Intuitions and counterintuitions of surveillance testing accuracy during an epidemic

Citation

Permanent link

Terms of Use

Share Your Story

Intuitions and counterintuitions of surveillance testing accuracy during an epidemic

James A. Hay*[1], Joel Hellewell[2], Xueting Qiu[1]

*_jhay@hsph.harvard.edu_

1. Center for Communicable Disease Dynamics, Department of Epidemiology, Harvard University T H Chan School of Public Health, Boston, Massachusetts, USA.

2. Centre for Mathematical Modelling of Infectious Diseases, London School of Hygiene & Tropical Medicine, UK

Main text

From 8 March 2021, students in English secondary schools began routine asymptomatic testing for SARS-CoV-2 infection using rapid lateral flow tests (LFTs). Students were tested at school three times during the first two weeks after their return and subsequently given two tests per week to use at home [1]. The strategy aims to identify and prompt the isolation of infected students, keeping SARS-CoV-2 transmission within schools minimal.

Two primary concerns have been raised over the use of LFTs in this context. First, the observed test positivity rate within schools is lower than SARS-CoV-2 prevalence estimates from cross-sectional surveys such as those run by the Office for National Statistics (ONS) or REACT [2–4]. This raises the question of whether many infections within schools are being missed due to poor LFT sensitivity. Second, as the prevalence of SARS-CoV-2 in school aged children continues to decline, the rate of false positive results could lead to considerable unnecessary isolation for students and their close contacts.

Here, we provide a more nuanced view of test positivity rates and isolation using a simple model of repeated LFT testing that accounts for 1) declining SARS-CoV-2 prevalence and incidence, 2) isolation of detected true positives from the tested population, and 3) test sensitivity as it varies over the course of infection. The model is available to test under alternative assumptions and parameters in the accompanying online tool.

Systematically testing all individuals in a population with any test, including qPCR, underestimates true prevalence without adjustment for imperfect sensitivity (Figure 1A). Furthermore, test positivity observed under a regular testing program will necessarily be lower than the prevalence observed in random cross-sectional LFT/qPCR surveys. Initially, routine LFT screening is expected to return the same percentage positive as in a random cross-sectional LFT survey, but then drops systematically lower in the following weeks. This arises because the number of true positives left to be detected may be continually depleted through isolation of existing infections, counteracting the replenishment of true positives via incident infections; increased test frequency can lead to depletion of infections faster than they can be replenished. As a result, the percentage of positive tests under a regular screening strategy gives a metric that sits somewhere between incidence and prevalence (the black dots sit between the grey bars and red line in Figure 1A) — a low observed test positivity compared to cross-sectional surveillance is therefore expected under regular screening, and should not be used to evaluate the success of the strategy.

Understanding how true positives are detected and isolated also requires a detailed view of test sensitivity. LFT sensitivity has been reported at 50.1% in the context of school testing, reflecting the overall positive percentage agreement (PPA) with qPCR for infected individuals tested at an unknown day post infection [5]. However, test sensitivity is not a single static value, but depends on the quantity of viral material within the host which increases and then decreases over the course of an infection (Supplementary Materials Figure S1). When individuals are tested prospectively, infections that are missed due to low viral loads on the first test day are likely to have higher viral loads and corresponding increased sensitivity at the next test. Thus, very few infections escape detection when individuals are routinely tested every three days, compounded by the fact that at least one test day will likely coincide with the period of high viral load (Supplementary Materials Figure S2) [6].

