

1 **Increasing testing throughput and case detection with a pooled-sample Bayesian**
2 **approach in the context of COVID-19**

3 Rodrigo Noriega^{1*}, Matthew H. Samore^{2,3}.

4 ¹Department of Chemistry, University of Utah, Salt Lake City, UT USA.

5 ²Division of Epidemiology, University of Utah, Salt Lake City, UT USA.

6 ³Division of Epidemiology, Veterans Affairs Salt Lake City Health Care System, Salt Lake City,
7 Utah

8 *Correspondence to: noriega@chem.utah.edu

9 **Abstract:** Rapid and widespread implementation of infectious disease surveillance is a critical
10 component in the response to novel health threats. Molecular assays are the preferred method to
11 detect a broad range of pathogens with high sensitivity and specificity. The implementation of
12 molecular assay testing in a rapidly evolving public health emergency can be hindered by resource
13 availability or technical constraints. In the context of the COVID-19 pandemic, the applicability of
14 a pooled-sample testing protocol to screen large populations more rapidly and with limited
15 resources is discussed. A Bayesian inference analysis in which hierarchical testing stages can have
16 different sensitivities is implemented and benchmarked against early COVID-19 testing data.
17 Optimal pool size and increases in throughput and case detection are calculated as a function of
18 disease prevalence. Even for moderate losses in test sensitivity upon pooling, substantial increases
19 in testing throughput and detection efficiency are predicted, suggesting that sample pooling is a
20 viable avenue to circumvent current testing bottlenecks for COVID-19.

21

22 Emerging infectious diseases pose a global hazard to public health, as exemplified by the COVID-
23 19 pandemic. Key epidemiologic strategies for control of community spread include contact tracing,
24 case isolation, ring containment, and social distancing (1–7). The use of microbiological testing to
25 identify disease cases is a crucially important element of these strategies. Some countries, including
26 the US, experienced a shortage of kits needed for COVID-19 diagnosis, which resulted in the
27 imposition of restrictive criteria to manage the selection of patients for testing. Constraints in the
28 supply of kits had a particularly significant impact on testing of mildly symptomatic individuals, as
29 well as asymptomatic contacts of confirmed cases. For some facilities that have been overwhelmed
30 by demand for testing as the pandemic progressed, test throughput continues to be a limiting factor
31 (8–10). Strategies for screening more individuals with a reduced burden on resources are highly
32 desirable. Using a Bayesian formalism, a hierarchical testing protocol based on sample pooling is
33 discussed. Anticipated benefits include easing the demand of constrained resources and enabling
34 more efficient detection of a larger number of cases.

35 Molecular assays are the predominant testing method for viral and bacterial pathogens (11–14).
36 Specifically, nucleic acid detection assays typically employ real-time polymerase chain reaction
37 (RT-PCR) for DNA targets and reverse-transcription real-time PCR (rRT-PCT) for RNA targets
38 (15, 16). The popularity of such testing platforms is due to **1)** their high sensitivity and specificity,
39 **2)** the widespread access to sequencing and synthesis technologies for the identification of nucleic
40 acid target sequences and probes, and **3)** the development of fast, user-friendly, and cost-effective
41 equipment. While nucleic acid assays have powered a revolution in diagnostics and delivery of care
42 for individual patients, their application in large-scale infectious disease surveillance is hampered
43 partly by low throughput at a population level.

44 The information content of a diagnostic test can be evaluated with a Bayesian probability formalism
45 in the context of an individual sample or for repeated sampling from the same patient (17–19) by

46 taking into account the probability of detecting a positive case (assay sensitivity, or identification
47 rate P_{id}) and the probability of a positive result from healthy samples (false positive rate P_{fp}).
48 Bayesian inference requires the assessment of a “prior” probability to the presence of disease in a
49 sample, $P(D)$, which is updated to a “posterior” probability given a positive or negative test result
50 with conditional probabilities $P(D|+)$ or $P(D|-)$ (**Eq. 1.a-b**).

$$51 \quad P(D|+) = \frac{P(D) \cdot P(+|D)}{P(+)} = \frac{P(D) \cdot P_{id}}{P(+)} \quad (\text{Eq. 1a}) \quad P(D|-) = \frac{P(D) \cdot P(-|D)}{P(-)} = \frac{P(D) \cdot (1 - P_{id})}{1 - P(+)} \quad (\text{Eq. 1b})$$

