Deploying wearable sensors for pandemic mitigation

Nathan Duarte B.A.Sc.[1], Rahul K. Arora B.H.Sc.[2,3], Graham Bennett B.A.[4], Meng Wang Ph.D.[5], Michael P. Snyder Ph.D.[5], Jeremy R. Cooperstock Ph.D.[1], Caroline E. Wagner Ph.D.[6]*

1. Department of Electrical and Computer Engineering, Faculty of Engineering, McGill University, Montreal, Canada.
2. Department of Community Health Sciences, University of Calgary, Calgary, Canada.
3. Institute of Biomedical Engineering, University of Oxford, Oxford, United Kingdom.
4. Department of Economics, Faculty of Arts, McGill University, Montreal, Canada.
5. Department of Genetics, Stanford University School of Medicine, Stanford University, California, United States of America.
6. Department of Bioengineering, Faculty of Engineering, McGill University, Montreal, Canada.

*** Corresponding Author**: Caroline E. Wagner (caroline.wagner@mcgill.ca; +1-438-399-7911)

ABSTRACT

Wearable sensors can continuously and passively detect potential respiratory infections, before or absent symptoms. However, the population-level impact of deploying these devices during pandemics is unclear. We built a compartmental model of Canada's second COVID-19 wave and simulated wearable sensor deployment scenarios, systematically varying detection algorithm accuracy, uptake, and adherence. With current detection algorithms and 4% uptake, we found that deploying wearable sensors could have averted 9% of second wave SARS-CoV-2 infections, though 29% of this reduction is attributed to incorrectly quarantining uninfected device users. Improving detection specificity and offering confirmatory rapid tests each minimized incorrect quarantines and associated costs. With a sufficiently low false positive rate, increasing uptake and adherence became effective strategies for scaling averted infections. We concluded that wearable sensor deployment can meaningfully contribute to pandemic mitigation; in the case of COVID-19, technology improvements or supporting measures are required to reduce social and economic costs to acceptable levels.

Key Words: wearable sensors, public health, asymptomatic, presymptomatic, SARS-CoV-2, COVID-19

INTRODUCTION

Infectious disease outbreaks can have devastating health and economic consequences. Effective public health strategies are crucial for limiting transmission and minimizing these harms. One approach to controlling viral spread during pandemics – a "Find, Test, Trace, Isolate" (FTTI) strategy – relies on identifying and quarantining infectious individuals[1]. However, the COVID-19 pandemic has demonstrated that FTTI systems reliant on lab-based tests are often limited by missed hidden infection chains, which result from presymptomatic and asymptomatic transmission, and slow test result turnaround times[2,3]. Digital contact tracing and frequent use of rapid tests have the potential to fill these gaps, but both approaches have faced implementation barriers: inadequate participation levels, concerns around privacy and feasibility, and limited test availability[4–6].

Wearable sensors have already been established as tools to detect deviations from users' physiological baselines[7]. Recent findings suggest that wearable sensors may also be able to detect infections caused by respiratory pathogens, including SARS-CoV-2, before or absent symptoms[8–10]. Alavi *et al*, for example, developed an algorithm that analyzes patterns in smartwatch-captured overnight resting heart rate and provides real-time alerts of potential presymptomatic and asymptomatic SARS-CoV-2 infection[10]. If such algorithms were widely deployed, wearable sensors could be promising tools for outbreak mitigation, enabling FTTI systems to more rapidly identify (and subsequently isolate) infectious individuals, particularly those without symptoms. Wearable sensors would also offer the unique benefit of passive monitoring, which minimizes required user engagement, and could operate in privacy-preserving fashion because sensor data would not need to be shared with a centralized database. With these potential benefits in mind, several studies have focused on developing wearable sensor-based infectious disease detection algorithms; however, to the best of our knowledge, the real-world impact of deploying these devices for pandemic mitigation has yet to be explored.

In this study, we assessed how wearable sensors can help reduce the burden of infection during a pandemic with the overarching goal of guiding future research investment and policy. To do so, we used COVID-19 as an example and explored counterfactual scenarios in which these devices were deployed to combat Canada's second wave. We built a compartmental epidemiological model in which wearable devices notify users of potential infection and prompt users to seek a confirmatory lab-based test, self-isolating while waiting for the result. We aimed to (1) assess the baseline impact of deploying currently available detection algorithms, (2) investigate how detection accuracy and behavioural parameters influence this impact, and (3) explore a complementary strategy in which rapid antigen tests are used to confirm wearable-informed notifications of potential infection.

METHODS

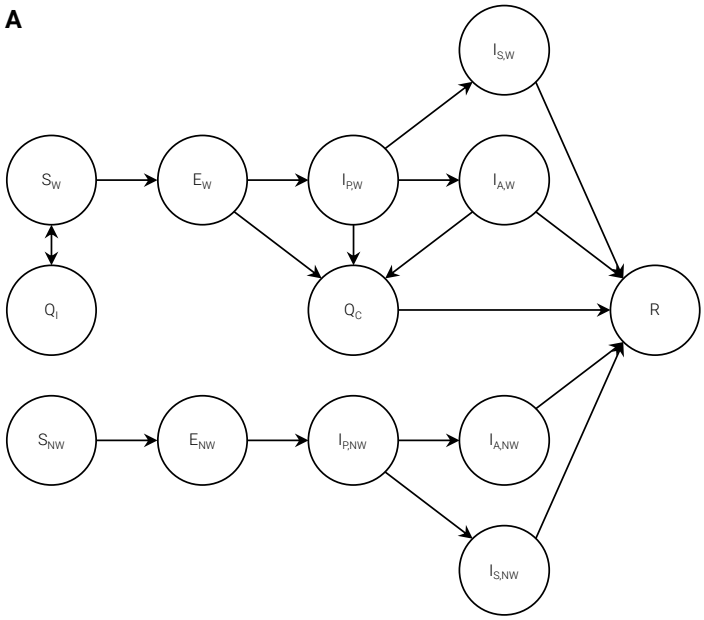
Counterfactual Scenarios

We simulated Canada's second COVID-19 wave (September 1, 2020 to February 20, 2021). This time window allowed us to capture the dynamics of wearable sensor deployment during an acute phase of the pandemic, prior to broad vaccine availability, and at a time when the technology would have been ready and deployable.

First, we considered a baseline scenario in which wearable device users can download an application with currently available detection algorithms[10]. Then, we investigated the impact of technology and behavioural parameters: detection sensitivity and specificity; uptake, defined as the proportion of the population that has downloaded the application and uses their wearable device often enough; and adherence, defined as the proportion of users who comply with recommended next steps after a positive notification. Finally, we explored a complementary intervention where wearable device users with a positive notification are offered a confirmatory rapid antigen test before they are prompted to seek a lab-based test and self-isolate.

Model description

We built a compartmental model based on a *Susceptible*, *Exposed*, *Infectious*, *Removed* (*SEIR*) framework (Figure 1). We split the *Infectious* state into three sub-states: *Presymptomatic*, *Asymptomatic*, and *Symptomatic*. All infected individuals enter the *Presymptomatic* infectious state after a latency period following exposure; some go on to develop symptoms (*Symptomatic*) while others do not (*Asymptomatic*).



B

Compartment	Symbol
Susceptible	S
Exposed	E
Infectious (presymptomatic)	I_p
Infectious (asymptomatic)	I_a
Infectious (symptomatic)	I_s
Removed	R
Quarantined (correctly)	Q_c
Quarantined (incorrectly)	Q_i

Figure 1: Compartmental model. Subscript "W" denotes a wearable user and "NW" indicates otherwise. Two key model equations are presented below; remaining parameters, equations, and assumptions are outlined in Supplementary Appendix Section 1.

To incorporate wearable sensor deployment, we stratified *Susceptible*, *Exposed*, and *Infectious* states by whether individuals are device users or not. Wearable device users can enter *Quarantined* states if they are notified of potential infection, and if they adhere to this notification by seeking a confirmatory lab-based test and self-isolating while awaiting the result. We made the simplifying assumption that users notified of potential infection either adhered to all recommended next steps or ignored the notification. We set the nominal lab-based test turnaround time to 2 days, and explored faster and slower turnaround times in Supplementary Appendix Section 5[11]. *Susceptible* users could be *Incorrectly Quarantined* due to a false positive notification and would re-enter the *Susceptible* state after receiving their lab-based test result. *Exposed* and *Infectious* device

users would be *Correctly Quarantined* and would enter the *Removed* state (a longer quarantine until recovery) after their lab-based test confirms infection.

We assumed perfect lab-based test accuracy and did not account for potential reinfections, which were likely negligible during the simulation period[12]. We did not include a pathway for *Symptomatic* device users to enter *Quarantined* states because it is likely that a large fraction of these individuals already underwent some degree of self-isolation – behaviour accounted for in the historical transmission rate (β). In some scenarios, we also included a step where compliant users take a confirmatory rapid antigen test: if positive, we assumed they then take a lab-based test, self-isolating while awaiting the result; if negative, we assumed they return to normal behaviour. We investigated the impact of antigen test sensitivity in Supplementary Appendix Section 5.

Simulation approach

To perform simulations, we first extracted the historical transmission rate (β) from the Institute for Health Metrics and Evaluation (IHME) infection model, a time series "nowcasting" model for incidence of infection in Canada (π)[13]. The IHME model estimates π by analyzing trends in confirmed cases, hospitalizations, and deaths; it validates the inferred π using seroprevalence data. We downloaded these data from the IHME website on December 7, 2021. We used modelled values for the incidence of infection because confirmed case counts are subject to incomplete case ascertainment and thereby underestimate the extent of infection[14]. We calculated β from π using Equation (1); this time series for the historical transmission rate incorporates all policy (e.g., restrictions, business closures, testing availability) and behavioural (e.g., adherence to restrictions, quarantines) measures that occurred. N represents the size of the entire population and λ represents the transmission potential in infected individuals without symptoms relative to those with symptoms.

$$\beta = \frac{\pi N}{S(\lambda I_p + \lambda I_a + I_s)} \tag{1}$$

$$\pi = \frac{\beta}{N}\Big[\lambda(aI_{p,w} + I_{p,nw} + aI_{a,w} + I_{a,nw}) + (I_{s,w} + I_{s,nw})\Big]\Big[S_w + S_{nw}\Big] \tag{2}$$

We then applied β according to Equation (2) to simulate counterfactual scenarios. Because some *Susceptible*, *Exposed*, and *Infectious* device users undergo self-isolation in the simulations, overall transmission decreases even though the transmission rate for all others remains unchanged. In Supplementary Appendix Section 5, we investigated the possibility that users not notified of potential infection act in a riskier fashion relative to historical behaviour[15]. a is a coefficient used to study this possibility and is nominally set to 1. Other model parameters, equations, and assumptions are presented in Supplementary Appendix Section 1.

We modeled asymptomatic prevalence (ρ), detection algorithm sensitivity (σ_w) and specificity (ν_w), and adherence (ψ) as beta-distributed random variables because these parameters were important sources of variance in our assessment of wearable sensors as pandemic mitigation tools. We sampled these variables and used the resulting values to generate an epidemic trajectory. We repeated this process 1,000 times, using these Monte Carlo simulations to model uncertainty in our estimates.