The removal of true infections from the tested population also has implications for the proportion, but not absolute number, of false positive test results. The probability of a positive LFT result being a true positive is given by the positive predictive value (PPV), which depends on the specificity of LFTs and the prevalence of SARS-CoV-2 in the population. Both LFTs and qPCR are highly specific: LFT specificity has been cited at 99.6%, but given the continued decline in percentage of tests returning positive, specificity may be even higher [7,8]. One study estimated LFT specificity to be 99.97% after accounting for imperfect qPCR sensitivity [5]. However, because the pool of true positives declines faster than population prevalence through the isolation of detected infections, the positive predictive value of the regular screening strategy decreases faster than in a random surveillance strategy, though the negative predictive value is systematically higher than under random surveillance (Figure 1B; Supplementary Material Figure S3). This low PPV must be weighed against the absolute number of false positive LFTs, which is constant and low depending on the test specificity, leading to a relatively small number of individuals incorrectly isolating at a given time (Figure 1C&D). However, we note that the number isolating is considerably higher when a large number of contacts are also isolated following detection of a single positive.

Taken together, these factors demonstrate the challenges of naively interpreting true and false positivity rates when repeatedly testing and isolating individuals in light of declining SARS-CoV-2 prevalence and infection-time-varying test sensitivity. Although the overall ability of sustained mass screening strategies, as carried out in English schools, to detect and isolate infected individuals is very high, the costs, benefits and risk of false negatives arising from confirmatory and test-to-release antigen or molecular testing will require ongoing evaluation.
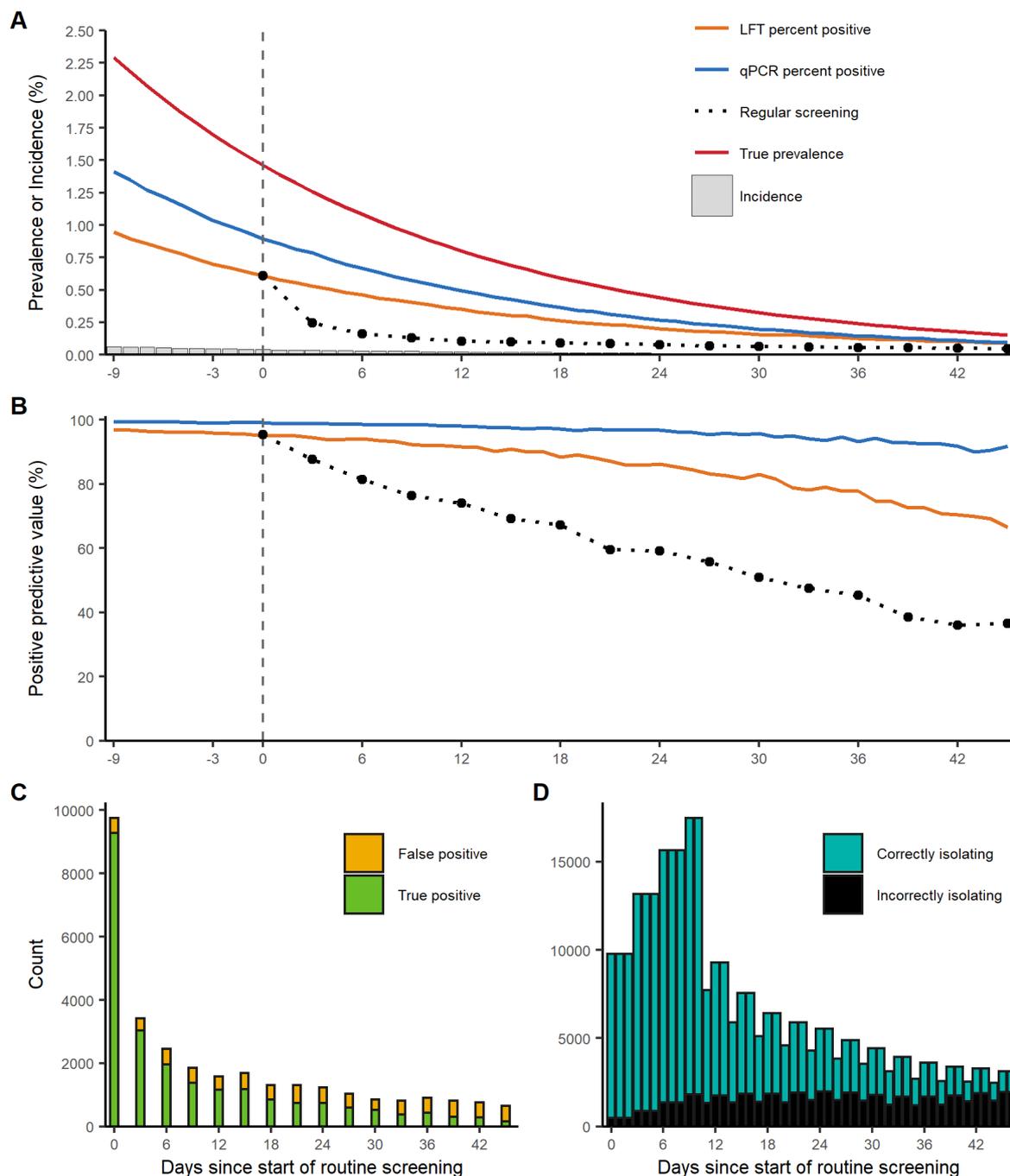
Figure 1. Test positivity and positive predictive value decline alongside prevalence, but the number of false positives and subsequent isolations are stable. Simulation uses a population of 1.5 million individuals to match the number of secondary school pupils tested in their first week of return to school, with prevalence and test characteristics chosen to broadly reflect the situation in England [9]. True prevalence ranged from 2.29% on the first day of the strategy to 0.153% on the last day. LFT specificity is set to 99.97% and qPCR specificity to 99.99%. Full simulation assumptions and parameters are shown in *Supplementary Material 1: Methods*. The repeated screening strategy is initiated when true prevalence is at approximately 1.5% (vertical dashed lines). (A) Percentage of tests returning positive under random cross-sectional testing with LFT or qPCR and no isolation of positives, or 3-day LFT screening with isolation of positives compared to true prevalence and incidence. Note that daily incidence ranges from 6.33 to 0.405 infections per 10,000 people (Supplementary Material