52 where $P(+)$ and $P(-)$ are the overall probabilities of the test yielding a positive or negative result,
53 respectively. These tools can be extended to the somewhat counterintuitive situation where a
54 diagnostic test is conducted on a sample pooled from multiple individuals. The motivation for
55 sample pooling is to screen multiple patients simultaneously and reduce the burden on testing
56 facilities working with limited resources. Pooling schemes have been developed since their
57 introduction in the 1940s for syphilis testing, and have been applied to screen for and estimate
58 prevalence rates of a variety of diseases (20-24). Here, a simple two-step hierarchical protocol first
59 introduced by Dorfman is considered: *Samples from N_p patients are collected and randomly pooled*
60 *into groups of n individual samples each. Pooled samples are interrogated with the diagnostic test.*
61 *If pooled testing yields a negative result, no further testing is conducted. If pooled testing yields a*
62 *positive result, all patients in that pool are tested individually.*

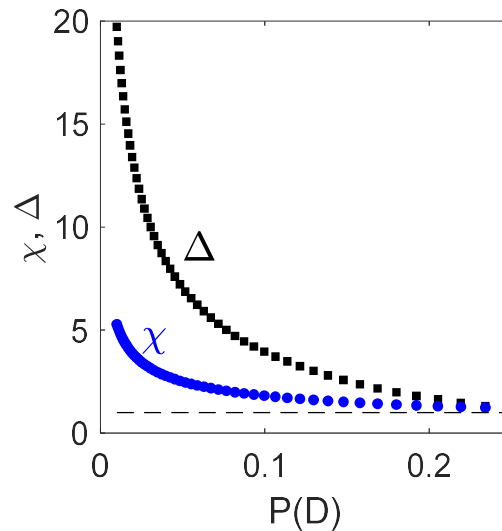
63 While statistical approaches have been focused on characterizing the performance of various
64 pooling schemes, not all of them include non-ideal test parameters (25). Moreover, testing
65 characteristics at the pooled- and individual-sample levels can be different. Here, the Bayesian
66 inference approach in **Eq. 1.a-b** is modified to include differences in the assay sensitivity and
67 overall probability of a positive result in pooled vs. individual tests (**Eq. S1-2**). Sensitivity loss is
68 included as a reduction in the identification rate of pooled-sample tests by a scaling factor γ . Due to
69 the exceptional specificities of nucleic acid assays, the false positive rate is assumed to remain

70 unaffected. Importantly, the posterior probability assessed from a positive pooled test is used as a
71 prior for follow-up individual tests, which yields additional information and enhances the Bayesian
72 inference assessment of those cases (**Eq. S3**).

73 For a population of N_p patients divided into sample pools of size n , the average number of tests is
74 the number of initial pooled tests plus the expected number of follow-up tests (**Eq. S4**). Throughput
75 increase χ is expressed as the effective number of individuals screened by each diagnostic test (**Eq.**
76 **S5**). The individual- and pooled-sample test characteristics determine the pool size that optimizes
77 screening throughput as a function of average disease prevalence in the tested population.

78 The advantages of pooled-sample screening are discussed in the context of the rapidly evolving
79 COVID-19 pandemic (caused by the novel coronavirus SARS-CoV-2) (8, 26). A recent rRT-PCR
80 assay for COVID-19 reports an identification rate of 95% and no false positives after testing 310
81 samples including other respiratory pathogens (27). Given reported specificities of commercially-
82 available respiratory panel assays (>99%), an estimated 1% false positive rate was included in the
83 model results reported here. A moderate reduction in the identification rate for a pooled sample
84 ($\gamma = 0.9$) was assumed – this variable is discussed in detail below. Consistent with similar
85 implementations of Dorfman-type testing algorithms (28), substantial increases in testing
86 throughput are predicted for low disease prevalence rates ($P(D) \leq 8\%$, **Fig. 1**), where throughput
87 more than doubles and optimal pool sizes are $4 \leq n \leq 12$. At intermediate prevalence rates $0.1 \leq$
88 $P(D) \leq 0.2$, the increase in throughput is moderate yet substantial (>30% increase in throughput)
89 and pool sizes are small ($n = 3$). For high prevalence rates, pooling yields no improvement
90 (optimal pool size $n = 1$). Average disease prevalence can be re-assessed as information is gained
91 for the tested population to re-optimize pool size. Dynamic self-tuning is a feature of Bayesian
92 inference, a significant asset when a close feedback loop is desirable.