Outcome measures

We calculated the number of averted infections and the percent reduction in the burden of infection to quantify the health impact of wearable sensor deployment. We measured the number of days incorrectly spent in quarantine per month per device user (a consequence of false positive notifications) as the primary indicator of the strategy's social burden. To assess resource consumption, we quantified the number of additional lab-based tests (and rapid antigen tests, where applicable) required each day, on average. Finally, to provide more complete intuition around offering rapid antigen tests as a

complementary intervention, we generated a first-order estimate of net healthcare expenditures in applicable scenarios, subtracting savings from costs. We approximated savings as the costs avoided by averting hospitalizations: we multiplied averted infections, the infection hospitalization ratio, and the average cost per hospitalization. We inferred the infection hospitalization ratio by dividing the total number of COVID-19 hospitalizations by the total number of infections during the simulation period. We approximated costs as expenditures required to fund lab-based tests and rapid antigen tests. Further details on this economic analysis are provided in Supplementary Appendix Section 3. We adhered to CHEERS guidelines for health economics cost benefit analyses[16].

RESULTS

Baseline impact of wearable sensor deployment

We first investigated the baseline scenario in which detection algorithms that currently exist are made publicly available for device users to download and use (Figure 2)[10]. Upon notification of potential presymptomatic or asymptomatic infection, users are prompted to seek a confirmatory lab-based test, self-isolate while awaiting the result (nominally, for 2 days), and quarantine until recovery if positive. We used the nominal parameters in Supplementary Table 2, setting uptake, adherence, detection sensitivity, and detection specificity to 4%, 50%, 80%, and 92%, respectively.

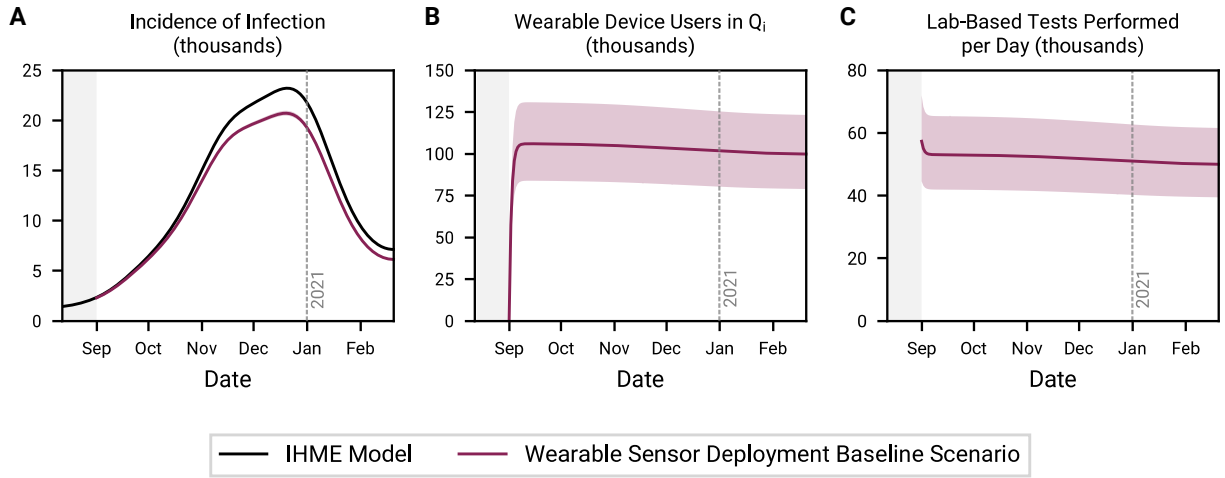


Figure 2: Baseline scenario for wearable sensor deployment. Time series depiction of the incidence of infection (A), the number of wearable device users incorrectly in quarantine (B), and the daily demand for lab-based tests (C). Uptake, adherence, detection sensitivity, and detection specificity are set to 4%, 50%, 80%, and 92%, respectively.

With 4% uptake, 219,700 (95% CI: 198,700–241,600) infections could have been averted during Canada's second COVID-19 wave – a 9.4% (95% CI: 8.5–10.3%) reduction in the burden of infection. However, the social costs were high: between ~75,000 and ~125,000 device users were incorrectly self-isolating on any given day (Figure 2b). Moreover, between ~40,000 and ~65,000 additional lab-based tests were required each day (Figure 2c), corresponding to a 51.8% (95% CI: 41.0–64.0%) increase in demand relative to historical volumes. Historically, ~101,000 lab-based tests were performed each day, on average, during the simulation time frame[13,17]. The number of individuals incorrectly in quarantine and daily demand for lab-based tests were steady over time because they depend on the number of *Susceptible* device users, adherence, and detection specificity; the gradual decrease can be attributed to the flow of users into the *Removed* state.

Tradeoff between detection algorithm sensitivity and specificity

After their initial release on technology platforms, health detection algorithms can be updated and improved as more real-world data are collected. However, it is often challenging to dramatically raise detection sensitivity and specificity at the same time. We explored the implications of this tradeoff (Figure 3), varying detection sensitivity and specificity while keeping uptake and adherence constant at 4% and 50%, respectively.

Increasing detection sensitivity increased the number of averted infections by prompting more *Infectious* users to self-isolate (Figures 3a and 3b). On the other hand, increasing specificity had a two-part effect. First, as specificity approached 100%, the number of days incorrectly spent self-isolating approached zero (Figure 3c); sensitivity had negligible impact on incorrect quarantines. Second, by virtue of decreasing the number of incorrect quarantines, increasing specificity resulted in a larger pool of *Susceptible* individuals; in turn, fewer infections were averted. However, despite this second effect, incorrect quarantines were not central to the strategy's public health impact. In the baseline scenario presented above (80% detection sensitivity, 4% uptake, and 50% adherence), a 6.7% (5.9–7.3%) reduction in the burden of infection was still achievable with

perfect detection specificity. As well, while 28.6% (95% CI: 18.3–38.2%) of averted infections could be attributed to incorrect quarantines in the baseline scenario, this proportion decreased as sensitivity improved (Supplementary Figure 4).

In theory, increasing detection sensitivity would increase demand for lab-based tests; however, this effect paled in comparison to the number of lab-based tests prompted by false positive notifications (Figure 3d). Lab-based test demand expectedly decreased as detection specificity increased.

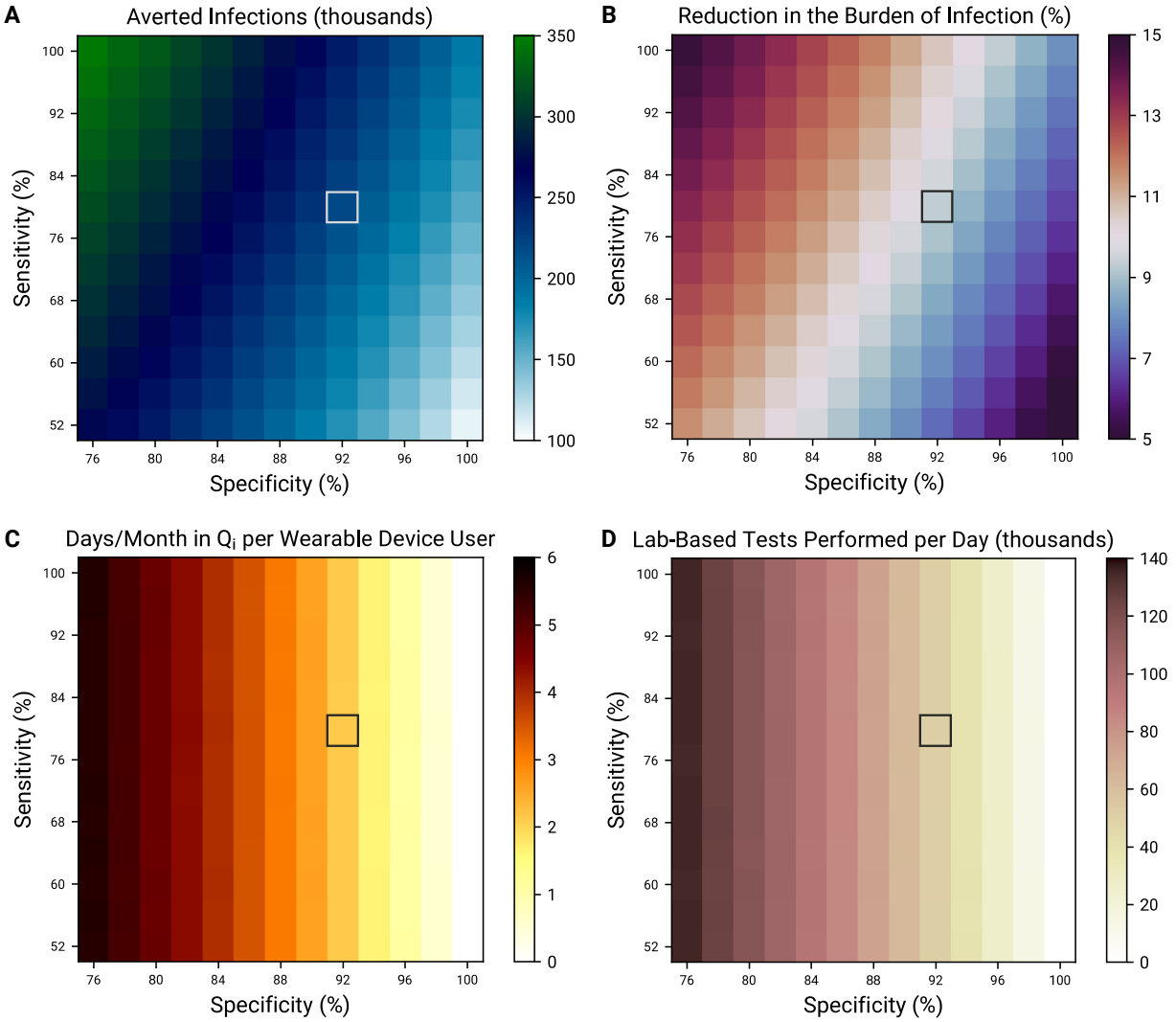


Figure 3: Tradeoff between detection sensitivity and specificity. Averted infections (A), reduction in the burden of infection (B), days incorrectly spent in quarantine per month per device user (C), and average daily demand for lab-based tests (D), all over the entire simulation period, as a function of detection sensitivity and specificity. Grey boxes denote nominal sensitivity (80%) and specificity (92%).

Impact of increasing uptake

Ensuring that public health measures reach sufficient levels of uptake has been a continued challenge through the COVID-19 pandemic. Digital contact tracing and vaccination efforts have demonstrated that well-constructed policies – for example, incentivizing participation – can improve uptake of measures[18,19]. Here, we explored the role of uptake to provide relevant context for the design of wearable sensor deployment policies (Figure 4; Supplementary Figure 1). We estimated that uptake would fall between 0.5% and 7.5% (Supplementary Tables 2-4) at baseline but chose to present outcomes at all levels of uptake to illustrate emergent phenomena. We also explored multiple technology scenarios, setting "high" detection

sensitivity and specificity at 96.0% and 98.4%, respectively; we based these increases on the respective goals of capturing 20% more infections and reducing the false positive rate by 80% relative to nominal values. We kept adherence constant at 50%.

In all technology scenarios, increasing uptake averted more infections, though with diminishing returns (Figures 4a and 4b). As expected, improving detection specificity resulted in fewer averted infections when uptake is held constant; this effect was most pronounced at ~30% to ~60% uptake. The number of days incorrectly spent in quarantine per month per device user remained constant as a function of uptake but decreased from ~2.15 to ~0.45 when detection specificity was increased (Figure 4c). This ~80% decrease was consistent with our definition of "high specificity" demonstrating that detection specificity directly influences the burden of incorrect quarantines on device users. The average daily demand for lab-based tests scaled linearly with uptake, but at a slower rate with improved detection specificity (Figure 4d).