Figure S4). Black dots show the observed test positivity from the 3-day LFT screening tests performed on that day. (B) Positive predictive value (PPV) of random cross-sectional testing with LFT or qPCR, or 3-day LFT screening with removal of positives. PPV ranged from 95.3% to 36.5% for the 3-day LFT screening strategy. (C) Number of positives detected on each day of the 3-day LFT screening strategy. (D) Number of individuals in isolation over time under the 3-day LFT screening strategy, stratified by whether they were detected as a true or false positive.

References

1. Coronavirus (COVID-19) asymptomatic testing in schools and colleges. [cited 25 Mar 2021]. Available: https://www.gov.uk/government/publications/coronavirus-covid-19-asymptomatic-testing-in-schools-and-colleges/coronavirus-covid-19-asymptomatic-testing-in-schools-and-colleges

2. Coronavirus (COVID-19) infection survey, UK: 19 Mar 2021 [cited 25 Mar 2021]. Available: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/coronaviruscovid19infectionsurveypilot/19march2021

3. Riley S, Wang H, Eales O, Haw D, Walters CE, Ainslie KEC, et al. REACT-1 round 9 final report: Continued but slowing decline of prevalence of SARS-CoV-2 during national lockdown in England in February 2021. medRxiv; 2021. doi:10.1101/2021.03.03.21252856

4. Weekly statistics for rapid asymptomatic testing in England: 4 March to 10 March 2021. [cited 25 Mar 2021]. Available: https://www.gov.uk/government/publications/weekly-statistics-for-nhs-test-and-trace-england-4-march-to-10-march-2021/weekly-statistics-for-rapid-asymptomatic-testing-in-england-4-march-to-10-march-2021

5. Lateral flow device specificity in phase 4 (post-marketing) surveillance. [cited 25 Mar 2021]. Available: https://www.gov.uk/government/publications/lateral-flow-device-specificity-in-phase-4-post-marketing-surveillance/lateral-flow-device-specificity-in-phase-4-post-marketing-surveillance