93 Besides testing throughput, it is informative to assess the increased ability to detect cases using a
94 pooled-sample vs individual-sample scheme, which can be accomplished by comparing the ratio of
95 detected to missed cases in each protocol. Importantly, resource constraints are incorporated into
96 this comparison by accounting for unscreened cases in the standard 1:1 scheme (**Eq. S6-8**). The
97 relative increase in the detection-to-miss ratio between the pooled and standard 1:1 schemes
98 exhibits even more significant gains than those observed for testing throughput (**Fig. 1**).



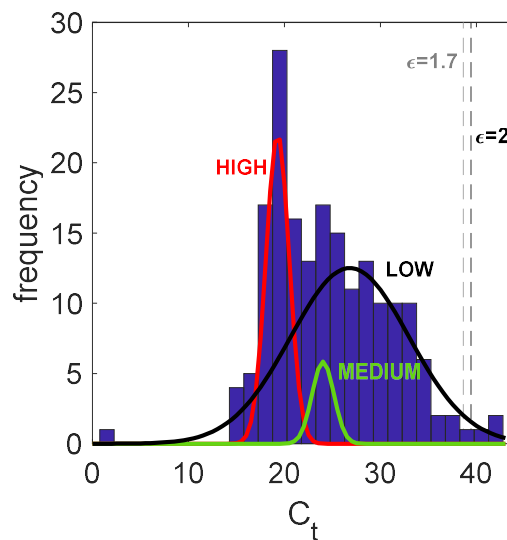
99

100 **Fig. 1.** Number of patients screened by each diagnostic test (χ , blue circles) and the relative increase
101 in detection-to-miss ratio for a pooled scheme (Δ , black squares), as a function of average disease
102 prevalence. Gray line is the no-pooling scheme reference.

103 Loss in screening power associated with sample pooling is assessed by **1**) estimating the portion of
104 cases that would have been identified in an individual test but missed by pooled screening, and **2**)
105 determining the information gained for patients whose pooled screening result was negative.

106 A reduction in overall pathogen concentration due to pooling in conditions of low disease
107 prevalence can decrease the test's identification rate, although it is manageable when targeting
108 infectious diseases for which typical pathogen concentrations are non-negligible (23). This concern
109 is examined with a set of 186 positive rRT-PCR diagnostic test results for COVID-19 (**Fig. 2**),

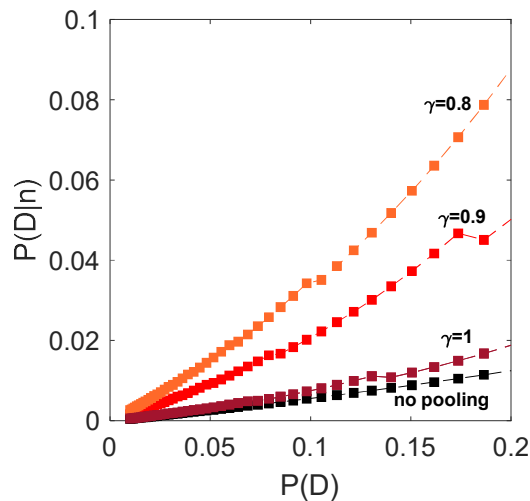
110 which include nasopharyngeal swab, oropharyngeal swab, and bronchoalveolar lavage (29). Tests
111 have a mean cycle threshold value of $\langle C_t \rangle = 24.6$, and a positive test result is defined as a rRT-PCR
112 reaction with $C_t \leq 42$ – in agreement with early reports for nasal swab COVID-19 rRT-PCR tests
113 with $\langle C_t \rangle = 24.3$ and viral loads of 1.4×10^6 copies/mL (30). From these data, the portion of
114 samples that would have had a positive test result even if pooled with an entirely healthy population
115 is estimated using a pool size $n = 12$ and rRT-PCR geometric efficiencies of $\epsilon = 1.7 - 2$ per
116 cycle. A pooled screening protocol would have detected 95.7-96.8% of these cases (178-180 out of
117 186). Further support for moderate sensitivity loss can be achieved by dividing the distribution of C_t
118 values for the test-positive samples into three subpopulations. For the same pooling and rRT-PCR
119 efficiencies stated previously, >95% of the broadest, lowest-load population would be detected.