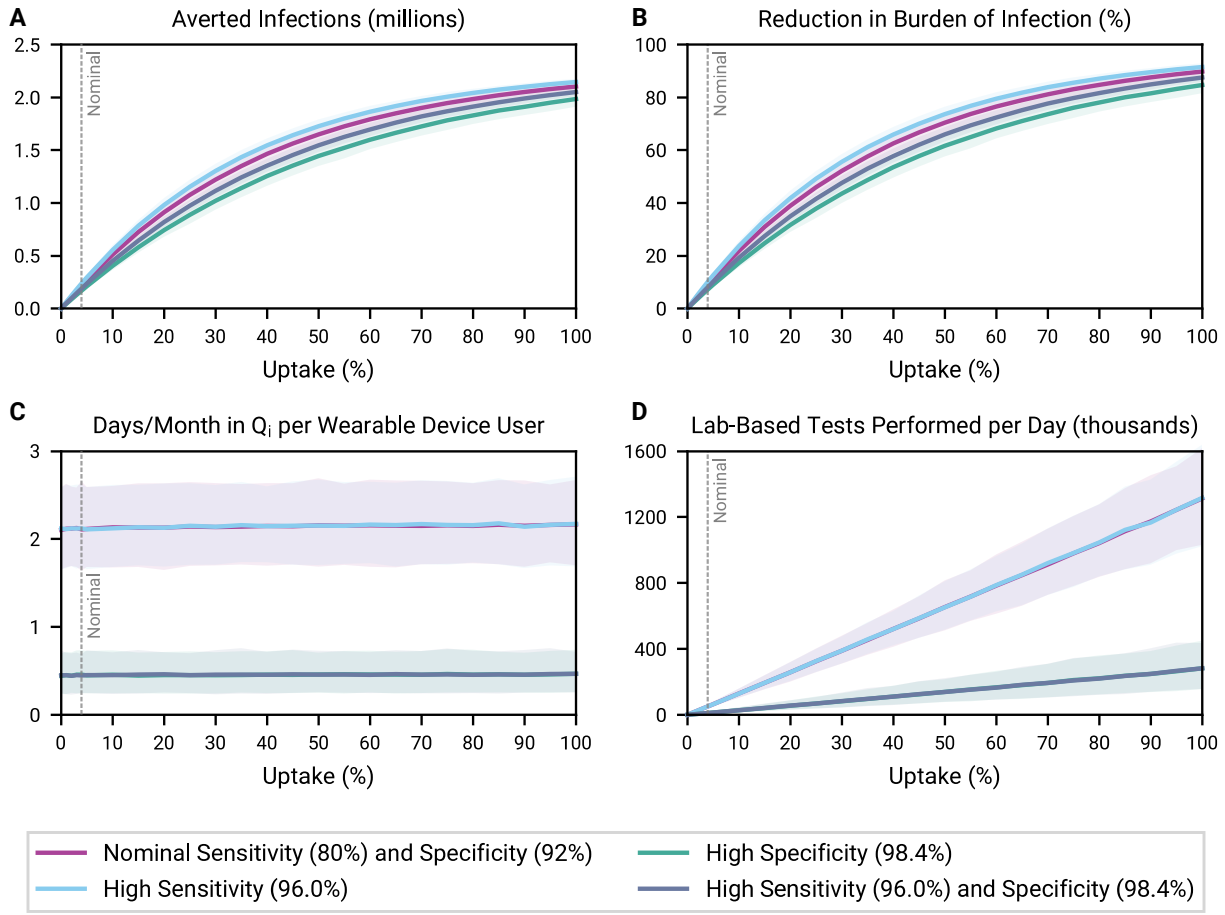


Figure 4: Impact of increasing uptake under different technology assumptions. Averted infections (A), reduction in the burden of infection (B), days incorrectly spent in quarantine per month per device user (C), and average daily demand for lab-based tests (D), all over the entire simulation period, as a function of increasing uptake. Grey dashed lines denote nominal uptake (4%). In the High Sensitivity and High Specificity scenarios, detection specificity and sensitivity are kept at their nominal values, respectively. In (C) and (D), the "Nominal Sensitivity and Specificity" and "High Sensitivity" curves overlap, and the "High Specificity" and "High Sensitivity and Specificity" curves overlap.

Impact of increasing adherence

Adherence to public health guidelines also impacts the success of pandemic control measures. Targeted policies – for example, compensating individuals in self-isolation – could help improve compliance with public health recommendations[20]. Here, we explored the role of adherence in wearable sensor deployment strategies (Figure 5; Supplementary Figure 1). We

chose to present outcomes at all levels of adherence because compliance with public health guidelines could vary greatly, for example, from as low as 14% to as high as 86% (Supplementary Table 2). We kept uptake constant at 4% and assessed multiple technology scenarios using the same definitions of "high" detection sensitivity and specificity as above.

Adherence meaningfully impacted the burden of infection (Figure 5b). For example, with nominal detection sensitivity and specificity, increasing adherence among participating wearable device users from 20% to 80% tripled the achieved reduction in the burden of infection, raising it from 4.2% (95% CI: 3.6–4.7%) to 13.6% (95% CI: 12.4–14.8%). However, increasing the proportion of users who comply with notifications also magnified the consequences of false positive notifications: the number of days incorrectly spent self-isolating per month per user (Figure 5c) and the demand for lab-based tests (Figure 5d) grew proportionally with adherence. These social and resource costs grew at a slower rate with improved detection specificity.

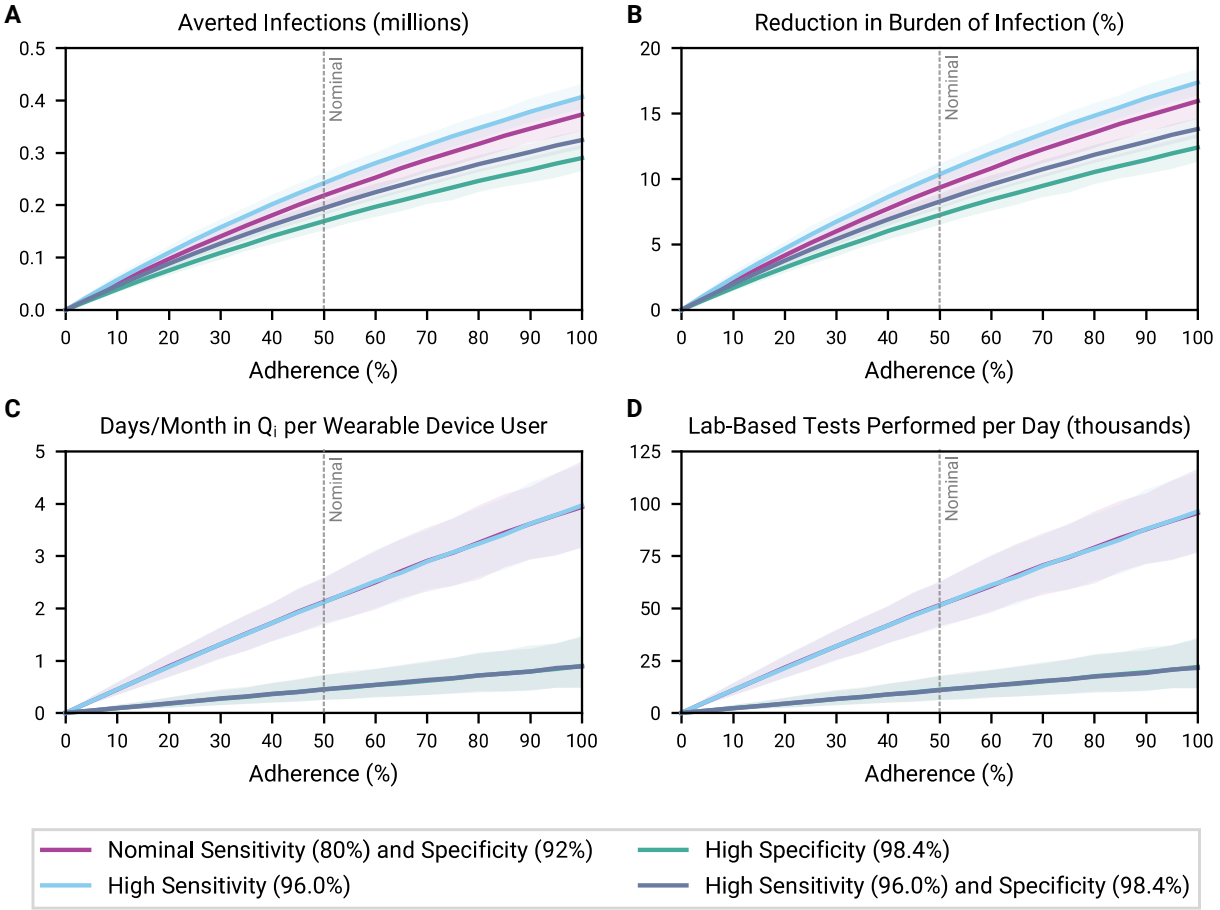


Figure 5: Impact of increasing adherence under different technology assumptions. Averted infections (A), reduction in the burden of infection (B), days incorrectly spent in quarantine per month per device user (C), and average daily demand for lab-based tests (D), all over the entire simulation period, as a function of increasing adherence. Grey dashed lines denote nominal uptake (4%). In the High Sensitivity and High Specificity scenarios, detection specificity and sensitivity are kept at their nominal values, respectively. In (C) and (D), the "Nominal Sensitivity and Specificity" and "High Sensitivity" curves overlap, and the "High Specificity" and "High Sensitivity and Specificity" curves overlap.

Impact of offering confirmatory rapid antigen tests

Our earlier findings suggested that false positive notifications of potential infection were the primary cause of unnecessary quarantines and lab-based tests. Improving detection specificity was one way to decrease false positive notifications. Here, we investigated whether offering confirmatory rapid antigen tests to users with a positive notification could also contribute to reducing unnecessary quarantines and lab-based tests (Figure 6; Table 2; Supplementary Table 6). We considered multiple

scenarios, each with either low levels of uptake (0.5%) or adherence (14%), nominal levels of uptake (4%) or adherence (50%), or high levels of uptake (12.5%) or adherence (86%). We examined these scenarios in the cases of nominal detection sensitivity and specificity, and of "high" detection sensitivity and specificity (using the same definitions of "high" as above).

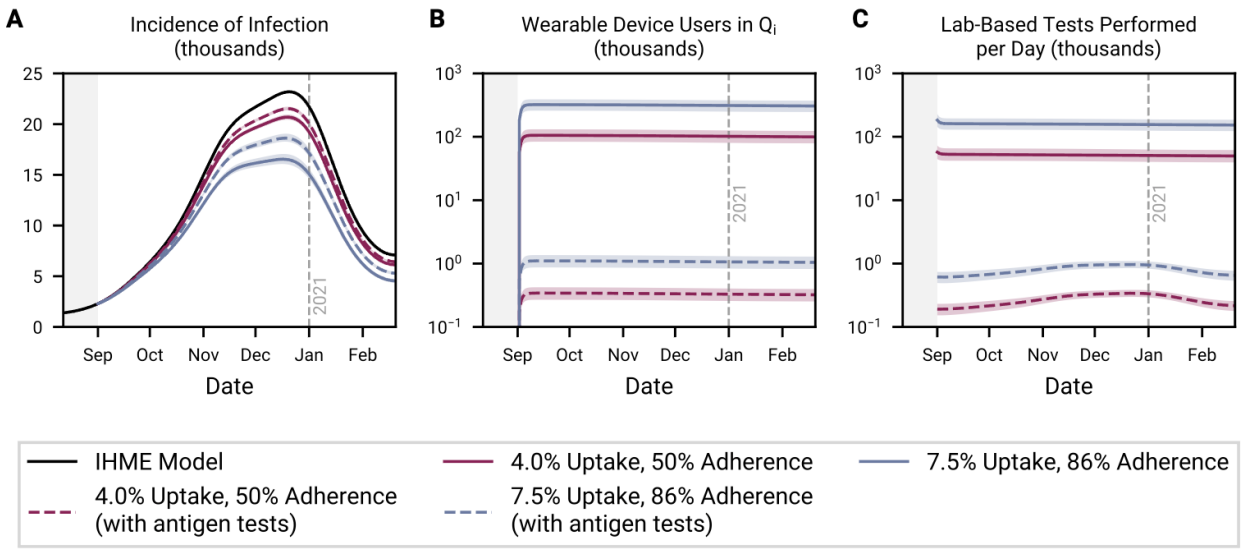


Figure 6: Wearable sensor deployment with confirmatory antigen tests. Time series depiction of the incidence of infection (A), the number of wearable device users incorrectly in quarantine (B), and the daily demand for lab-based tests (C). Detection sensitivity and specificity are set to their nominal values of 80% and 92%, respectively.