6. Smith RL, Gibson LL, Martinez PP, Ke R, Mirza A, Conte M, et al. Longitudinal assessment of diagnostic test performance over the course of acute SARS-CoV-2 infection. medRxiv; 2021. doi:10.1101/2021.03.19.21253964

7. LFD mass testing in English schools - additional evidence of high test specificity. 4 Mar 2021 [cited 25 Mar 2021]. Available: https://cmmid.github.io/topics/covid19/mass-testing-schools.html

8. Dinnes J, Deeks JJ, Berhane S, Taylor M, Adriano A, Davenport C, et al. Rapid, point-of-care antigen and molecular-based tests for diagnosis of SARS-CoV-2 infection. Cochrane Database Syst Rev. 2021 [cited 25 Mar 2021]. doi:10.1002/14651858.CD013705.pub2

9. Torjesen I. What do we know about lateral flow tests and mass testing in schools? BMJ. 2021;372: n706. doi:10.1136/bmj.n706

10. Kucirka LM, Lauer SA, Laeyendecker O, Boon D, Lessler J. Variation in False-Negative Rate of Reverse Transcriptase Polymerase Chain Reaction-Based SARS-CoV-2 Tests by Time Since Exposure. Ann Intern Med. 2020;173: 262–267. doi:10.7326/M20-1495