120
121 **Fig. 2.** Cycle threshold values for rRT-PCR reactions for confirmed COVID-19 cases (27),
122 described by subpopulations with low, medium, and high viral load. Sensitivity cutoffs shown as
123 vertical lines.

124 The main benefit of hierarchical pooled-sample testing protocols is the ability to screen a larger
125 portion of the population and detect more positive cases. The relative increase in case detection is
126 given by the ratio of the number of cases detected in a pooled setting and the number of cases

127 detected if the same number of tests were used on single patients – the multiplicative factor in
128 testing throughput scaled by the loss in sensitivity, $\chi \cdot \gamma$. Loss in sensitivity also leads to more
129 screened-yet-missed cases – even in the absence of pooling, diagnostic tests yield false negatives
130 and increasing those odds should not be considered lightly. In the range of disease prevalence with
131 moderate to high increase in throughput and compared to individual testing of every patient, a $\gamma =$
132 0.9 sensitivity loss leads to a relatively low increase in posterior probability of disease after a
133 negative test outcome, $P(D|–)$ (**Fig. 3**). While sensitivity losses decrease the ability to confidently
134 screen healthy individuals, the threshold for this tradeoff depends on the situation where the
135 screening protocol is deployed. Effective risk communication to screened individuals is needed to
136 prevent an outsized sense of security after a negative pooled-test result – e.g., vigilance to symptom
137 development triggers individual testing.



138
139 **Fig. 3.** Probability of an individual being infected even though their pooled-sample test gave a
140 negative result, as a function of background disease prevalence and for different values of
141 sensitivity loss.

142 This protocol is amenable to HIPAA regulations – in fact, pooling has been implemented by state
143 laboratories in the recent past (31)– and requires limited additional sample processing. Refining the
144 dependence of sensitivity loss with pool size and coupling to modeling strategies that inform

145 population sampling and prior probabilities will provide further screening improvements. From an
146 epidemiological surveillance standpoint, increased detection of positive cases in a larger portion of
147 the population denotes that a greater fraction of infectious individuals can be isolated. However,
148 pooled testing is less useful in an inpatient clinical setting where the highest sensitivity is needed to
149 minimize risk of hospital transmission from non-isolated patients. As with any change in the
150 delivery of medical care, a discussion including community stakeholders is paramount.

151 In summary, a pooled testing strategy has the potential to enhance comprehensive surveillance of
152 SARS-CoV-2 particularly when test kits are in short supply. The benefits of surveillance are
153 greatest in the early phases of community spread. Thus, improving the capacity for high-throughput
154 testing has the highest impact when prevalence is low enough that pooled sampling is most
155 beneficial. The ratio of confirmed COVID-19 cases to tests performed varies by country, but it
156 appears that aggressive testing strategies yield a low enough prevalence to benefit from pooled-
157 sample screening – e.g., South Korea’s is 3% (8.3k/270k as of 3/16/2020) (32). While the
158 development of clinical prediction rules and non-testing screening are critical to any
159 epidemiological response, dealing with a novel disease for which data is still sparse and testing
160 capabilities are limited means that maximizing the impact of each individual test can benefit the
161 continued refinement of our strategy.

162

163 **Acknowledgments:** Valuable discussions with Dr. Kimberly Hanson and Dr. Lindsay T. Keegan.

164 **Funding:** none; **Author contributions:** R.N. conceived the idea, performed the analysis, prepared
165 the figures, and drafted the manuscript. R.N. and M.H.S. developed the idea and edited the

166 manuscript; **Competing interests:** Authors declare no competing interests; **Data and materials**

167 **availability:** All data is available in the main text or the supplementary materials.