Table 2: Impact of offering confirmatory antigen tests under different technology assumptions. 95% confidence intervals are listed in parentheses. Supplementary Table 6 depicts outcomes in analogous scenarios without rapid antigen tests.

Uptake (%)	Adherence (%)	Averted Infections (thousands)	Reduction in Burden of Infection (%)	Days/Month in Q_i per User (thousands)	Additional Lab-Based Tests Performed per Day	Additional Rapid Tests Performed per Day (thousands)	Net Savings ($ million)
Nominal Detection Sensitivity (80%) and Specificity (92%) Scenario							
0.5	14	5.8 (5.0–6.6)	0.2 (0.2–0.3)	1.92 (1.48–2.46)	10 (8 – 12)	1.9 (1.5–2.5)	-0.1 (-0.5–0.3)
0.5	50	18.6 (16.4–20.5)	0.8 (0.7–0.9)	6.85 (5.27–8.55)	34 (29 – 40)	6.9 (5.3–8.6)	-0.9 (-2.5–0.6)
0.5	86	29.0 (26.0–31.4)	1.2 (1.1–1.3)	11.79 (9.20–14.70)	56 (48 – 65)	11.9 (9.3–14.9)	-2.4 (-5.2–0.1)
4.0	14	45.9 (39.2–53.2)	2.0 (1.7–2.3)	1.91 (1.47–2.42)	81 (67 – 95)	15.5 (11.9–19.6)	-0.7 (-4.4–2.7)
4.0	50	145.0 (128.9–159.7)	6.2 (5.5–6.8)	6.89 (5.40–8.49)	270 (230 – 312)	55.7 (43.7–68.7)	-8.3 (-20.6–3.1)
4.0	86	222.1 (200.9–240.9)	9.5 (8.6–10.3)	11.79 (9.35–14.87)	437 (376 – 511)	95.3 (75.7–120.2)	-21.7 (-44.9–-1.7)
7.5	14	85.9 (72.7–99.4)	3.7 (3.1–4.2)	1.90 (1.41–2.46)	150 (123 – 180)	28.9 (21.4–37.3)	-1.1 (-8.3–5.5)
7.5	50	265.4 (235.6–291.7)	11.3 (10.1–12.4)	6.85 (5.32–8.55)	494 (420 – 574)	103.8 (80.7–129.6)	-16.7 (-41.7–4.3)
7.5	86	401.8 (363.6–433.0)	17.1 (15.5–18.5)	11.86 (9.34–14.63)	800 (681 – 926)	179.9 (141.6–221.7)	-45.9 (-84.1–-10.3)
High Detection Sensitivity (96.0%) and Specificity (98.4%) Scenario							
0.5	14	6.9 (6.0–7.8)	0.3 (0.3–0.3)	0.38 (0.20–0.61)	6 (5 – 7)	0.4 (0.2–0.6)	1.7 (1.4–2.0)
0.5	50	21.6 (20.2–23.0)	0.9 (0.9–1.0)	1.36 (0.73–2.22)	20 (18 – 23)	1.4 (0.8–2.3)	5.1 (4.2–5.8)
0.5	86	33.2 (31.5–34.8)	1.4 (1.3–1.5)	2.34 (1.27–3.72)	31 (27 – 35)	2.4 (1.3–3.8)	7.6 (6.3–8.7)
4.0	14	54.7 (48.0–61.6)	2.3 (2.0–2.6)	0.38 (0.20–0.62)	50 (42 – 58)	3.2 (1.7–5.1)	13.2 (10.9–15.5)
4.0	50	168.4 (158.1–177.9)	7.2 (6.7–7.6)	1.36 (0.73–2.25)	151 (134 – 174)	11.1 (6.1–18.3)	39.5 (33.1–44.9)
4.0	86	254.2 (242.0–264.9)	10.8 (10.3–11.3)	2.37 (1.24–3.68)	228 (200 – 261)	19.3 (10.2–29.9)	57.5 (46.6–66.5)
7.5	14	101.6 (90.0–114.8)	4.3 (3.8–4.9)	0.38 (0.20–0.63)	92 (79 – 109)	5.9 (3.1–9.6)	24.6 (20.6–28.9)
7.5	50	306.6 (288.0–324.2)	13.1 (12.3–13.8)	1.36 (0.74–2.16)	270 (237 – 309)	20.8 (11.5–33.0)	71.7 (59.0–82.0)
7.5	86	455.4 (433.6–474.1)	19.4 (18.5–20.2)	2.40 (1.34–3.78)	399 (347 – 462)	36.7 (20.6–57.5)	101.4 (81.6–118.0)

The use of antigen tests reduced the number of days incorrectly spent in quarantine by ~300-fold by increasing the "effective specificity" of the strategy (Figure 6b). That is, with antigen tests, the likelihood of a *Susceptible* user being incorrectly

prompted to quarantine on a given day fell from $(1 - \nu_w)$ to the product of $(1 - \nu_w)$ and $(1 - \nu_a)$, where ν_w and ν_a are detection algorithm specificity and antigen test specificity, respectively. In earlier scenarios (Figures 4a and 5a), averted infections were decreased by improving detection specificity more than they were increased by improving detection sensitivity; fewer infections were averted in the high detection sensitivity and specificity scenarios than in the nominal detection sensitivity and specificity scenarios. Here, the specificity contributed by the antigen tests diminished the relative impact of improving detection specificity on averted infections: the "effective specificity" of the strategy was 99.976% with nominal detection specificity and 99.995% with high detection specificity (Supplementary Table 2)[21]. Instead, improving detection sensitivity was what increased the number of averted infections. Importantly, antigen tests had the secondary effect of decreasing the strategy's "effective sensitivity" – the product of antigen test sensitivity (91.1%) and detection algorithm sensitivity[21].

Offering confirmatory rapid antigen tests also decreased the demand for lab-based tests by ~200-fold, providing the double benefit of alleviating the burden on testing infrastructure and decreasing costs (Figure 6c). We earlier found that in a baseline scenario (4% uptake, 50% adherence, 80% detection sensitivity, 92% detection specificity), between ~40,000 and ~65,000 additional lab-based tests would be required each day (Figure 2c). Here, in an analogous scenario, average daily demand dropped to 270 (95% CI: 230–312) additional lab-based tests, with 55,700 (95% CI: 43,700–68,700) antigen tests being performed each day instead (Table 2). Antigen tests were also more cost-effective tools than lab-based tests for confirming *Susceptible* users with false positive notifications were not infectious. Accounting for the costs of lab-based tests and savings associated with averted hospitalizations, we approximated that a baseline scenario would involve $825.7 (95% CI: $645.2–1,035.7) million in net healthcare expenditures (Supplementary Table 6). Here, we approximated that the analogous scenario would involve $8.3 (95% CI: -3.1–20.6) million in net expenditures (Table 2). Unnecessary lab-based tests drove this ~10-fold difference in costs.

DISCUSSION

We used a counterfactual model of Canada's second COVID-19 wave to demonstrate that wearable sensors capable of detecting infectious diseases before or absent symptoms can be useful tools for pandemic mitigation. Through continuous and non-invasive monitoring of physiological parameters, these devices can conceivably help FTTI systems identify hidden infection chains with minimal delay and without active user engagement or broad sharing of user data. We showed that (1) deploying today's detection algorithms could have meaningfully reduced the second wave burden of infection, but with substantial social and resource costs; (2) improving detection algorithm specificity and offering confirmatory rapid antigen tests can help minimize unnecessary quarantines and lab-based tests; and (3) once false positive notifications are minimized, increasing uptake and adherence become effective strategies to scale the number of averted infections.

In theory, wearable sensor deployment reduces the burden of infection by decreasing the pool of infectious individuals (a function of detection algorithm sensitivity). Here, we found that detection specificity played an unexpectedly large role as well, with false positive notifications of potential infection prompting unnecessary quarantines and thereby decreasing the pool of susceptible individuals. Thus, although prioritizing uptake and adherence as part of a wearable sensor deployment strategy could mitigate a substantial number of infections, the unsustainable growth of associated costs should also be considered. In a baseline scenario, without improvements to detection specificity, every user would spend over two days a month on average incorrectly quarantining, and ~40,000 to ~65,000 additional confirmatory lab-based tests would be required each day. The social and economic harm caused by solely promoting uptake or adherence without improvements to detection specificity would likely undermine public confidence in and compliance with a wearable-based pandemic mitigation strategy[22]. Alavi *et al* found that many false positives were due to the detection algorithm identifying lifestyle-driven changes in resting heart rate (e.g., after intense exercise or alcohol consumption); accounting for these factors using more advanced algorithms may be one way to target improved detection specificity[10].

We found that the inclusion of confirmatory antigen testing was a valuable mechanism, beyond improving detection specificity, to increase the "effective specificity" of the strategy and decrease the overall false positive rate. The inclusion of antigen testing decreased days incorrectly spent in quarantine by ~300-fold and brought the additional demand on lab-based testing infrastructure to justifiable levels. In general, antigen tests were more cost-effective and immediate tools than lab-based tests to confirm a *Susceptible* device user was not infectious. However, even with the inclusion of antigen tests, improvements to detection specificity still had value. Relative to scenarios with nominal detection specificity, with "high" detection specificity, we observed a ~4-fold reduction in days incorrectly spent in quarantine per month per user, a ~2-fold reduction in lab-based tests performed each day, and a ~5-fold reduction in antigen tests used each day. Importantly, a strategy in which antigen tests support the deployment of wearable sensors is notably different from one involving frequent use of rapid antigen tests for surveillance testing[23]. On their own, broad antigen test-based screening approaches require tremendous manufacturing volumes, infrastructure, and funding[24]. Conversely, wearable sensors can non-invasively detect infections without active user engagement, reducing the effort required to participate. Further, in the scenarios explored here, detection algorithm specificity governed how efficiently rapid antigen tests were deployed.

Our work has important limitations. First, we assumed that SARS-CoV-2 epidemiology and wearable device use were homogenous within the population. Our determination of the transmission rate from epidemiological data inherently results in a "well-mixed" approximation for this value averaged over population-level heterogeneities such as age and super-spreading. As well, the COVID-19 pandemic has disproportionately impacted low-income and minority groups, while younger and wealthier individuals are more likely to own wearable devices[25,26]. Future studies could consider a policy where the incentive to download the application is increased to become a subsidy for purchasing a wearable device, reducing the participation barrier. Second, we used an existing model for the incidence of infection which has its own assumptions and limitations[13]. Third, we made the simplifying assumption that all users without symptoms (and that no users with symptoms) could benefit from wearable-informed prompts to take a confirmatory test and self-isolate. Fourth, we did not consider how uptake or adherence might vary over time or with detection accuracy[18,22]. Fifth, we used median values for SARS-CoV-2 infection parameters (e.g., latent period) and did not account for reinfections or vaccinations. Finally, we did not account

for administrative costs or second order savings (e.g., avoided worker's compensation payouts for 14-day quarantines).