Acknowledgements
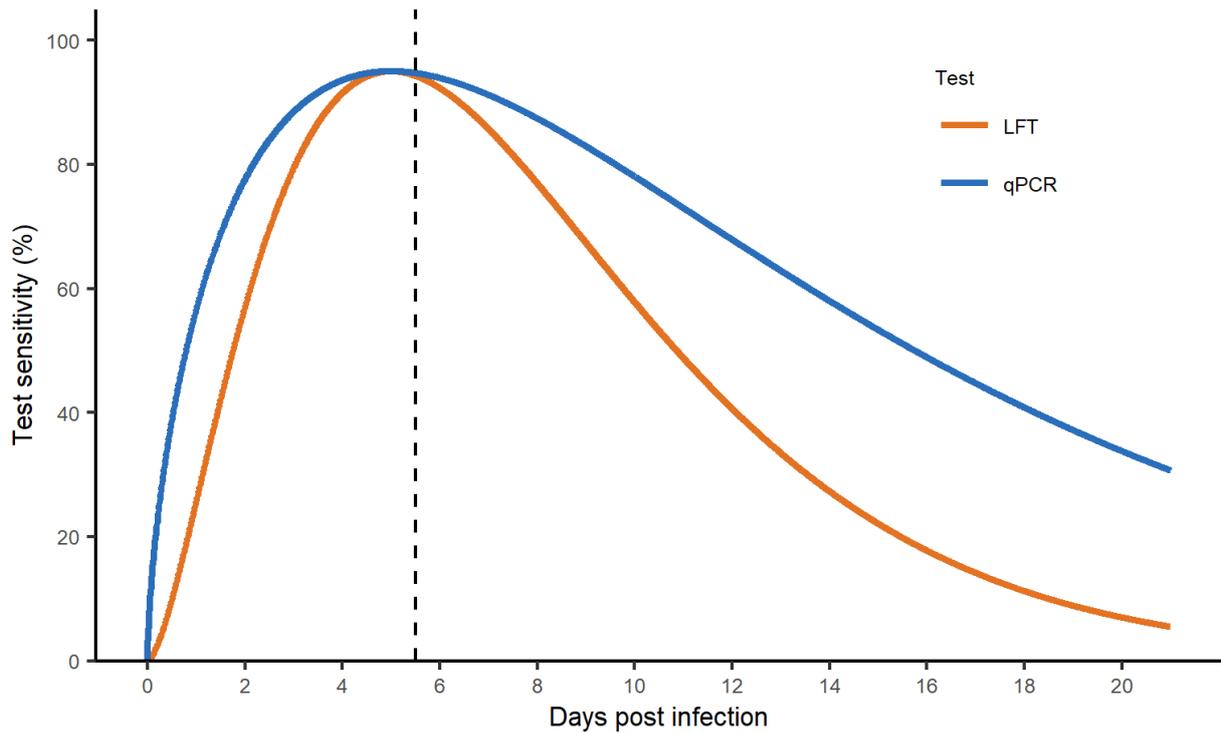
Supplementary Figure 1. Assumed test sensitivity over time post infection. That is, the probability of returning a positive test result given a sample is taken on a given day after infection. The dashed vertical line shows day 5.5 post infection as a benchmark for the typical time of symptom onset. Under these curves, the mean LFT and qPCR sensitivity over the 21 day period are 44% and 63% respectively. The mean sensitivities over days 2-12 post infection are 71% and 83% for LFT and qPCR respectively.
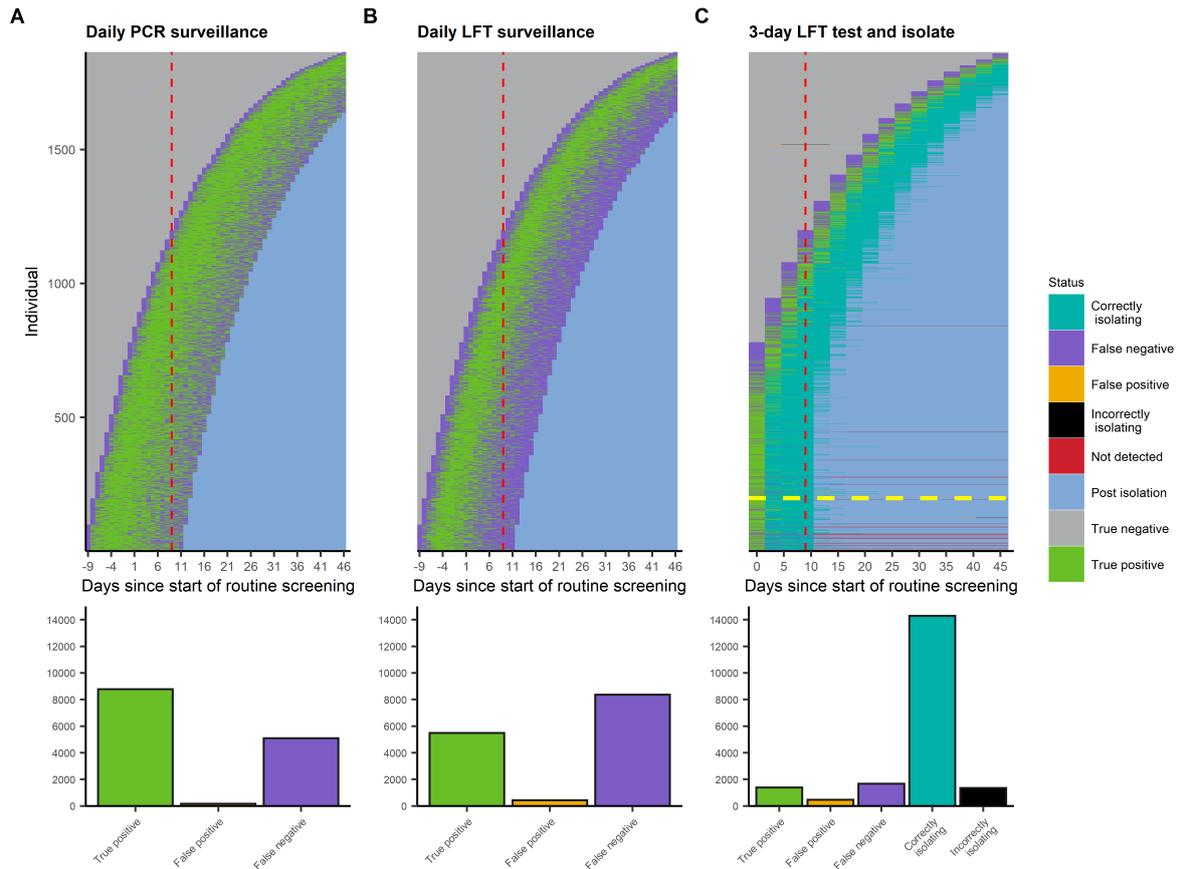
Supplementary Figure 2. (Top) Individual-level test results under complete surveillance with no isolation using (A) qPCR, (B) LFT or (C) testing all individuals every three days with an LFT followed