168 **References and Notes:**

- 169 1. C. Wells, D. Yamin, M. L. Ndeffo-Mbah, N. Wenzel, S. G. Gaffney, J. P. Townsend, L. A.
170 Meyers, M. Fallah, T. G. Nyenswah, F. L. Altice, K. E. Atkins, A. P. Galvani, Harnessing
171 Case Isolation and Ring Vaccination to Control Ebola. *PLOS Neglected Tropical Diseases*. **9**,
172 e0003794 (2015).
- 173 2. G. Chowell, S. Echevarría-Zuno, C. Viboud, L. Simonsen, J. Tamerius, M. A. Miller, V. H.
174 Borja-Aburto, Characterizing the Epidemiology of the 2009 Influenza A/H1N1 Pandemic in
175 Mexico. *PLoS Med.* **8** (2011), doi:10.1371/journal.pmed.1000436.
- 176 3. E. P. Fenichel, C. Castillo-Chavez, M. G. Ceddia, G. Chowell, P. A. G. Parra, G. J. Hickling,
177 G. Holloway, R. Horan, B. Morin, C. Perrings, M. Springborn, L. Velazquez, C. Villalobos,
178 Adaptive human behavior in epidemiological models. *PNAS*. **108**, 6306–6311 (2011).
- 179 4. R. M. May, A. R. McLean, J. Pattison, R. A. Weiss, R. M. Anderson, C. Fraser, A. C. Ghani,
180 C. A. Donnelly, S. Riley, N. M. Ferguson, G. M. Leung, T. H. Lam, A. J. Hedley,
181 Epidemiology, transmission dynamics and control of SARS: the 2002–2003 epidemic.
182 *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*. **359**,
183 1091–1105 (2004).
- 184 5. D. Klinkenberg, C. Fraser, H. Heesterbeek, The Effectiveness of Contact Tracing in Emerging
185 Epidemics. *PLOS ONE*. **1**, e12 (2006).
- 186 6. R. D. Smith, Responding to global infectious disease outbreaks: Lessons from SARS on the
187 role of risk perception, communication and management. *Social Science & Medicine*. **63**,
188 3113–3123 (2006).

- 189 7. S. Cauchemez, N. M. Ferguson, C. Wachtel, A. Tegnell, G. Saour, B. Duncan, A. Nicoll,
190 Closure of schools during an influenza pandemic. *The Lancet Infectious Diseases*. **9**, 473–481
191 (2009).
- 192 8. C. del Rio, P. N. Malani, COVID-19—New Insights on a Rapidly Changing Epidemic. *JAMA*
193 (2020), doi:10.1001/jama.2020.3072.
- 194 9. M. Lin, A. Beliavsky, K. Katz, J. E. Powis, W. Ng, V. Williams, M. Science, H. Groves, M. P.
195 Muller, A. Vaisman, S. Hota, J. Johnstone, J. A. Leis, What can early Canadian experience
196 screening for COVID-19 teach us about how to prepare for a pandemic? *CMAJ* (2020),
197 doi:10.1503/cmaj.200305.
- 198 10. Y. Bai, L. Yao, T. Wei, F. Tian, D.-Y. Jin, L. Chen, M. Wang, Presumed Asymptomatic
199 Carrier Transmission of COVID-19. *JAMA* (2020), doi:10.1001/jama.2020.2565.
- 200 11. S-S. Chang, W-H. Hsieh, T-S. Kiu, S-H. Lee, C-H. Wang, H-C. Chou, Y. H. Yeo, C-P. Tseng,
201 C-C. Lee, Multiplex PCR System for Rapid Detection of Pathogens in Patients with Presumed
202 Sepsis – A Systemic Review and Meta-Analysis, *PLOS ONE*. **8**(5), e62323 (2013).
- 203 12. S. C. Kehl, S. Kumar, Utilization of Nucleic Acid Amplification Assays for the Detection of
204 Respiratory Viruses. *Clinics in Laboratory Medicine*. **29**, 661–671 (2009).
- 205 13. J. B. Mahony, Detection of Respiratory Viruses by Molecular Methods. *Clinical Microbiology*
206 *Reviews*. **21**, 716–747 (2008).
- 207 14. W. Wu, Y.-W. Tang, Emerging Molecular Assays for Detection and Characterization of
208 Respiratory Viruses. *Clinics in Laboratory Medicine*. **29**, 673–693 (2009).