To our knowledge, this is the first study to explore the real-world impact of deploying wearable sensors for pandemic mitigation. Using the example of COVID-19, we demonstrated that these devices have the potential to support FTTI systems with real-time detection of presymptomatic and asymptomatic infections, and ultimately reduce the burden of infection. However, in cases with insufficient detection algorithm specificity, complementary interventions that reduce false positives are required to minimize social costs and resource demands. In the future, there is clear merit to further exploring how wearable sensors can be incorporated into FTTI systems for pandemic mitigation, and whether these devices continue to have public health utility once an endemic phase is reached, complementary to vaccines.

STATEMENTS

Acknowledgements: We would like to thank Chadi M. Saad-Roy (Princeton University) and Étienne Racine (McGill University) for guidance with modeling, Andrew Soltan (Oxford University) for insightful discussions, and Arash Alavi (Stanford University) for assistance with data interpretation.

Author Contributions: ND and GB contributed to the conceptualization; ND, RKA, GB, MW, and CEW contributed to the methodology; ND and GB contributed to the writing the original draft; ND contributed to software; ND and RKA contributed to visualization; ND, RKA, GB, MW, MPS, JRC, and CEW contributed to review and editing; ND, JRC, and CEW contributed to supervision.

Competing Interests: MPS is a co-founder and member of the scientific advisory board of Personalis, Qbio, January, SensOmics, Protos, Mirvie, NiMo, Onza and Oralome. He is also on the scientific advisory board of Danaher, Genapsys and Jupiter.

Materials and Correspondence: Correspondence and material requests should be addressed to Nathan Duarte and Caroline E. Wagner.

REFERENCES

1. Rajan, S., D. Cylus, J. & Mckee, M. What do countries need to do to implement effective 'find, test, trace, isolate and support' systems? *J. R. Soc. Med.* **113**, 245–250 (2020).
2. Contreras, S. *et al.* The challenges of containing SARS-CoV-2 via test-trace-and-isolate. *Nat. Commun.* **12**, 378 (2021).
3. Grantz, K. H. *et al.* Maximizing and evaluating the impact of test-trace-isolate programs: A modeling study. *PLOS Med.* **18**, e1003585 (2021).
4. Lewis, D. Why many countries failed at COVID contact-tracing — but some got it right. *Nature* (2020).
5. Rubin, R. The Challenges of Expanding Rapid Tests to Curb COVID-19. *JAMA* **324**, 1813 (2020).
6. Freeze, C. Supply crunch coming as demand for COVID-19 rapid testing soars. *The Globe and Mail* (2021).
7. Li, X. *et al.* Digital Health: Tracking Physiomes and Activity Using Wearable Biosensors Reveals Useful Health-Related Information. *PLOS Biol.* **15**, e2001402 (2017).
8. Grzesiak, E. *et al.* Assessment of the Feasibility of Using Noninvasive Wearable Biometric Monitoring Sensors to Detect Influenza and the Common Cold Before Symptom Onset. *JAMA Netw. Open* **4**, e2128534 (2021).
9. Mishra, T. *et al.* Pre-symptomatic detection of COVID-19 from smartwatch data. *Nat. Biomed. Eng.* **4**, 1208–1220 (2020).
10. Alavi, A. *et al.* Real-time alerting system for COVID-19 and other stress events using wearable data. *Nat. Med.* (2021) doi:10.1038/s41591-021-01593-2.
11. Government of Canada. *Testing for COVID-19: How and where we test for active infections*. https://www.canada.ca/en/public-health/services/diseases/2019-novel-coronavirus-infection/symptoms/testing/diagnosing.html (2021).
12. Lumley, S. F. *et al.* Antibody Status and Incidence of SARS-CoV-2 Infection in Health Care Workers. *N. Engl. J. Med.* **384**, 533–540 (2021).
13. Institute for Health Metrics and Evaluation (IHME). *COVID-19 Projections*. https://covid19.healthdata.org/projections (2020).

14. Bobrovitz, N. *et al.* Global seroprevalence of SARS-CoV-2 antibodies: A systematic review and meta-analysis. *PLOS ONE* **16**, e0252617 (2021).

15. Cleary, J. L., Fang, Y., Sen, S. & Wu, Z. *A Caveat to Using Wearable Sensor Data for COVID-19 Detection: The Role of Behavioral Change after Receipt of Test Results*. http://medrxiv.org/lookup/doi/10.1101/2021.04.17.21255513 (2021) doi:10.1101/2021.04.17.21255513.

16. Husereau, D. *et al.* Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement. *BMJ* **346**, f1049–f1049 (2013).

17. Government of Canada. *COVID-19 daily epidemiology update*. https://health-infobase.canada.ca/covid-19/epidemiological-summary-covid-19-cases.html.

18. Munzert, S., Selb, P., Gohdes, A., Stoetzer, L. F. & Lowe, W. Tracking and promoting the usage of a COVID-19 contact tracing app. *Nat. Hum. Behav.* **5**, 247–255 (2021).

19. Vavreck, L. $100 as Incentive to Get a Shot? Experiment Suggests It Can Pay Off. *The New York Times* (2021).

20. Bodas, M. & Peleg, K. Income assurances are a crucial factor in determining public compliance with self-isolation regulations during the COVID-19 outbreak – cohort study in Israel. *Isr. J. Health Policy Res.* **9**, 54 (2020).

21. Abbot. *In vitro diagnostic rapid test for qualitative detection of SARS-CoV-2 antigen (Ag)*. https://www.globalpointofcare.abbott/en/product-details/panbio-covid-19-ag-antigen-test.html.

22. Kaptchuk, G., Goldstein, D. G., Hargittai, E., Hofman, J. & Redmiles, E. M. How good is good enough for COVID19 apps? The influence of benefits, accuracy, and privacy on willingness to adopt. *ArXiv200504343 Cs* (2020).

23. Larremore, D. B. *et al.* Test sensitivity is secondary to frequency and turnaround time for COVID-19 screening. *Sci. Adv.* **7**, eabd5393 (2021).

24. Iacobucci, G. Covid-19: Government rolls out twice weekly rapid testing to all in England. *BMJ* n902 (2021) doi:10.1136/bmj.n902.

25. Liao, T. F. & De Maio, F. Association of Social and Economic Inequality With Coronavirus Disease 2019 Incidence and Mortality Across US Counties. *JAMA Netw. Open* **4**, e2034578 (2021).

26. Vogels, E. About one-in-five Americans use a smart watch or fitness tracker. *Pew Research Centre*.

409

Supplementary Appendix

1. Modeling assumptions, variables, and equations

Supplementary Table 1: SARS-CoV-2 characteristics. We obtained the presymptomatic infectious period by subtracting the latent period (3.69 days) from the incubation period (5 days)[1,2]. We obtained the asymptomatic and symptomatic infectious period by subtracting the presymptomatic infectious period (1.31 days) from the total infectious period (7 days)[1–3]. During Monte Carlo simulations, we modeled asymptomatic prevalence as a beta random variable with a mean of 0.4 and a sample size of 200; a sample size of 200 is greater than 75% of the study populations examined by Oran and colleagues[4].

Parameter	Symbol	Value
Latent period (exposure to infectiousness)	$1 / \alpha$	3.69 days[1]
Presymptomatic infectious period	$1 / \tau$	1.31 days[1,2]
Asymptomatic and symptomatic infectious period	$1 / \gamma$	5.69 days[1–3]
Asymptomatic prevalence	ρ	40%[4]
Transmission potential without symptoms relative to with symptoms	λ	55%[1]
Daily new infections	π	varies
Average number of transmissions per infectious person per day	β	varies

Supplementary Table 2: Wearable device and policy parameters.

Parameter	Symbol	Nominal Value	Notes
Wearable device uptake	θ	4%	We defined uptake as the percent of the population that owns a wearable device, has downloaded the detection/notification application, *and* uses the device enough to collect sufficient data for detection. We estimated uptake would range from 0.5% to 7.5% at baseline. Supplementary Tables 3 and 4 provide an example of this calculation. Device ownership: Estimates from 2018 place device ownership in Canada between 22% and 25%[5,6]. Download rate: A study found that the baseline download rate of Germany's national contact tracing application was between ~8% and ~11%; incentives could increase this rate[7]. Paré and colleagues found that 57% of owners regularly track their health with their device[5]. Here, we looked at download rates ranging from 10% to 60%. Utilization: In research studies where wearable devices were used to track health, usage rates ranged from as low as 24%, to 50% in the long run[8,9].
Adherence to wearable device notification	ψ	50%	We defined adherence as the proportion of wearable device users that comply with recommended next steps upon notification of potential infection. At baseline, next steps include seeking a confirmatory lab-based test, self-isolating while awaiting a result, and quarantining until recovery if the result is positive. With antigen tests, compliant users also take a confirmatory rapid antigen test prior to seeking a lab-based test. In Israel, at least ~53% of antigen test kit results were reported (~613,000 reports out of ~1,150,000 kits taken home)[10,11]. In Norway, up to 70% of those with a suspected diagnosis and up to 86% with a positive diagnostic test adhered to quarantine[12]. However, adherence could be as low as ~14%, if one considers Canada's COVID Alert contact tracing app's reporting rate in light of confirmed cases as of July 27, 2021[13]. During Monte Carlo simulations, we modeled adherence as a beta random variable with a mean of 0.5 and a sample size of 1723; 1723 was the sample size for the relevant experiment in the Norway study.
Detection algorithm sensitivity	σ_w	80%	We defined sensitivity as the proportion of infected yet asymptomatic wearable device users (i.e., *Exposed*, *Presymptomatic*, and *Asymptomatic*) who receive a notification of potential infection prior to recovering and entering the *Removed* compartment. Alavi and colleagues' NightSignal algorithm achieved a sensitivity of ~80%, which is a plausible value based on other efforts to develop similar algorithms[14,15]. During Monte Carlo simulations, we modeled sensitivity as a beta random variable with a mean of 0.80 and a sample size of 84; 84 was the size of the sample used to calculate the NightSignal algorithm's sensitivity.
Detection algorithm specificity	ν_w	92%	We defined specificity as the probability that, on a given day, a healthy (i.e., *Susceptible*) user does not receive a notification of potential infection. Alavi and colleagues' NightSignal algorithm gave potentially healthy users 0.0819 false positive notifications ("red alerts") per day on average, corresponding to a specificity of ~92%[14]. During Monte Carlo simulations, we modeled specificity as a beta random variable with a mean of 0.92 and a sample size of 818; alarm data from 818 potentially healthy users in Alavi and colleagues' dataset were used to calculate the false positive rate.

Detection algorithm sensitivity adjustment factor	κ	$(1/\alpha + 1/\tau + 1/\lambda)^{-1}$ days^{-1}	We assumed that σ_w is applied uniformly across *Exposed*, *Presymptomatic*, and *Asymptomatic* states such that by the time infectiousness ends, σ_w of infected users without symptoms are notified.
Detection algorithm specificity adjustment factor	χ	$(1)^{-1}$ days^{-1}	We assumed that ν_w is applied over a period of one day such that on any given day, *Susceptible* users receive an incorrect (i.e., false positive) notification of potential infection with probability $(1-\nu_w)$.
Rapid antigen test sensitivity	σ_a	91.1%	We used the lowest performance reported for Abbot's Panbio test[16]. Independent evaluations of this test suggest lower sensitivity in presymptomatic and asymptomatic individuals[17,18]. However, we reasoned that the receipt of a wearable-informed notification of potential infection raised the pretest probability of infection, thereby improving the negative predictive value (NPV) of the tests[19]. We performed a sensitivity analysis on the sensitivity of rapid antigen tests (Supplementary Figure 5).
Rapid antigen test specificity	ν_a	99.7%	We used the lowest performance reported for Abbot's Panbio test[16]. Independent evaluations suggest that this test's specificity is even higher[17,18].
Lab-based test turnaround time	ε	2 days	In general, 1-3 days based on Health Canada reporting[20].
Transmission rate increase factor	a	1	Used to test the scenario in which device users who do not receive a notification of potential infection act in a riskier fashion, effectively increasing their transmission rate[21].

