

1 **Open Development and Clinical Validation**  
2 **Of Multiple 3D-Printed Sample-Collection Swabs:**  
3 **Rapid Resolution of a Critical COVID-19 Testing Bottleneck**

4 Cody J Callahan BS<sup>a</sup>, Rose Lee MD<sup>bc</sup>, Katelyn E. Zulauf PhD<sup>bd</sup>, Lauren Tamburello MPH<sup>e</sup>,  
5 Kenneth P. Smith PhD<sup>bd</sup>, Joe Previtera AS<sup>f</sup>, Annie Cheng BA<sup>b</sup>, Alex Green BA<sup>bd</sup>, Ahmed Abdul  
6 Azim MD<sup>ci</sup>, Amanda Yano MHA<sup>g</sup>, Nancy Doraiswami RN BSN<sup>h</sup>, James E. Kirby MD PhD<sup>bd</sup>, and  
7 Ramy A. Arnaout MD DPhil<sup>bdjt</sup>

8 <sup>a</sup>Department of Radiology, Beth Israel Deaconess Medical Center, Boston, MA, USA 02215

9 <sup>b</sup>Clinical Microbiology Laboratories, Division of Clinical Pathology, Department of Pathology,  
10 Beth Israel Deaconess Medical Center, Boston, MA USA 02215

11 <sup>c</sup>Division of Infectious Disease, Department of Medicine, Beth Israel Deaconess Medical Center,  
12 Boston, MA, USA 02215

13 <sup>d</sup>Harvard Medical School, Boston, MA, USA 02115

14 <sup>e</sup>Division of Urologic Surgery, Department of Surgery, Beth Israel Deaconess Medical Center,  
15 Boston, MA, USA 02215

16 <sup>f</sup>Division of Respiratory Therapy, Beth Israel Deaconess Medical Center, Boston, MA, USA  
17 02215

18 <sup>g</sup>Department of Medicine, Beth Israel Deaconess Medical Center, Boston, MA, USA 02215

19 <sup>h</sup>Division of Perioperative Services, Department of Central Processing, Beth Israel Deaconess  
20 Medical Center, Boston, MA, USA 02215

21 <sup>i</sup>Division of Infection Control/Hospital Epidemiology, Silverman Institute for Healthcare Quality  
22 and Safety, Beth Israel Deaconess Medical Center, Boston, MA, USA 02215

23 <sup>j</sup>Division of Clinical Informatics, Department of Medicine, Beth Israel Deaconess Medical Center,  
24 Boston, MA USA 02215

25 <sup>†</sup>Corresponding author: +1-617-538-5681, [rarnaout@bidmc.harvard.edu](mailto:rarnaout@bidmc.harvard.edu).

26 **Abstract**

27 The SARS-CoV-2 pandemic has caused a severe international shortage of the nasopharyngeal  
28 swabs that are required for collection of optimal specimens, creating a critical bottleneck in the  
29 way of high-sensitivity virological testing for COVID-19. To address this crisis, we designed and  
30 executed an innovative, radically cooperative, rapid-response translational-research program  
31 that brought together healthcare workers, manufacturers, and scientists to emergently develop  
32 and clinically validate new swabs for immediate mass production by 3D printing. We performed  
33 a rigorous multi-step preclinical evaluation on 160 swab designs and 48 materials from 24  
34 companies, laboratories, and individuals, and shared results and other feedback via a public  
35 data repository (<http://github.com/rarnaout/Covidswab/>). We validated four prototypes through  
36 an institutional review board (IRB)-approved clinical trial that involved 276 outpatient volunteers  
37 who presented to our hospital's drive-through testing center with symptoms suspicious for  
38 COVID-19. Each participant was swabbed with a reference swab (the control) and a prototype,  
39 and SARS-CoV-2 reverse-transcriptase polymerase chain reaction (RT-PCR) results were  
40 compared. All prototypes displayed excellent concordance with the control ( $\kappa=0.85-0.89$ ). Cycle-  
41 threshold (Ct) values were not significantly different between each prototype and the control,  
42 supporting the new swabs' non-inferiority (Mann-Whitney U [MWU]  $p>0.05$ ). Study staff  
43 preferred one of the prototypes over the others and the control swab overall. The total time  
44 elapsed between identification of the problem and validation of the first prototype was 22 days.  
45 Contact information for ordering can be found at <http://printedswabs.org>. Our experience holds  
46 lessons for the rapid development, validation, and deployment of new technology for this  
47 pandemic and beyond.