- 209 15. M. J. Espy, J. R. Uhl, L. M. Sloan, S. P. Buckwalter, M. F. Jones, E. A. Vetter, J. D. C. Yao,
210 N. L. Wengenack, J. E. Rosenblatt, F. R. Cockerill, T. F. Smith, Real-Time PCR in Clinical
211 Microbiology: Applications for Routine Laboratory Testing. *Clinical Microbiology Reviews*.
212 **19**, 165–256 (2006).
- 213 16. I. M. Mackay, K. E. Arden, A. Nitsche, Real-time PCR in virology. *Nucleic Acids Res.* **30**,
214 1292–1305 (2002).
- 215 17. H. Sox, S. Stern, D. Owens, H. L. Abrams, “The use of diagnostic tests: A probabilistic
216 approach” in “*Assessment of Diagnostic Technology in Health Care: Rationale, Methods,*
217 *Problems, and Directions*” (National Academies Press (US), 1989).
- 218 18. R. S. Ledley, L. B. Lusted, Reasoning Foundations of Medical Diagnosis: Symbolic logic,
219 probability, and value theory aid our understanding of how physicians reason. *Science*. **130**, 9–
220 21 (1959).
- 221 19. H. C. Sox, Diagnostic Decision: Probability Theory in the Use of Diagnostic Tests: An
222 Introduction to Critical Study of the Literature. *Ann Intern Med.* **104**, 60 (1986).
- 223 20. R. Dorfman, The detection of defective numbers of large populations. *Annals of Mathematical*
224 *Statistics* **14**, 436–440 (1943).
- 225 21. C. Pilcher, S. Fiscus, T. Nguyen, E. Foust, L. Wolf, D. Williams, R. Ashby, J. O'Dowd, J.
226 McPherson, B. Stalzer, L. Hightow, W. Miller, J. Eron, M. Cohen, P. Leone, Detection of
227 acute infections during HIV testing in North Carolina. *New England Journal of Medicine* **352**,
228 1873–1883 (2005).
- 229 22. C. Lindan, M. Mathur, S. Kumta, H. Jerajani, A. Gogate, J. Schachter, J. Moncada, Utility of
230 pooled urine specimens for detection of *Chlamydia trachomatis* and *Neisseria gonorrhoeae* in

- 231 men attending public sexually transmitted infection clinics in Mumbai, India, by PCR. *Journal*
232 *of Clinical Microbiology* **43**, 1674–1677 (2005).
- 233 23. T. Van, J. Miller, D. Warshauer, E. Reisdorf, D. Jerrigan D, Humes R, Shult P., Pooling
234 nasopharyngeal/throat swab specimens to increase testing capacity for influenza viruses by
235 PCR. *Journal of Clinical Microbiology* **50**, 891–896 (2012).
- 236 24. M. S. Warasi, J. M. Tebbs, C. S. McMahan, C. R. Bilder, Estimating the prevalence of multiple
237 diseases from two-stage hierarchical pooling. *Statist. Med.* **35**, 3851– 3864 (2016).
- 238 25. H.-Y. Kim, M. G. Hudgens, J. M. Dreyfuss, D. J. Westreich, C. D. Pilcher, Comparison of
239 Group Testing Algorithms for Case Identification in the Presence of Test Error. *Biometrics* **63**,
240 1152-1163 (2007).
- 241 26. S. P. Layne, J. M. Hyman, D. M. Morens, J. K. Taubenberger, New coronavirus outbreak:
242 Framing questions for pandemic prevention. *Science Translational Medicine*. **12** (2020),
243 doi:10.1126/scitranslmed.abb1469.
- 244 27. V. M. Corman, O. Landt, M. Kaiser, R. Molenkamp, A. Meijer, D. K. Chu, T. Bleicker, S.
245 Brünink, J. Schneider, M. L. Schmidt, D. G. Mulders, B. L. Haagmans, B. van der Veer, S.
246 van den Brink, L. Wijsman, G. Goderski, J.-L. Romette, J. Ellis, M. Zambon, M. Peiris, H.
247 Goossens, C. Reusken, M. P. Koopmans, C. Drosten, Detection of 2019 novel coronavirus
248 (2019-nCoV) by real-time RT-PCR. *Eurosurveillance*. **25**, p. 2000045 (2020).
- 249 28. S. M. Samuels, The exact solution to the two-stage group-testing problem. *Technometrics* **20**,
250 497-500 (1978).

- 251 29. K. Hanson, ARUP COVID-19 rRT-PCR diagnostic testing results. Testing method is the EUA
252 approved Hologic Panther Fusion SARS-COV2 assay:
253 <https://www.fda.gov/media/136156/download>
- 254 30. W. Wang, Y. Xu, R. Gao, R. Lu, K. Han, G. Wu, W. Tan, Detection of SARS-CoV-2 in
255 Different Types of Clinical Specimens. *JAMA* (2020), doi:10.1001/jama.2020.3786.
- 256 31. J. L. Lewis, V. M. Lockary, S. Kobic, Cost Savings and Increased Efficiency Using a
257 Stratified Specimen Pooling Strategy for Chlamydia trachomatis and Neisseria gonorrhoeae.
258 *Sexually Transmitted Diseases* **39**, 46-48 (2012).
- 259 32. J. Cohen, K. Kupferschmidt, Countries test tactics in ‘war’ against COVID-19. *Science*. **367**,
260 1287–1288 (2020).

