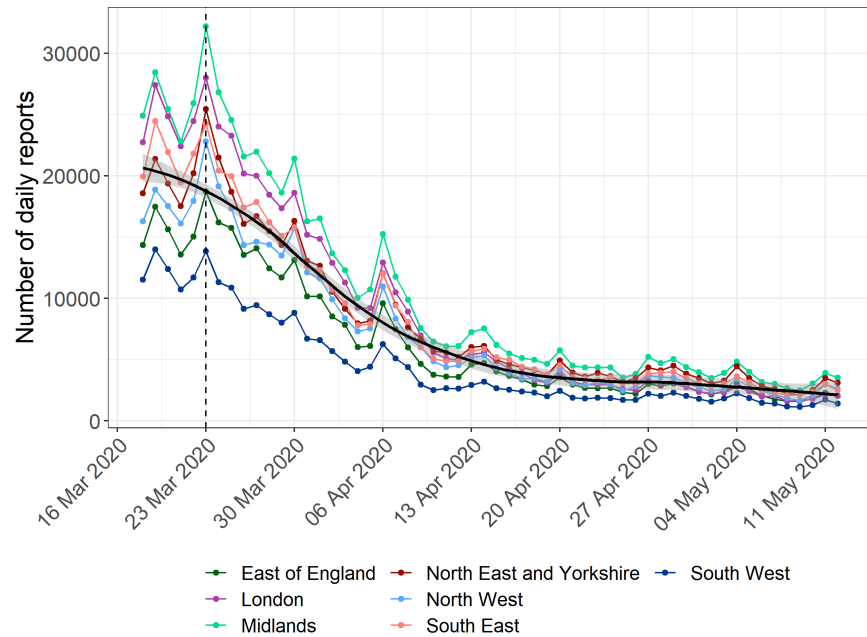
We calculated Pearson's correlation between the daily time series of deaths and NHS Pathways reports, lagging the reports from zero to thirty days. Approximate 95% confidence intervals for each correlation estimate were calculated by bootstrapping with 1,000 replicates. From this we identified an optimal lag at which the reports correlate most strongly with subsequent deaths. We then further evaluated the potential of NHS Pathways reports lagged at this value as a predictor, assuming a quasi-Poisson distribution for daily deaths.

All analyses were performed using the R software [13], and the code is publicly available from [https://github.com/qleclerc/nhs\\_pathways\\_report](https://github.com/qleclerc/nhs_pathways_report) and distributed under the MIT license.

## 3 Results

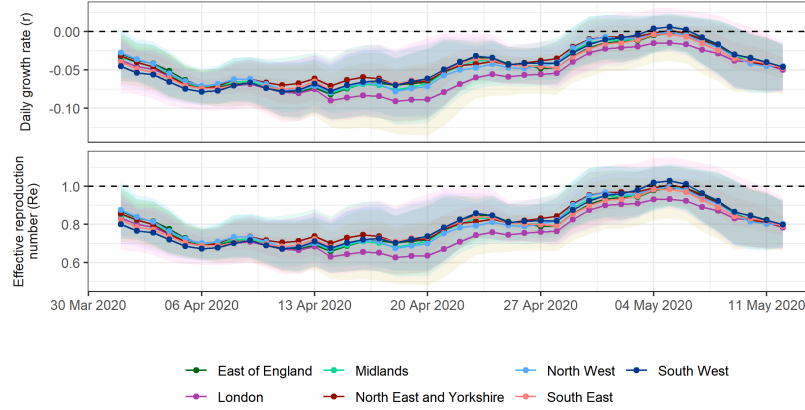
Overall reports of potential COVID-19 cases through NHS Pathways have been clearly decreasing in all NHS regions since 18/03/2020 and until approximately 13/04/2020, after which the trend seems to plateau

(Figure 1). Weekly spikes were consistently observed for all NHS regions, with increased reports on Mondays likely reflecting less reporting over weekends. The NHS COVID-19 daily deaths dataset shows daily deaths increased until approximately 10/04/2020, and have been decreasing since.



**Figure 1: Daily potential COVID-19 cases reported through NHS Pathways and reported COVID-19-related deaths, by NHS region.** Pathways data include calls to 111 and 999, as well as 111-online reports. Dates correspond to the date of case report and death report, respectively, with x-axis labels corresponding to Mondays. The solid black line and grey ribbon correspond to a lowess smoother and its 95% confidence interval. The start of the lockdown in England (23/03/2020) and date at which death data were truncated to avoid bias from reporting delay (20/04/2020) are highlighted by vertical lines.

Analyses over sliding time windows show that daily growth rates have likely been changing substantially during this period. Results show a marked decrease in growth rates ( $r$ ) and effective reproduction numbers ( $R_e$ ) until the 6th April 2020, after which these numbers remained low in all NHS regions for a period of about two weeks (Figure 2). The lowest  $r$  estimated was for Saturday 18 April 2020 in the London region at -0.08 (95% CI: -0.10 - -0.06), corresponding to a halving time of 8.45 days (95% CI: 6.77 - 11.20) and an  $R_e$  of 0.66 (95% CI: 0.59 - 0.73). Similar trends were observed across all NHS regions, with the exception of London which showed lower  $r$  (and  $R_e$ ) after the 13th of April.