## 48 **Introduction**

49 Since the emergence of the COVID-19 pandemic, more than 2.5 million cases have been  
50 diagnosed worldwide (1). These diagnoses were made using material collected from NP swabs,  
51 which provide the highest sensitivity for detecting SARS-CoV-2 infection during early infection  
52 using commercial RT-PCR-based assays. An NP swab is an FDA Class I exempt medical  
53 device roughly 15 cm in length and 2-3mm in diameter designed to collect secretions from the  
54 posterior nasopharynx (Figs. 1a, left and 1b, top). The head of the swab is generally coated with  
55 short synthetic filaments called flock. The swab is inserted into the nasopharynx, rotated several  
56 times to collect material, and then placed into a vial containing a few milliliters of transport  
57 media. A breakpoint on the shaft enables detachment and release of the head into the vial,  
58 which is then sealed and sent for testing.

59 The rapid spread of SARS-CoV-2 has resulted in severe shortages of NP swabs, due to both  
60 manufacturing stoppages resulting in decreased supply and the spread of the pandemic  
61 resulting in unprecedented demand (2). To address the swab shortage, hospitals and other  
62 testing centers have repurposed other commercially available swabs (e.g. throat, urogenital) to  
63 collect nasal epithelial mucosa for testing (Fig. 1a, second from left and 1b, second from top).  
64 However, such swabs are suboptimal for swabbing the nasopharynx due to differences in size  
65 and flexibility and the possibility they contain PCR-inhibitory materials (3, 4). Material from other  
66 sites has not yet been shown to be able to substitute for swabbing the nasopharynx.

67 One solution to the swab crisis is to design and 3D-print swabs. Advantages of 3D printing  
68 include simplicity (avoiding the multistep process of applying flock), the widespread availability  
69 of 3D printing capacity, and the ability to iterate prototypes rapidly (5). To resolve the swab-  
70 shortage crisis, we have been coordinating an open collaborative process that has brought  
71 together many medical centers, individuals, academic laboratories, and both new and well  
72 established manufacturers (6). As part of this process, we have been testing and continuously  
73 providing feedback on prototype swabs in order to proceed rapidly but safely toward the  
74 development of swabs that can be used clinically, at volumes equal to the need. The openness  
75 of the process was a conscious decision supported by a substantial body of scientific literature,  
76 including the previous experience of the present authors, that demonstrates the advantages of  
77 openness over closed or hybrid approaches (7–9). At our institution, this process has led to an  
78 ongoing clinical trial of several prototype swabs, the first results of which we report here.

## 79 **Materials and Methods**

80 **Process.** We created a public repository using GitHub, a free website most often used by  
81 programmers to co-develop computer code (<http://www.github.com/rarnaout/Covidswab>)<sup>6</sup>. We  
82 provided a clear description of the problem and updated the repository with whatever we  
83 learned and encouraged others to do the same. By tapping our personal and professional  
84 networks, we nucleated an ad hoc network of manufacturers that included companies,  
85 academic groups, and individuals. This grew to include other medical centers interested in  
86 helping develop and test new swabs. These other groups were given the ability to add to the  
87 repository as desired.

88 We devised a three-phase process consisting of preclinical evaluation (Phase I), production  
89 considerations (Phase II), and field testing (Phase III). We described these processes on the  
90 repository for all to see. We took high-resolution photographs of all prototypes and stored Phase  
91 I results in a Microsoft Excel (Microsoft Corporation, Redmond, WA, USA) spreadsheet that  
92 remains publicly available in the repository. All contributors could see each others' designs and  
93 our feedback and iterated accordingly.

94 We made our personal contact information freely available to facilitate communication and  
95 speed the delivery of prototypes. We involved representatives of our institution's nursing, legal,  
96 intellectual property, leadership, purchasing, human resources, communications, and  
97 contracting teams and the institutional review board early and often in order to facilitate open  
98 development, reassign idled staff to our process, and minimize lead times during the rapidly  
99 changing situation.