by 10-day isolation of positives. Each row shows one individual, ordered by infection date. Plot is subsampled to show 10% of the full population. Each row is an individual, coloured to show their infection status on each day of the simulation. Note that only individuals who are infected within the focal period are plotted, and all uninfected individuals would be plotted as predominantly grey bars with infrequent false positives. The vertical red dashed line shows a single observation time for the bar charts shown in the bottom row. In (C), the horizontal yellow line gives a boundary of old versus new infections: individuals below this line were infected in the nine days prior to the start of the regular screening strategy, and thus represent older infections. Most infections who are not detected before they recover occurred prior to the start of routine screening; there are very few individuals infected after the start of the screening strategy who escaped detection (see *Supplementary Material 1: Methods* for further results).
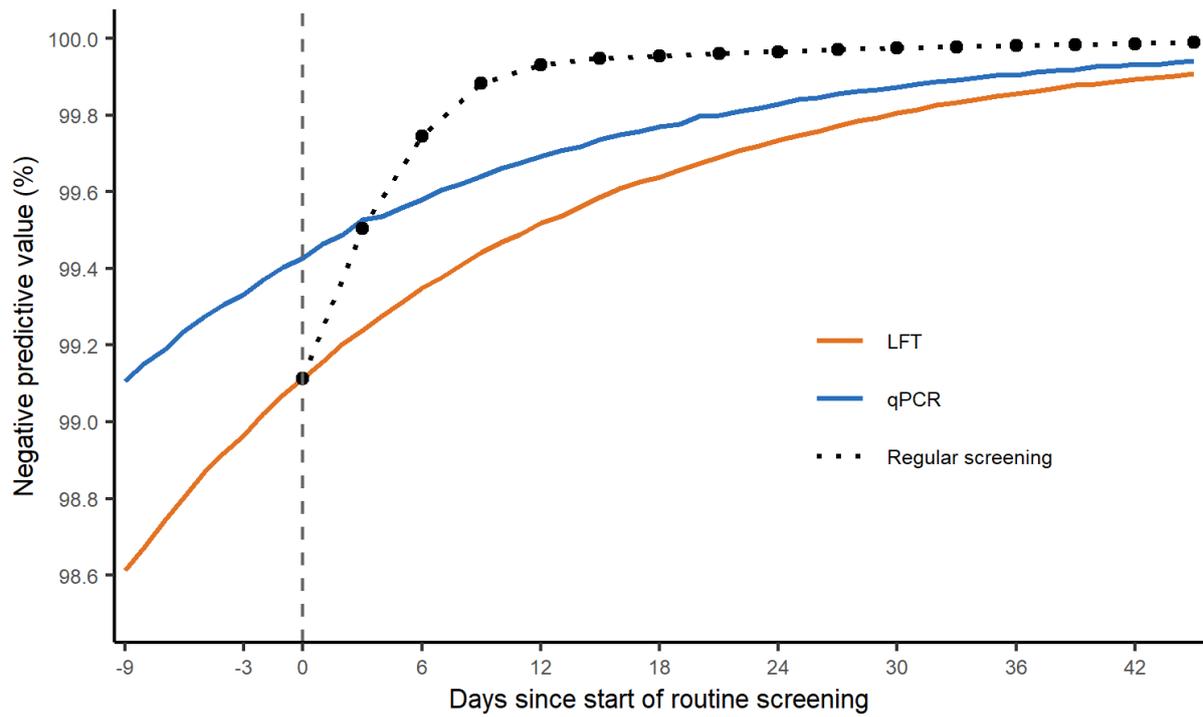
Supplementary Figure 3. Negative predictive value (NPV) of daily testing of all individuals with qPCR or LFT versus LFT testing and isolating positive individuals every three days. Dashed vertical line shows the time when the regular screening strategy is initiated. NPV ranged from 99.1% to 100% for the 3-day LFT screening strategy.
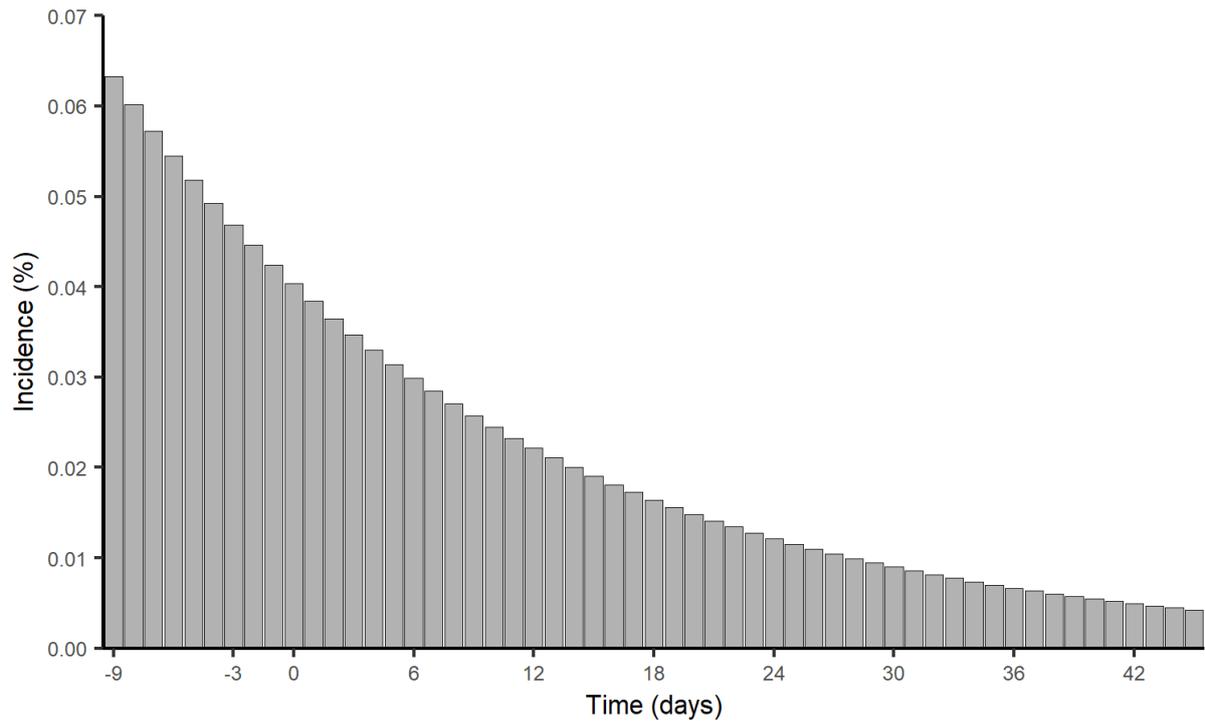
Supplementary Figure 4. Simulated daily incidence as a percentage of the entire population matching parameter assumptions in main text Figure 1.