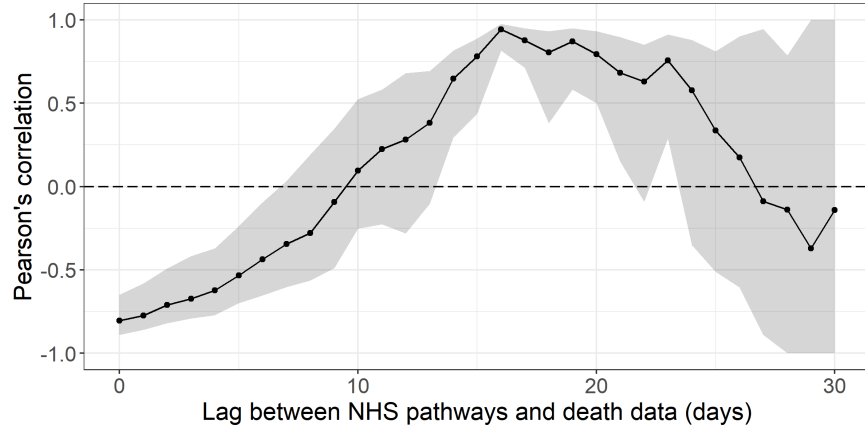
**Figure 2: Estimates of daily growth rates ( $r$ ) and effective reproduction numbers ( $Re$ ) for potential COVID-19 cases reported through NHS Pathways.** Dotted lines indicate the central estimate, and ribbons their 95% confidence intervals. Estimates are indicated at the end of the time window used for estimation, so that values of  $r$  and  $Re$  provided on a given day correspond to the 2 weeks leading up to that day.

Confidence intervals suggest  $r$  values remained lower than 0 (and  $Re$  lower than 1) in all regions until the 15th of April, consistent with a decrease in COVID-19 related reports. After this, values of  $r$  and  $Re$  seem to have gradually increased, to a point where there was no longer any strong evidence of a decrease in COVID-19 reports in any NHS region as of the 4th May 2020, as the 95% CIs of all growth rates include 0. The most recent estimate of  $r$  averaged over all NHS regions is -0.05 (95% CI: -0.08 - -0.02), corresponding to an  $Re$  of 0.79 (95% CI: 0.68 - 0.91).

The strongest correlation between NHS Pathways reports and deaths was obtained with a lag of 16 days (Pearson's correlation = 0.9; 95% CI: 0.6 - 0.99). Figure 3 illustrates the observed trend in correlation across all tested lags. Estimates become increasingly unstable for lags above 23 days as the number of points within the overlapping time window becomes small ( $n = 3$  at 30 days lag). There is however a clear, initially upward trend and subsequent plateau between 16 and 19 days, after which the strength of correlation appears to decrease. Further analysis suggests that this correlation could potentially also be seen by NHS regions separately (see Figure S6 Supplementary Material for further details).

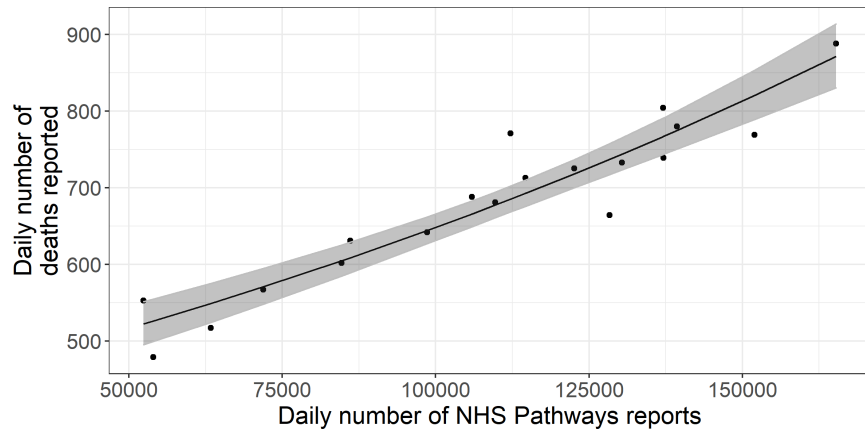
Fitting a quasi-Poisson GLM, we found that over 85.8% of the deviance in daily reported deaths could potentially be explained by NHS Pathways reports 16 days prior, with an average of 1.87 (bootstrap 95% CI: 1.66 - 2.01) additional deaths for every 1,000 potential COVID-19 cases reported in NHS Pathways 16 days before (intercept = 412, 95% CI: 376 - 452; % increase per 1000 notifications = 0.45, 95% CI: 0.37 - 0.53) (Figure 4).

Figure 3



**Figure 3: Pearson's correlation between deaths and potential COVID-19 cases reported through NHS Pathways, lagged between 0 and 30 days. 95% confidence intervals are calculated by bootstrapping with 1,000 replicates..**

Figure 4



**Figure 4: Daily total COVID-19 deaths reported in England between 03/04 and 21/04, against the number of potential COVID-19 cases reported through NHS Pathways with a lag of 16 days (between 18/03 and 05/04). The black line and grey ribbon correspond to predictions from the regression model and associated 95% confidence intervals. The coefficient of determination indicates that 85.8% of the deviance in reported deaths is linearly explained by NHS Pathways reports.**

## 4 References