100 **Phase I: Preclinical evaluation.** *Design.* An infectious disease physician, clinical pathologist  
101 (clinical microbiologist), and respiratory therapist tested each prototype swab for design and  
102 mechanical properties (Fig. 1c-d). These included size measurements of the head, neck, shaft,  
103 and breakpoint (requirement of ~15cm to reach the posterior nasopharynx; head diameter of 1-  
104 3.2mm to pass into the mid-inferior portion of the inferior turbinate and be able maneuver  
105 appropriately without catching on anatomical variants such as septal spurs or a deviated nasal  
106 septum); surface properties such as smoothness (with roughness leading to an unpleasant feel  
107 and risk of bleeding); flexibility vs. brittleness of the head, neck, shaft, and breakpoint (to avoid  
108 fracture during use); durability (e.g. ability to tolerate 20 rough repeated insertions into a 4-mm-  
109 inner-diameter clear plastic tube curved back on itself with a curve radius of ~3 centimeters;

110 ability to bend tip and neck 90 degrees without breaking; ability to restore to initial form following  
111 bend of 45 degrees; Fig. 1d); strength (resist breakage under rough but reasonable  
112 manipulation); and other factors as applicable (e.g. stickiness, smell).

113 *Collection sufficiency.* We assessed the ability to collect sufficient material for testing using  
114 Gram stain of a swab of the interior cheek smeared onto a standard microscopy slide as a  
115 surrogate for NP swabbing and comparison to Gram stain of a swab of the interior cheek using  
116 Copan Diagnostics, Inc. (Mantua, Italy) model 501CS01 NP swab as the control (Fig. 1c). Slides  
117 were heat fixed and Gram stained according to the BD BBL gram stain test kit protocol (10).  
118 Slides were examined at 40x magnification for the presence of both epithelial cells and bacteria.  
119 Prototypes were passed if they collected a comparable quantity of the material as the control.

120 *PCR compatibility.* We tested PCR compatibility by incubating the head overnight in 3 mL of  
121 modified CDC VTM (Hank's balanced salt solution containing: 2% heat inactivated FBS,  
122 100µg/mL gentamicin, 0.5µg/mL fungizone, and 10mg/L Phenol red (11)) to allow any PCR-  
123 inhibitory material to leach into the medium, spiking 1.5mL with 200 copies/mL of control SARS-  
124 CoV-2 amplicon target (representing 2 times the limit of detection on our system), vortexing, and  
125 testing using the Abbott RealTime SARS-CoV-2 Assay on an Abbott m2000 RealTime System  
126 platform (12), following the same protocol as for clinical testing. PCR-positive prototypes  
127 passed.

128 **Phase II: Production considerations.** We considered stability to autoclaving by repeating  
129 Phase I testing on post-autoclaved materials; manufacturers' short-term strategies for individual  
130 packaging; and manufacturers' stated ability to produce at least 10,000 swabs per day (at the  
131 time roughly a week's worth of swabs for a mid-sized testing center) within a week's notice. We  
132 considered differences in supply chain to minimize the risk of future crises.

133 **Phase III: Field testing.** *Trial design and oversight.* COVIDSwab is an adaptive trial for  
134 evaluating the performance of prototypes compared to the control (see above). Participants  
135 under clinical suspicion for COVID-19 who were scheduled for standard clinical SARS-CoV-2  
136 RT-PCR testing with a control swab were asked also to be swabbed afterward with a single  
137 prototype. Prototypes were collected and tested until at least 10 positive and 10 negative results  
138 on control swabs were obtained (13). Sample collection was performed by trained nursing or  
139 respiratory-therapy staff ("study staff") overseen by the respiratory therapy department at  
140 BIDMC. The Clinical Microbiology laboratories oversaw data collection. This study was reviewed

141 and approved by the institutional review board of Beth Israel Deaconess Medical Center  
142 (protocol number 2020P000323).

143 *Participants.* Participants were individuals clinically suspected of COVID-19 who were brought  
144 to the drive-through/walk-up (“drive-through”) COVID-19 testing site at BIDMC. Adults over 18  
145 years of age were given a participant information sheet by study staff and asked whether they  
146 would agree to being swabbed with a prototype swab performed by a trained nurse or  
147 respiratory therapist in addition to the control swab required for testing. Individuals with known  
148 thrombocytopenia of <50,000 platelets/ $\mu$ l were excluded from the study to avoid risk of mild  
149 bleeding.

150 *Trial procedures.* Prototype swabs were individually packaged and autoclaved at BIDMC for  
151 sterilization according to manufacturer protocols. Swabbing was performed per standard  
152 protocol. Participants were first swabbed with the control swab, then the prototype. Choice of  
153 naris for each swab was left to study staff and the participant. Approximately half of all drive-  
154 through arrivals participated. Control and prototype swabs were placed in separate vials of VTM  
155 and transported to the BIDMC Clinical Microbiology Laboratories where each sample was tested  
156 on the Abbott m2000 SARS-CoV-2 RT-PCR platform as per standard clinical protocol.