Supplementary Materials 1: Methods

We used a simple discrete time model to simulate infection and test states in a population over a short period of exponentially growing or declining infection incidence. The model generates infections, prevalence, overall LFT and qPCR test positivity, and isolation status of detected individuals under a strategy of repeated LFT testing. In the main text, model parameters were chosen to reflect exponentially declining incidence with prevalence and population size set to represent the number of screened secondary school aged children in England. We describe the three key components of the model:

1. The incidence and prevalence of infections.
2. Characteristics of test sensitivity and specificity.
3. Testing strategy and subsequent isolation.

Incidence of new infections

We assume that new infections, $I(t)$, arise each day under an exponential growth curve with a specified growth rate and cumulative incidence:

$$I(t) = N\psi \frac{e^{\beta t}}{\sum_{i=1}^{t_{max}} e^{\beta i}}$$

Where N is the population size (set to 1.5 million); t_{max} is the duration of the simulation period (set to 97 days, with the first 42 days discarded as burn-in; note that the simulation runs for 55 days but the 3-day LFT screening strategy only begins 9 days after burn-in); β is the exponential growth rate (set to -0.05); and ψ is the cumulative incidence between $t = 1$ and $t = t_{max}$ (set to 0.1). We assume that each infected individual remains truly positive (i.e., contributes towards true prevalence) for γ days, set to 21 days in the main text.

Test characteristics

Each infected individual is tested daily with qPCR and LFT to find the number of tests returning positive under daily, complete cross-sectional surveillance. For infected individuals, each day post infection, the result of a test is simulated as a Bernoulli trial with probability of returning a true positive equal to $P_{LFT}(x)$ or $P_{qPCR}(x)$, representing the probability of returning a positive LFT or qPCR result conditional on having been infected x days prior. $P_{LFT}(x)$ and $P_{qPCR}(x)$ were defined by scaled gamma distributions:

$$f(x) = \frac{1}{\Gamma(k)\theta^k} x^{k-1} e^{\frac{-x}{\theta}}$$

$$P_{LFT/qPCR}(x) = min(\frac{f(x)}{\lambda * max(f(x))}, p_{max})$$

Where $f(x)$ is the probability density function of the gamma distribution with shape parameter k and scale parameter θ; λ is an arbitrary scaling parameter that sets the maximum value for $f(x)$ at the mode; and p_{max} is the maximum sensitivity. Note that λ can take values above 1 to generate a plateau of maximally sensitive test days, truncated by p_{max}.

For uninfected individuals, false positive results are simulated from Bernoulli trials with probability equal to one minus the specificity, $1 - sp_{LFT}$ or $1 - sp_{qPCR}$. For the main text results, we assumed that $sp_{LFT} = 99.97\%$ and $sp_{qPCR} = 99.99\%$.

The gamma distribution was chosen to generate a flexible parametric curve representing how test sensitivity increases then decreases over the course of an infection [6]. If the LFT is taken around the time of peak viral load then sensitivity will be high, whereas if the sample is taken three weeks after infection onset then sensitivity will be far lower [10]. qPCR sensitivity demonstrates the same time dependence as LFTs but with higher sensitivity early and late in the infection due to its very low analytical limit of detection. qPCR also returns a positive test result for individuals who are no longer infectious but may still shed viral RNA for a long time after recovery.

Surveillance strategies

As a benchmark, we simulated test results assuming that all individuals are tested each day of the simulation with qPCR or LFT, with no isolation of detected positives. We compared this to the results of testing all individuals with an LFT every ρ days, where $\rho = 3$ in the main text. In this strategy, individuals who are detected as positive enter isolation for 10 days following detection, and can not be tested again. True positives in isolation are categorised as "correctly isolating", whereas false positives in isolation are categorised as "incorrectly isolating". True infections who are not detected at any point in their infection are categorised as "not detected".

Undetected infections

Under the simulation, there were a total of 150,000 infections. The vast majority of these occurred in the discarded burn-in phase of the simulation: 7050 infections arose in the nine days prior to initiation of routine screening, and 11,200 thereafter. 190 of the infections from the nine days prior to routine screening and 270 infections thereafter were never detected. However, 250/270 of these latter infections occurred in the final week of the simulation, and may have been captured if routine screening continued. Only 20 of the 10600 infections which arose within the screening period and more than one week prior to the end of the simulation were missed.