Model equations, in addition to Equation (1), used to extract transmission rate (β) from the IHME infection model:

$$\frac{dS}{dt} = -\pi \tag{3}$$

$$\frac{dE}{dt} = \pi - \alpha E \tag{4}$$

$$\frac{dI_p}{dt} = \alpha E - \tau I_p \tag{5}$$

$$\frac{dI_a}{dt} = \rho\tau I_p - \gamma I_a \tag{6}$$

$$\frac{dI_s}{dt} = (1-\rho)\tau I_p - \gamma I_s \tag{7}$$

$$\frac{dR}{dt} = \gamma I_a + \gamma I_s \tag{8}$$

Model equations, in addition to Equation (2), used to run simulations:

$$\frac{dS_w}{dt} = -\theta\pi - \psi\chi(1-\nu_w)(1-\nu_a)S_w + \frac{Q_i}{\epsilon} \tag{9}$$

$$\frac{dE_w}{dt} = \theta\pi - \psi\kappa\sigma_w\sigma_a E_w - \alpha E_w \tag{10}$$

$$\frac{dI_{p,w}}{dt} = \alpha E_w - \psi\kappa\sigma_w\sigma_a I_{p,w} - \tau I_{p,w} \tag{11}$$

$$\frac{dI_{a,w}}{dt} = \rho\tau I_{p,w} - \psi\kappa\sigma_w\sigma_a I_{a,w} - \gamma I_{a,w} \tag{12}$$

$$\frac{dI_{s,w}}{dt} = (1-\rho)\tau I_{p,w} - \gamma I_{s,w} \tag{13}$$

$$\frac{dS_{nw}}{dt} = -(1-\theta)\pi \tag{14}$$

$$\frac{dE_{nw}}{dt} = (1-\theta)\pi - \alpha E_{nw} \tag{15}$$

$$\frac{dI_{p,nw}}{dt} = \alpha E_{nw} - \tau I_{p,nw} \tag{16}$$

$$\frac{dI_{a,nw}}{dt} = \rho\tau I_{p,nw} - \gamma I_{a,nw} \tag{17}$$

$$\frac{dI_{s,nw}}{dt} = (1-\rho)\tau I_{p,nw} - \gamma I_{s,nw} \tag{18}$$

$$\frac{dR}{dt} = \gamma(I_{a,w} + I_{s,w} + I_{a,nw} + I_{s,nw}) + \frac{Q_c}{\epsilon} \tag{19}$$

$$\frac{dQ_i}{dt} = \psi\chi(1-\nu_w)(1-\nu_a)S_w - \frac{Q_i}{\epsilon} \tag{20}$$

$$\frac{dQ_c}{dt} = \psi\kappa\sigma_w\sigma_a(E_w + I_{p,w} + I_{a,w}) - \frac{Q_c}{\epsilon} \tag{21}$$

2. Estimating ranges for uptake

Below, Supplementary Tables 3 and 4 delineate how we estimated plausible ranges for uptake in a baseline scenario. First, we multiplied the download rate by the proportion of the population that owns a device to obtain a plausible range for the proportion of the population that owns the application (Supplementary Table 3). Then, we calculated a range for uptake by multiplying the proportion of individuals that own the application by expected levels of utilization (Supplementary Table 4). Accounting for utilization was a necessary step because not all individuals who download the application to their device use it enough to provide sufficient data for the algorithm to function correctly[14]. We concluded that in a baseline scenario, uptake would likely range from 0.5% to 7.5%. Assumptions around parameter values are listed in Supplementary Table 2 above.

Supplementary Table 3: Calculation of the proportion of individuals who own the application.

Application Ownership	*Device Ownership*		
Download Rate	*22.0%*	*23.5%*	*25.0%*
10.0%	2.2%	2.4%	2.5%
35.0%	7.7%	8.2%	8.8%
60.0%	13.2%	14.1%	15.0%

Supplementary Table 4: Calculation of the proportion of individuals who own and use the application.

Uptake	*Utilization*		
Application Ownership	*24.0%*	*37.0%*	*50.0%*
2.2%	0.5%	0.8%	1.1%
8.6%	2.1%	3.2%	4.3%
15.0%	3.6%	5.6%	7.5%

3. Estimating economic impact

To generate a first-order estimate of net healthcare expenditures in particular scenarios, we subtracted savings from expenditures. Savings included costs avoided from averting hospitalizations. Expenditures included the costs of lab-based tests and antigen tests, where applicable. We calculated the number of averted hospitalizations and tests from our simulations using Equations (22) to (24) below. We multiplied these volumes by the costs in Supplementary Table 5, as in Equation (25), to generate a first-order approximation of net healthcare expenditures.

$$n_{averted\ hosp} = Pr\{hospitalization|infection\} n_{averted\ inf} \tag{22}$$

$$n_{lab\ tests} = \psi \sum_t \left[\chi(1-\nu_w)(1-\nu_a)S_w + \kappa\sigma_w\sigma_a(E_w + I_{p,w} + I_{a,w}) \right] \tag{23}$$

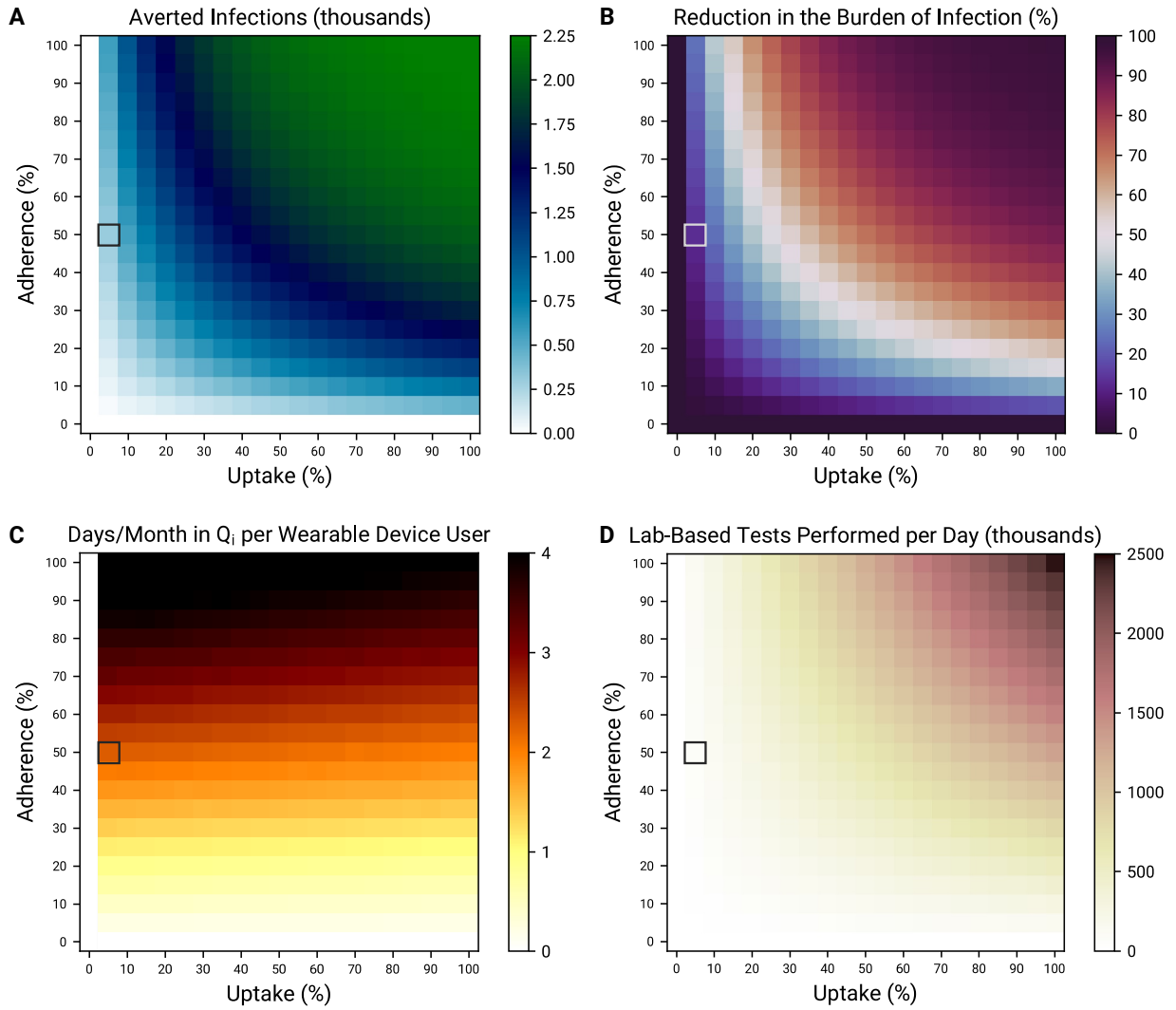
$$n_{rapid\ tests} = \psi \sum_t \left[\chi(1-\nu_w)S_w + \kappa\sigma_w(E_w + I_{p,w} + I_{a,w}) \right] \tag{24}$$

$$net\ expenditures = 100\ n_{lab\ tests} + 5\ n_{rapid\ tests} - 23,471\ n_{averted\ hosp} \tag{25}$$

Supplementary Table 5: Costs used in economic impact calculations and associated assumptions.

Parameter	Value	Notes
Average cost of a COVID-19 hospitalization	$23,471	Average cost between January 2020 and March 2021; includes normal visits and ICU visits[22]. Data downloaded on December 16, 2021.
Infection hospitalization ratio	1.31%	We calculated the infection hospitalization ratio by dividing the total number of COVID-19 hospitalizations by the total number of infections, all within the simulation period[23]. We offset the hospitalization counts by 7 days to align them with infections.
Cost of an antigen test	$5	Based on rates achieved when tests are purchased at scale[24].
Cost of a diagnostic test	$100	Large variation in this figure; used rule of thumb[25].

4. Additional results



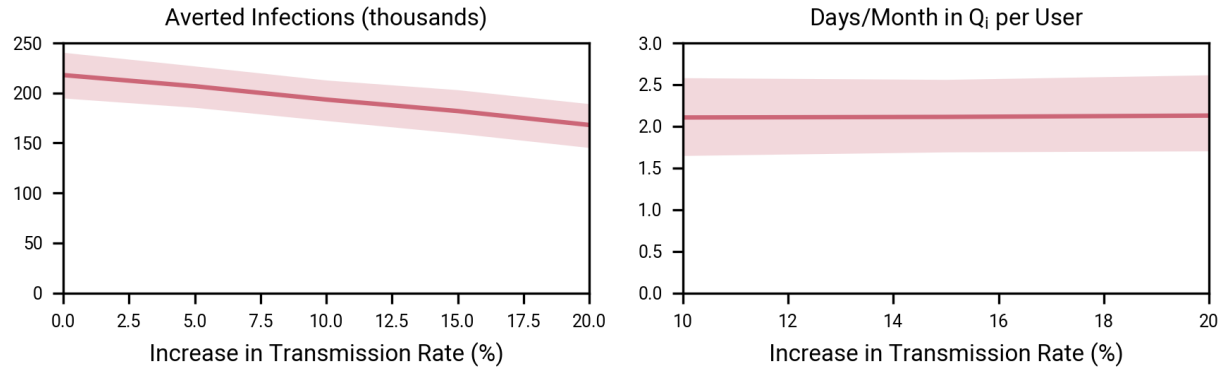
Supplementary Figure 1: Impact of simultaneously increasing uptake and adherence. Averted infections (A), reduction in the burden of infection (B), days incorrectly spent in quarantine per month per device user (C), and average daily demand for lab-based tests (D), all over the entire simulation period, as a function of uptake and adherence. Grey boxes denote nominal sensitivity (80%) and specificity (92%).