157 *Statistical analyses.* RT-PCR results are reported categorically as either positive or negative.  
158 We tested categorical concordance using Cohen’s kappa (14). For each positive test, the Ct  
159 value (the RT-PCR cycle number at which the sample first turns positive) was obtained from the  
160 Clinical Microbiology Laboratories. Higher values reflect lower viral load in the sample.

161 We tested for systematic bias in Ct values by comparing values for controls vs. prototypes using  
162 MWU (15). This tested the null hypothesis that values for controls and prototypes are drawn  
163 from the same underlying distribution;  $p>0.05$  was interpreted as no bias. For discordant  
164 (positive control/negative prototype or vice versa) samples, the negative was assigned a Ct  
165 value of 37, the total number of cycles run. As a second test for bias, we compared (again by  
166 MWU) the distribution of differences in Ct values between control and prototype swabs to the  
167 distribution of differences between two control swabs taken within 24 hours (quality-control data  
168 independent of our study). This tested the null hypothesis that the differences between control  
169 and prototype swabs and the differences between two control swabs are drawn from the same  
170 underlying distribution;  $p>0.05$  was interpreted as no bias.

171 To quantify relative preferences among the prototypes, we gave study staff members printouts  
172 of all six possible pairs of swabs (a “round robin”), in randomized order, and for each pair asked  
173 them to circle their preference (A-B testing). We collated the results and assessed preferences.

## 174 **Results**

175 *Open process.* In the first days of the development effort GitHub repository (6) was established  
176 to serve a public resource and knowledge base. We updated the repository continuously with  
177 design information and test results. These updates included high-resolution images of  
178 prototypes submitted to us for testing (6), a public database of results of our Phase I testing,  
179 and periodic updates and guidance based on our experiences. Open communication facilitated  
180 rapid design iteration by providing anyone interested with a way to quickly understand the  
181 required specifications and to learn from each other’s experiences.

182 *Phase I testing.* To date we have evaluated 48 materials and 160 designs submitted to us for  
183 testing by 4 individuals, 2 laboratories, and 18 companies, for a total of 24 manufacturers.  
184 Seven (4.4%) have passed Phase I testing. Most failures were either for inappropriate materials,  
185 including some that were sticky or brittle, or for inappropriate designs, including those with sharp  
186 heads. Prototypes from 19 manufacturers went through at least two iterations, with a maximum  
187 of 28 prototypes from one manufacturer (Prototype 4 below; Fig. 1). The rate-limiting steps were  
188 receipt of new prototypes, with slow mail delivery during the pandemic being a major  
189 contributor, and PCR-compatibility testing, as testing patient samples took priority over testing  
190 prototypes. Communication with and responsiveness by manufacturers were considered  
191 outstanding.

192 *Phase II and III prototypes.* Four prototypes passed Phase II testing, all of which have  
193 completed our Phase III clinical trial: these are prototypes from the 3D-printing manufacturers  
194 Resolution Medical (with technology from Carbon3D), EnvisionTec, Origin.io, and HP Inc.  
195 (Prototypes 1-4, respectively; Fig. 1a). Like control swabs, the prototypes were 15-16cm in  
196 length with 1-3cm length radially symmetric heads 2-3mm in diameter, a thin neck 4-7cm long  
197 and 1-2mm in diameter, and a thicker shaft 2-4mm in diameter, with a breakpoint most often 7-  
198 8cm from the tip of the head. The materials were plastics and resins such as Keysplint Soft.  
199 Head design evolved over many iterations to increase surface area. Designs generally featured  
200 either a polygonal matrix connected to a central, tapered strut with multiple branch points or else  
201 some form of spiral (Fig. 1b). Manufacturers were able to balance sample collection (Fig. 1c),

202 stiffness, and surface texture. Variations of a longitudinal central strut allowed for varying  
203 degrees of stability, flexibility, and impact cushioning (Fig. 1d).

204 *Sample and data acquisition.* We collected and tested control and prototype swab pairs from  
205 276 participants. Approximately half of the patients tested at our drive-through testing center  
206 participated. Because testing runs were batched and the COVID-19 status of participants was  
207 not known prior to testing, the number of control-positives usually exceeded the minimum  
208 requirement of 10 (range, 10-19). Total collection time was 2-3 days per prototype. The  
209 frequency of control-positive tests was 18%, generally increasing by prototype as the pandemic  
210 worsened in and around Boston.