261 **Materials and Methods**

262 Bayesian inference implementation

263 The process for assessing the posterior probabilities for individuals whose pooled test yielded a
 264 positive or negative result (**Eq. S1.a-b**)

265
$$P(D|+) = \frac{P(D) \cdot P_{id}^{pool}}{P_{pool}(+)} \quad (\text{Eq. S1.a}) \qquad P(D|-) = \frac{P(D) \cdot (1 - P_{id}^{pool})}{1 - P_{pool}(+)} \quad (\text{Eq. S1.b})$$

266 includes the a reduction in the identification rate of the pooled-sample diagnostic test by a factor γ
 267 compared to an individual test, so that $P_{id}^{pool} = \gamma \cdot P_{id}^{ind}$. One must also consider the probability of a
 268 positive result in a pooled sample. Assuming that every individual has a prior probability of being
 269 infected equal to some background average $P(D)$, the probability of a positive test result in a pool
 270 of size n is the probability of having a nonzero number of positive individual samples times the
 271 pooled-test identification rate plus the probability of a completely-healthy sample pool yielding a
 272 false positive (**Eq. S2**):

273
$$P_{pool}(+) = [1 - (1 - P(D))^n] \cdot P_{id}^{pool} + P_{fp} \cdot (1 - P(D))^n \quad (\text{Eq. S2})$$

274 The posterior probability $P(D|+)$ from the pooled test with a positive outcome can be used as a
 275 prior for the follow-up individual test Bayesian inference (**Eq. S3.a-b**).

276
$$P(D|+ +) = \frac{P(D|+) \cdot P_{id}}{P_{ind}(+)} = \frac{P(D|+) \cdot P_{id}}{P(D|+) \cdot P_{id} + (1 - P(D|+)) \cdot P_{fp}} \quad (\text{Eq. S3.a})$$

277

278
$$P(D|+ -) = \frac{P(D|+) \cdot (1 - P_{id})}{1 - P_{ind}(+)} \quad (\text{Eq. S3.b})$$

279 where the overall probability of an individual's test being positive includes the probability of
 280 correctly identifying a positive sample $P(D|+) \cdot P_{id}$ plus the probability of a negative sample
 281 yielding a false positive $(1 - P(D|+)) \cdot P_{fp}$.

282 For a population of N_p patients divided into sample pools of size n , the average number of tests
283 $\langle N_{test} \rangle$ is given by the number of initial pooled tests plus the expected number of follow-up tests
284 (Eq. S4):

$$285 \quad \langle N_{test} \rangle = \frac{N_p}{n} + n \cdot \frac{N_p}{n} \cdot P_{pool}(+) = N_p \cdot \left(\frac{1}{n} + P_{pool}(+) \right) \quad (\text{Eq. S4})$$

286 The throughput increase is expressed as a multiplicative factor representing how many individuals
287 were screened by the use of each diagnostic test (Eq. S5).

$$288 \quad \chi = \frac{N_p}{\langle N_{test} \rangle} = \left(\frac{1}{n} + P_{pool}(+) \right)^{-1} \quad (\text{Eq. S5})$$

289
290 Full vs. partial optimization. In a fully optimized scheme a pooled-sample's positive result triggers
291 a cascade of smaller-sized pools, minimizing the number of tests performed. However, an
292 intermediate improvement is chosen due to possible adverse outcomes of a lengthier process with
293 several pooling and Bayesian inference steps – e.g., possible delay of necessary care, increased
294 exposure to infected individuals.

295
296 Advantages in testing throughput and case detection

297 We can estimate the ratio of detected-to-missed (including unscreened) cases for each protocol
298 with the following simplified analysis:

299 **In the pooled scheme,** the number of detected cases are the total number of screened individuals
300 with the disease $\langle N_{test} \rangle \chi P(D)$ multiplied by the effective detection probability $\gamma \cdot (P_{id}^{ind})^2$, which
301 includes the sequential detection probability at the pooled stage and the follow-up individual stage.
302 The number of missed cases would thus be the difference between detected and total cases,
303 $\langle N_{test} \rangle \chi P(D) \left[1 - \gamma \cdot (P_{id}^{ind})^2 \right]$. The number of false positives is the number of individuals in all-
304 healthy pools yielding a false positive and whose individual test is also a false positive –