Supplementary Table 6: Wearable sensor deployment scenarios under different technology assumptions. 95% confidence intervals are listed in parentheses. This table serves as a counterpart to Table 2 so that analogous scenarios with and without confirmatory antigen tests can be compared.

Uptake (%)	Adherence (%)	Averted Infections (thousands)	Reduction in Burden of Infection (%)	Days/Month in Q_i per User	Additional Lab-Based Tests Performed per Day (thousands)	Net Savings ($ million)
Nominal Detection Sensitivity (80%) and Specificity (92%) Scenario						
0.5	14	8.8 (7.5–10.1)	0.4 (0.3–0.4)	0.62 (0.47–0.78)	1.9 (1.4 – 2.4)	-30.0 (-38.2– -22.5)
0.5	50	28.5 (25.7–31.4)	1.2 (1.1–1.3)	2.13 (1.68–2.63)	6.4 (5.1 – 8.0)	-102.8 (-128.8– -79.9)
0.5	86	44.7 (40.6–48.6)	1.9 (1.7–2.1)	3.46 (2.76–4.18)	10.5 (8.4 – 12.7)	-167.8 (-204.5– -132.2)
4.0	14	69.9 (60.0–80.7)	3.0 (2.6–3.4)	0.62 (0.48–0.80)	15.2 (11.8 – 19.4)	-240.7 (-311.3– -183.5)
4.0	50	218.9 (197.1–241.0)	9.3 (8.4–10.3)	2.13 (1.69–2.64)	51.6 (40.9 – 64.1)	-825.7 (-1,035.7– -645.2)
4.0	86	335.5 (306.7–363.2)	14.3 (13.1–15.5)	3.46 (2.79–4.16)	84.0 (67.7 – 100.9)	-1,349.6 (-1,637.9– -1,073.9)
7.5	14	129.1 (110.3–148.0)	5.5 (4.7–6.3)	0.62 (0.48–0.79)	28.4 (22.0 – 36.0)	-452.3 (-580.5– -343.8)
7.5	50	394.1 (358.8–428.2)	16.8 (15.3–18.3)	2.13 (1.69–2.60)	96.9 (77.0 – 118.2)	-1,555.8 (-1,918.6– -1,215.5)
7.5	86	589.9 (543.3–633.0)	25.2 (23.2–27.0)	3.47 (2.79–4.20)	157.8 (126.8 – 191.1)	-2,548.3 (-3,111.5– -2,021.9)
High Detection Sensitivity (96.0%) and Specificity (98.4%) Scenario						
0.5	14	8.1 (7.0–9.2)	0.3 (0.3–0.4)	0.13 (0.07–0.21)	0.4 (0.2 – 0.6)	-4.4 (-8.4– -1.3)
0.5	50	25.1 (23.4–26.9)	1.1 (1.0–1.1)	0.45 (0.24–0.69)	1.4 (0.7 – 2.1)	-16.0 (-28.7– -5.4)
0.5	86	38.5 (36.3–40.9)	1.6 (1.5–1.7)	0.76 (0.40–1.24)	2.3 (1.2 – 3.8)	-28.5 (-53.1– -10.0)
4.0	14	63.4 (55.6–71.1)	2.7 (2.4–3.0)	0.13 (0.07–0.21)	3.1 (1.7 – 5.0)	-33.9 (-66.6– -10.7)
4.0	50	194.1 (180.8–207.5)	8.3 (7.7–8.9)	0.44 (0.24–0.71)	10.8 (5.9 – 17.4)	-127.5 (-237.4– -45.6)
4.0	86	292.2 (275.4–309.1)	12.5 (11.8–13.2)	0.77 (0.41–1.21)	18.7 (10.2 – 29.6)	-233.8 (-419.1– -90.3)
7.5	14	117.7 (103.6–133.6)	5.0 (4.4–5.7)	0.13 (0.07–0.20)	5.8 (3.1 – 9.2)	-64.2 (-122.7– -18.3)
7.5	50	352.2 (328.2–378.0)	15.0 (14.0–16.1)	0.45 (0.24–0.72)	20.5 (11.3 – 32.7)	-246.8 (-449.3– -88.9)
7.5	86	520.4 (493.7–549.2)	22.2 (21.1–23.4)	0.77 (0.43–1.22)	35.4 (20.0 – 55.7)	-452.2 (-796.5– -190.4)

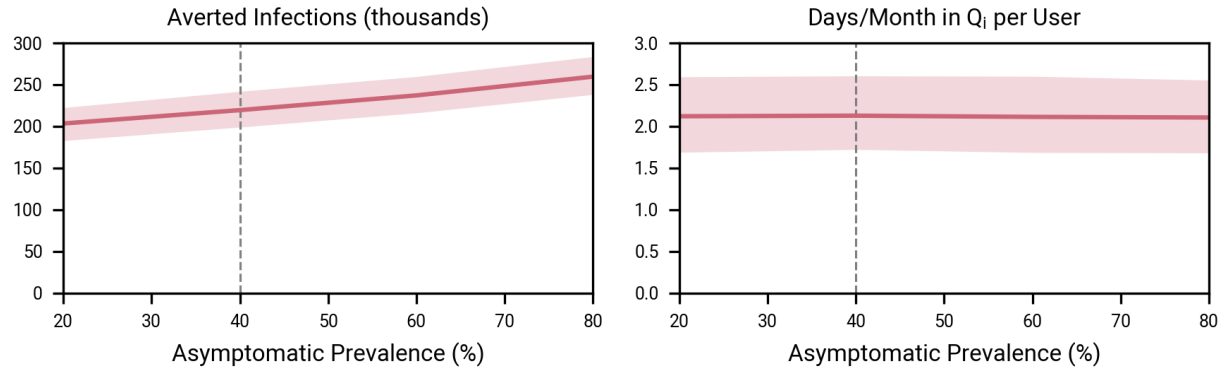
5. Sensitivity analyses

First, we looked at whether, instead of remaining constant, the transmission rate of device users who do not receive a notification of potential infection might in fact increase due to a sense of false confidence (Supplementary Figure 2)[21]. We modulated *a* in Equation (2) to do so; a 5% increase in the transmission rate, for example, meant *a* was set to 1.05. Increase in transmission among device users relative to historical levels reduced the number of averted infections. The number of incorrect quarantines was not impacted. The implication of this finding is that public health leaders would need to communicate the limitations of wearable sensors with respect to detecting infections and emphasize that a lack of a notification does not rule out potential infection.



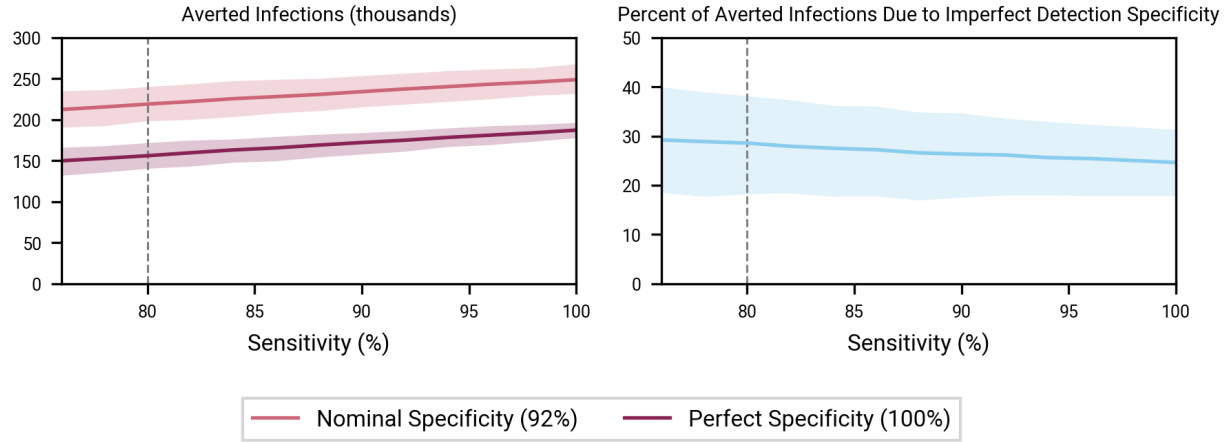
Supplementary Figure 2: Impact of increased transmission among device users who are not notified of potential infection in a baseline scenario. We assumed 4% uptake, 50% adherence, and that transmission among non-users was unchanged.

Second, we looked at the impact of asymptomatic prevalence (Supplementary Figure 3). It was important to perform this analysis because our model only accounts for notifications sent to presymptomatic and asymptomatic individuals, yet there is still a lack of consensus on a specific value for the asymptomatic prevalence[4]. As expected, with greater asymptomatic prevalence, more individuals could benefit from wearable device use and more infections could be averted. The success of this strategy was not dependent on a particular value of asymptomatic prevalence. The number of incorrect quarantines was not impacted.



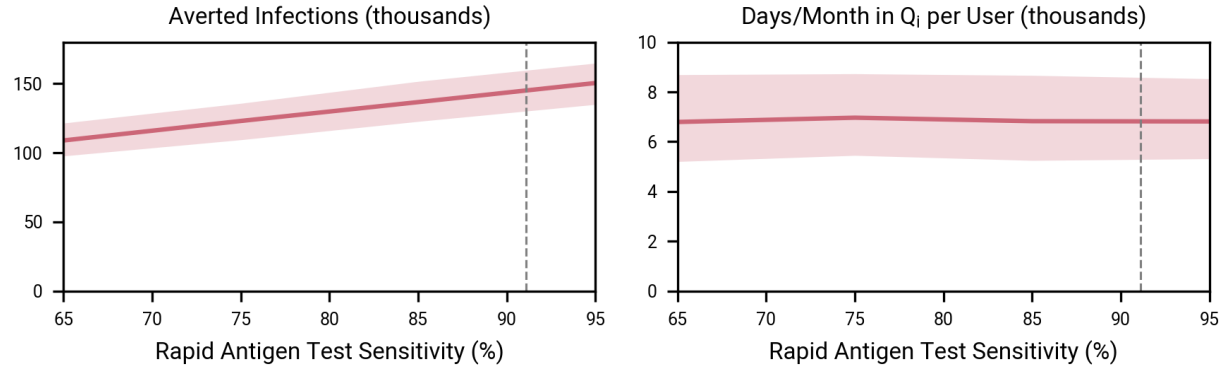
Supplementary Figure 3: Impact of asymptomatic prevalence in a baseline scenario. We assumed 4% uptake and 50% adherence. We continued to model asymptomatic prevalence as a beta-distributed random variable (Supplementary Table 1). The dashed grey line represents nominal asymptomatic prevalence (40%).