211 *Comparison.* All four prototypes exhibited a high degree of concordance with the control swab,  
212 with kappas of 0.88, 0.85, 0.89, and 0.88, respectively (Fig. 2a). For convenience we use the  
213 terminology of true positives, true negatives, false positives, and false negatives, with the  
214 control swab result considered the provisional gold standard. Prototypes exhibited 0-1 false  
215 positives and 1-2 false negatives. However, since control swabs are known to be an imperfect  
216 gold standard (<100% sensitivity) and because PCR positives are more likely to reflect true  
217 infection than error, false positives were interpreted as identifying missed infections; indeed,  
218 false positives were referred to clinical care teams as clinically actionable, as per IRB protocol.  
219 Of note, discordant cases were always associated with high Ct values, reflecting low viral load  
220 (Fig. 2b). For example, for Prototype 4, the control swab for one of the two false negatives had a  
221 Ct of 31.47, just shy of 31.50, our hospital's cutoff for reportability (corresponding approximately  
222 to a single virion per mL of VTM); in addition, testing of this false negative was delayed by 16  
223 hours because of prioritizing patient samples, which can result in decreased signal.

224 To better assess possible performance differences between control and prototype swabs, we  
225 compared Ct values for control-prototype pairs for which at least one was positive (assigning the  
226 maximum-possible Ct to negatives; see Methods). Specifically, we asked whether the Ct values  
227 for the prototype swabs were systematically different from those of the control swabs.

228 Systematically higher values for prototype swabs would suggest that they may underperform  
229 control swabs, notwithstanding the high kappa values. A p-value of >0.05 indicates no statistical  
230 difference. Although there were more datapoints below the 1:1 line than above it (Fig. 2b),  
231 statistical testing revealed no evidence for underperformance, with MWU p-values of 0.36, 0.26,  
232 0.42, and 0.31 for Prototypes 1-4, respectively (Fig. 2b). This result supports the conclusion that  
233 the prototypes are non-inferior to the control.



234 As an additional assessment of non-inferiority, we compared the difference in Ct values  
235 observed between control and prototype swabs to the differences between replicates of control  
236 swabs. Independent of our clinical trial, there were 88 cases in which a patient, in the course of  
237 clinical care, was swabbed twice within 24 hours (mean±stdev, 15±7 hours), during the time  
238 period of our study. In 11 of these cases, at least one of the two swabs was positive for SARS-  
239 CoV-2. There were two disagreements between replicate swab tests, resulting in a kappa of  
240 0.90, similar to what was observed in our study for each prototype (kappa=0.85-0.89). Also as in  
241 our study, the Ct values for the first swab and second swab were not significantly different  
242 (MWU p-value of 0.18). Finally, the differences between Ct values for the first and second  
243 control swabs were comparable to the differences between control and prototype swabs (MWU  
244 p-values of 0.31, 0.26, 0.47, and 0.44 for Prototypes 1-4; Fig. 2b).

245 *Staff and participant preferences.* A written staff survey showed a preference for Prototype 4,  
246 then Prototypes 2 and 3, then Prototype 1. There was a slight preference for the control swab  
247 over Prototype 4 (Fig. 3a). In narrative feedback, Prototype 4, which underwent the largest  
248 number of revisions through our process (28), was described as comparable to the control swab  
249 (Fig. 3b).

250 *Availability.* Swabs are available to order. Details can be found on the GitHub repository in the  
251 updates at <https://github.com/rarnaout/Covidswab/tree/master/BIDMC>.

## 252 Discussion

253 The COVID-19 pandemic has forced healthcare providers to seek alternative sources of critical  
254 materials affected by supply-chain disruptions and increases in demand. The situation has  
255 forced providers to innovate under extraordinary time-pressure. Over the course of our study we  
256 received numerous anecdotal reports of swab shortages at hospitals across the United States  
257 and in Europe, necessitating urgent stopgap solutions. Scientific literature on time-sensitive  
258 innovation suggests that open, collaborative, decentralized processes outperform closed or  
259 proprietary ones (7–9). Here we report the success of such a process, going from the  
260 identification of the swab crisis to multiple clinically validated prototypes capable of high-volume  
261 manufacture beginning at 22 days. Notably, none of the prototypes tested were flopped, yet their  
262 performance was statistically indistinguishable from the flopped control swab.