305 $\langle N_{test} \rangle \chi (1 - P(D))^n P_{fp}^2$ – plus the healthy individuals pooled with a diseased sample that is
 306 detected at the pooled stage and whose follow up test is a false positive, $\langle N_{test} \rangle \chi \gamma P_{id}^{ind} [1 -$
 307 $(1 - P(D))^n - P(D)] P_{fp}$. The detection-to-miss ratio θ_{pool} can be expressed as

$$308 \quad \theta_{pool} = \frac{\langle N_{test} \rangle \chi P(D) \gamma \cdot (P_{id}^{ind})^2}{\langle N_{test} \rangle \chi P(D) [1 - \gamma \cdot (P_{id}^{ind})^2]} = \frac{\gamma \cdot (P_{id}^{ind})^2}{1 - \gamma \cdot (P_{id}^{ind})^2} \quad \text{(Eq. S6)}$$

309 **In a standard 1:1 scheme**, the number of detected cases are the number of tested individuals with
 310 the disease $\langle N_{test} \rangle P(D)$ times the detection efficiency of the individual test P_{id}^{ind} . The number of
 311 missed cases is thus $\langle N_{test} \rangle P(D) (1 - P_{id}^{ind})$. The number of false positives is $\langle N_{test} \rangle [1 -$
 312 $P(D)] P_{fp}$. The number of unscreened cases that carry the disease is $\langle N_{test} \rangle (\chi - 1) P(D)$. The
 313 detection-to-miss ratio – considering the population sampled in the pooled scheme – is given by

$$314 \quad \theta_{ind} = \frac{\langle N_{test} \rangle P(D) P_{id}^{ind}}{\langle N_{test} \rangle P(D) (1 - P_{id}^{ind}) + \langle N_{test} \rangle (\chi - 1) P(D)} = \frac{P_{id}^{ind}}{\chi - P_{id}^{ind}} \quad \text{(Eq. S7)}$$

315 The relative increase in the detection-to-miss ratio between the pooled and standard 1:1 schemes is
 316 thus given by

$$317 \quad \Delta = \frac{\theta_{pool}}{\theta_{ind}} = \frac{\gamma P_{id}^{ind} (\chi - P_{id}^{ind})}{1 - \gamma \cdot (P_{id}^{ind})^2} \quad \text{(Eq. S8)}$$

318 To further illustrate this, consider a disease including asymptomatic-yet-infectious individuals
 319 and for which clinical predictions are at an early stage, preventing effective triage against conditions
 320 presenting similar symptoms. A screening point must decide how to use limited resources (e.g.,
 321 4,000 available tests) to detect the maximum number of cases in a population at 5% risk. Using the
 322 test characteristics described in the main text ($P_{id}^{ind} = 0.95, P_{fp} = 0.01, \gamma = 0.9$), it is possible to
 323 compare the pooled-sample screening protocol vs. standard 1:1: testing. On average, a pooled-
 324 sample approach allows testing 10,000 individuals, detecting 406 cases while missing 94 and
 325 yielding 18 false positives; conversely, a standard 1:1 approach would test 4,000 individuals and

326 detect 190 cases, miss 10, yield 38 false positives, and leave 300 positive cases untested and thus
327 undetected.

328 What if the rRT-PCR test sensitivity is much lower than expected – say $P_{id}^{ind} = 0.70$? Let us
329 also consider a much more pessimistic estimate for sensitivity loss, $\gamma = 0.8$ and no change in the
330 false positive rate ($P_{fp} = 0.01$). For the same 5% prevalence in the population, the screening point
331 with access to 4,000 tests would screen 12,550 individuals (627 of which are infected) using a
332 pooling scheme (for these new parameters, optimal pool size is $n = 7$). It would detect 246 and
333 miss 381. A 1:1 testing scheme would identify 140 cases, miss 60, and leave 427 untested. Not
334 surprisingly, the number of false negatives increases substantially in both scenarios due to the lower
335 starting point for the test's sensitivity. However, in a pooled scheme 381 infected individuals are
336 still at risk of spreading the disease in the community, while in the 1:1 scheme 487 infected
337 individuals remain at risk of further community spread. As mentioned in the main text, effective
338 risk communication is a critical component of any large-scale screening effort with imperfect tests.
339 Symptomatic individuals should be considered at increased risk even after a negative test result, and
340 other diagnostic avenues could be used (e.g., chest CT).