Third, we investigated how the decrease in the number of *Susceptible* individuals resulting from incorrect quarantines influenced the number of averted infections in a baseline scenario (Supplementary Figure 4). We compared the number of averted infections with nominal (92%) and perfect (100%) detection specificity as detection sensitivity increased. In a baseline scenario, a meaningful proportion – 28.6% (95% CI: 18.3–38.2%) in the case of nominal detection sensitivity – of averted infections were driven by incorrect quarantines. This proportion decreased with increasing sensitivity. Consistent with earlier findings, it remains important to minimize unnecessary quarantines – social costs can be substantially reduced while still averting a meaningful number of infections, especially if detection sensitivity can be improved in parallel.



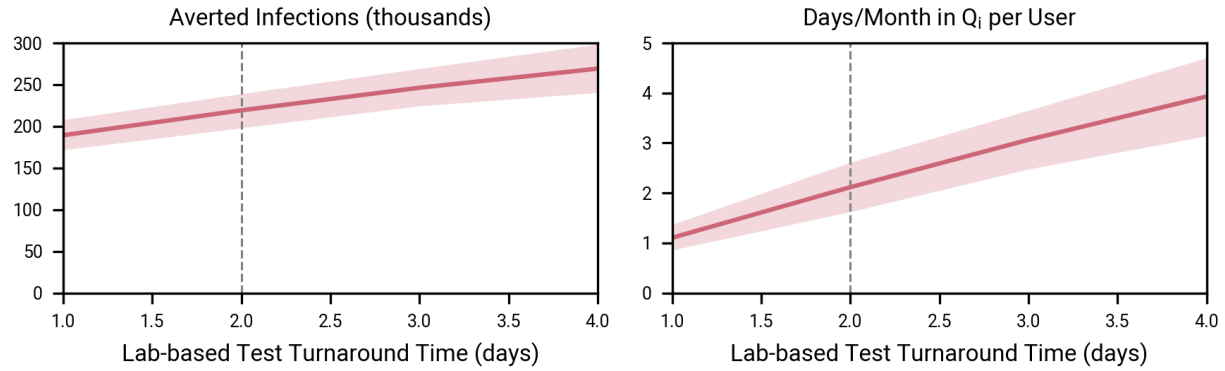
Supplementary Figure 4: Impact of incorrect quarantines on averted infections in a baseline scenario. We assumed 4% uptake and 50% adherence. The dashed grey line depicts nominal detection algorithm sensitivity (80%).

Fourth, we investigated the impact of lower rapid antigen test sensitivity (Supplementary Figure 5). First, fewer infections were averted overall – as discussed, using antigen tests to minimize incorrect quarantines increased the pool of susceptible individuals. Second, averted infections grew linearly with test sensitivity: a ~10% increase in test sensitivity resulted in a ~14,000 increase in averted infections. These two effects result in a tradeoff between missing more infectious individuals (imperfect antigen test sensitivity) and decreasing false positive prompts to seek a lab-based test and self-isolate while awaiting the results (near perfect antigen test specificity). We believe the use of antigen tests as a complementary mechanism could be justified. Although fewer infections are averted relative to wearable sensor deployment without antigen tests, hundreds of thousands of infections are still averted relative to the counterfactual scenario – and in a resource-efficient and socially acceptable fashion. Certainly, infected individuals with a false negative antigen test might act in a riskier fashion – even still, there remains an opportunity to avert a substantial number of infections (Supplementary Figure 2). We also point out that improving detecting algorithm sensitivity could help counteract the effect of imperfect antigen test sensitivity (Table 2).



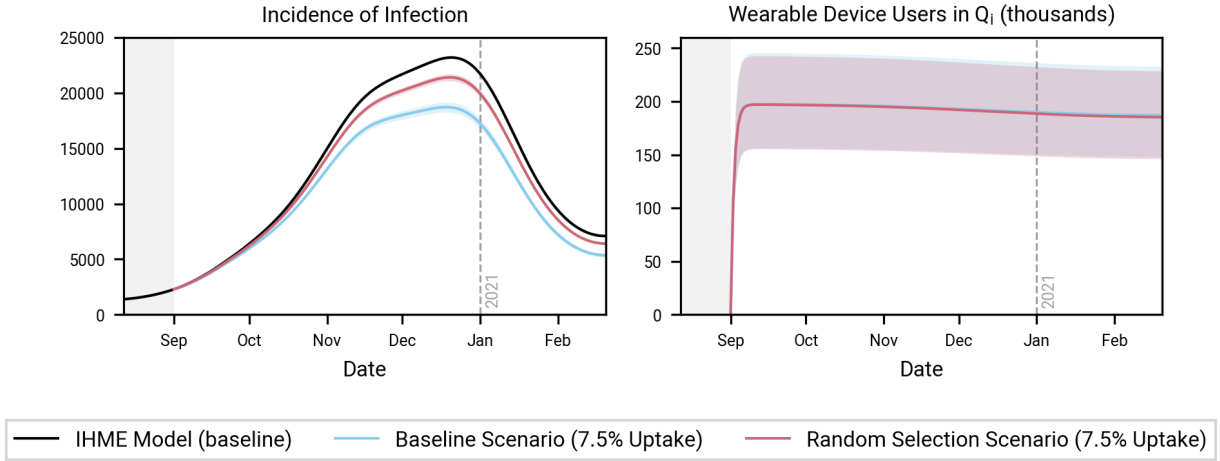
Supplementary Figure 5: Impact of antigen test sensitivity. We assumed 4% uptake and 50% adherence. The dashed grey line represents nominal antigen test sensitivity (91.7%).

Fifth, we explored the impact of lab-based test turnaround time to determine whether minimizing this variable should be a policy priority (Supplementary Figure 6). Throughout this study, we set turnaround time to its nominal value of two days. With longer turnaround times, individuals incorrectly in quarantine would remain there longer, further decreasing the pool of *Susceptible* individuals. As expected, more infections are consequently averted. However, there are other mechanisms (e.g., improving detection sensitivity) for increasing averted infections that are not accompanied by as large of an increase to social costs. Thus, as is already the case outside the context of wearable sensor deployment, it would make sense to minimize turnaround time.



Supplementary Figure 6: Impact of lab-based test turnaround time. We assumed 4% uptake and 50% adherence.

Finally, to understand the impact of detection algorithm sensitivity on the incidence of infection, we investigated the scenario in which device users are randomly asked to seek a lab-based test and self-isolate while awaiting the result (Supplementary Figure 7). We set the likelihood of being asked to quarantine to the false positive rate, still treating detection specificity as a beta-distributed random variable and setting it to its nominal value. As expected, the reduction in detection sensitivity substantially reduces the number of averted infections. The number of individuals incorrectly in quarantine is unimpacted because it is governed by detection specificity, not sensitivity.



Supplementary Figure 7: Impact of randomly asking device users to seek a lab-based test and quarantine while waiting for the result. We assumed 7.5% uptake, still with 50% adherence, to illustrate the relative difference in the incidence of infection.

6. References

1. Li, R. *et al.* Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). *Science* **368**, 489–493 (2020).
2. Lauer, S. A. *et al.* The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application. *Ann. Intern. Med.* **172**, 577–582 (2020).
3. Zhou, R. *et al.* Viral dynamics in asymptomatic patients with COVID-19. *Int. J. Infect. Dis.* **96**, 288–290 (2020).
4. Oran, D. P. & Topol, E. J. Prevalence of Asymptomatic SARS-CoV-2 Infection: A Narrative Review. *Ann. Intern. Med.* **173**, 362–367 (2020).
5. Paré, G., Leaver, C. & Bourget, C. Diffusion of the Digital Health Self-Tracking Movement in Canada: Results of a National Survey. *J. Med. Internet Res.* **20**, e177 (2018).
6. Sujay Vailshery, L. *Activity fitness tracker household penetration rate in Canada in 2018, by province*. https://www.statista.com/statistics/1036752/canada-activity-fitness-tracker-ownership-rate-by-province/ (2021).
7. Munzert, S., Selb, P., Gohdes, A., Stoetzer, L. F. & Lowe, W. Tracking and promoting the usage of a COVID-19 contact tracing app. *Nat. Hum. Behav.* **5**, 247–255 (2021).
8. Silva de Lima, A. L. *et al.* Feasibility of large-scale deployment of multiple wearable sensors in Parkinson's disease. *PLOS ONE* **12**, e0189161 (2017).
9. Radin, J. M., Wineinger, N. E., Topol, E. J. & Steinhubl, S. R. Harnessing wearable device data to improve state-level real-time surveillance of influenza-like illness in the USA: a population-based study. *Lancet Digit. Health* **2**, e85–e93 (2020).
10. Trabelsi Haddad, T. Returning to school, again: 'Operation Antigen 2' is underway, and only negatives will be admitted to institutions. *ynet* (2021).
11. Trabelsi Haddad, T. 613,000 reported, about 550 positive: Antigen results - and those that will arrive tomorrow without testing. *ynet* (2021).
12. Carlsen, E. Ø., Caspersen, I. H., Trogstad, L., Gjessing, H. K. & Magnus, P. *Public adherence to governmental recommendations regarding quarantine and testing for COVID-19 in two Norwegian cohorts*. http://medrxiv.org/lookup/doi/10.1101/2020.12.18.20248405 (2020) doi:10.1101/2020.12.18.20248405.
13. Government of Canada. *How many people are using COVID Alert*. https://www.canada.ca/en/public-health/services/diseases/coronavirus-disease-covid-19/covid-alert.html#a6.

14. Alavi, A. *et al.* Real-time alerting system for COVID-19 and other stress events using wearable data. *Nat. Med.* (2021) doi:10.1038/s41591-021-01593-2.

15. Gadaleta, M. *et al.* Passive detection of COVID-19 with wearable sensors and explainable machine learning algorithms. *Npj Digit. Med.* **4**, 166 (2021).

16. Abbot. *In vitro diagnostic rapid test for qualitative detection of SARS-CoV-2 antigen (Ag)*. https://www.globalpointofcare.abbott/en/product-details/panbio-covid-19-ag-antigen-test.html.

17. Torres, I., Poujois, S., Albert, E., Colomina, J. & Navarro, D. Evaluation of a rapid antigen test (Panbio™ COVID-19 Ag rapid test device) for SARS-CoV-2 detection in asymptomatic close contacts of COVID-19 patients. *Clin. Microbiol. Infect.* **27**, 636.e1-636.e4 (2021).

18. Winkel, B. *et al.* Screening for SARS-CoV-2 infection in asymptomatic individuals using the Panbio COVID-19 antigen rapid test (Abbott) compared with RT-PCR: a prospective cohort study. *BMJ Open* **11**, e048206 (2021).

19. Woloshin, S., Patel, N. & Kesselheim, A. S. False Negative Tests for SARS-CoV-2 Infection — Challenges and Implications. *N. Engl. J. Med.* **383**, e38 (2020).

20. Government of Canada. *Testing for COVID-19: How and where we test for active infections*. https://www.canada.ca/en/public-health/services/diseases/2019-novel-coronavirus-infection/symptoms/testing/diagnosing.html (2021).

21. Cleary, J. L., Fang, Y., Sen, S. & Wu, Z. *A Caveat to Using Wearable Sensor Data for COVID-19 Detection: The Role of Behavioral Change after Receipt of Test Results*. http://medrxiv.org/lookup/doi/10.1101/2021.04.17.21255513 (2021) doi:10.1101/2021.04.17.21255513.

22. Canadian Institute for Health Information. *COVID-19 Hospitalization and Emergency Department Statistics, 2021–2022*. (2021).

23. Institute for Health Metrics and Evaluation (IHME). *COVID-19 Projections*. https://covid19.healthdata.org/projections (2020).

24. World Health Organization. Global partnership to make available 120 million affordable, quality COVID-19 rapid tests for low- and middle-income countries. (2020).

25. Kliff, S. Most Coronavirus Tests Cost About $100. Why Did One Cost $2,315? *The New York Times* (2020).