263 The urgency of the situation, the configuration of the manufacturing ecosystem, and human  
264 nature contributed to several observations and shortcomings worth mentioning. First, 3D

265 printing has important advantages in a crisis, including the ability to iterate designs and output  
266 swabs rapidly. It remains to be seen how complementary manufacturing techniques, each with  
267 advantages and disadvantages relative to 3D printing, will contribute in a more mature market  
268 and less urgent setting. Second, in any cooperative process there is a temptation to “defect,” i.e.  
269 taking without giving back. Individuals and manufacturers may well exploit open knowledge for  
270 competitive advantage (16). This is a known price of openness that can disincentivize  
271 cooperation, absent social or structural mechanisms to enforce norms; managing this  
272 temptation took considerable effort by all. Third, ideally the study size would have been larger,  
273 and there would have been a better null model than replicates separated by many hours, to  
274 which to compare our results. Possible sources of variance in our study include differences in  
275 secretions or viral burden between nares and the possibility that the first (control) swab left less  
276 material for the second (prototype) when the same naris was used for both swabs. Despite  
277 these potential issues, our statistical tests supported analytical non-inferiority for all four  
278 prototypes. And fourth, we note our “round-robin” A-B testing survey was useful in summarize  
279 preferences, although the narrative comments seemed often to be more positive than the round  
280 robin suggested. A possible explanation is that the control swab was preferred in large part  
281 simply due to its being familiar, and preferred only narrowly (if often).

282 Like the control swab, the prototype swabs we tested can be improved upon, and manufacturers  
283 are currently doing so. The same is true for other prototypes we may test through our ongoing  
284 clinical trial. Especially in a crisis, perfect is the enemy of good enough. The pandemic  
285 continues to change quickly, and bottlenecks will likely continue to appear unpredictably. The  
286 constant requirement is the ability to respond in a timely fashion under this extraordinary  
287 pressure. We hope our experience, based on past scientific work on cooperation and  
288 innovation, will provide a useful case study for how to iterate and produce a clinically validated  
289 medical manufacture under the pressure of an ongoing pandemic, work on which others will  
290 hopefully improve as we continue to fight COVID-19 together

## 291 **Acknowledgements**

292 The authors would like to thank Abigail Bakken, Alisa Chau, Monique Cole, Zachary Fitts,  
293 Jonathan Goldie, Lan Huynh, Christina Lexinger-Holahan, Lorinda Longhi, Restituto Miranda,  
294 Pavel Gorelik, Jenny Hu, Ofer Mazor, Goutam Reddy, Karen Robinson, Stefan Riedel, Christina  
295 Yen, Summer Decker, Don Ingber, Pawan Jolly, Kit Parker, Adama Sesay, Craig Broady, John  
296 Burpo, Daniel Davis, Joe DeSimone, Annette Friskopp, Ric Fulop, Grant Michael Gonzalez,

297 Markus Greiner, Marie Herring, Matthew Hurley, Hardik Kabaria, Shawn Patterson, David  
298 Lakatos, Ben Linville-Engler, Oren Mechanic, Richard Novak, Jifei Ou, Michael Papish, Steve  
299 Pollack, Chris Prucha, Christian Reed, Isabel Sanz, Al Siblani, Lihua Zhao, Stephanie  
300 Dominguez, Stuart Steinbock, Greg Mark, Nic Scarfo, John Bentley, Nira Pollock, Paula  
301 Watnick, Colleen Baker, Carol Daugherty, Clementina DiMonda, Virginia Dolan, Isaac Fombuh,  
302 Kylie Griffin, Amy Guadagnoli, Allison Wang, Dan Eiref, Johan Sunryd, Jonathan Ford, Sherry  
303 Yu, and Kelsey Ladit for participation in the open process that made this study possible.

304 K.E.Z. was supported by a National Institute of Allergy and Infectious Diseases training grant  
305 (T32AI007061).

306 K.P.S. was supported by the National Institute of Allergy and Infectious Diseases of the National  
307 Institutes of Health under award number F32AI124590.

308 This work was conducted with support from Harvard Catalyst | The Harvard Clinical and  
309 Translational Science Center (National Center for Advancing Translational Sciences, National  
310 Institutes of Health Award UL 1TR002541) and financial contributions from Harvard University  
311 and its affiliated academic healthcare centers. The content is solely the responsibility of the  
312 authors and does not necessarily represent the official views of Harvard Catalyst, Harvard  
313 University and its affiliated academic healthcare centers, or the National Institutes of Health.

314 This work was supported by Carbon3D, Envisiontec, Origin, and HP to cover the cost of the IRB  
315 and PCR testing.

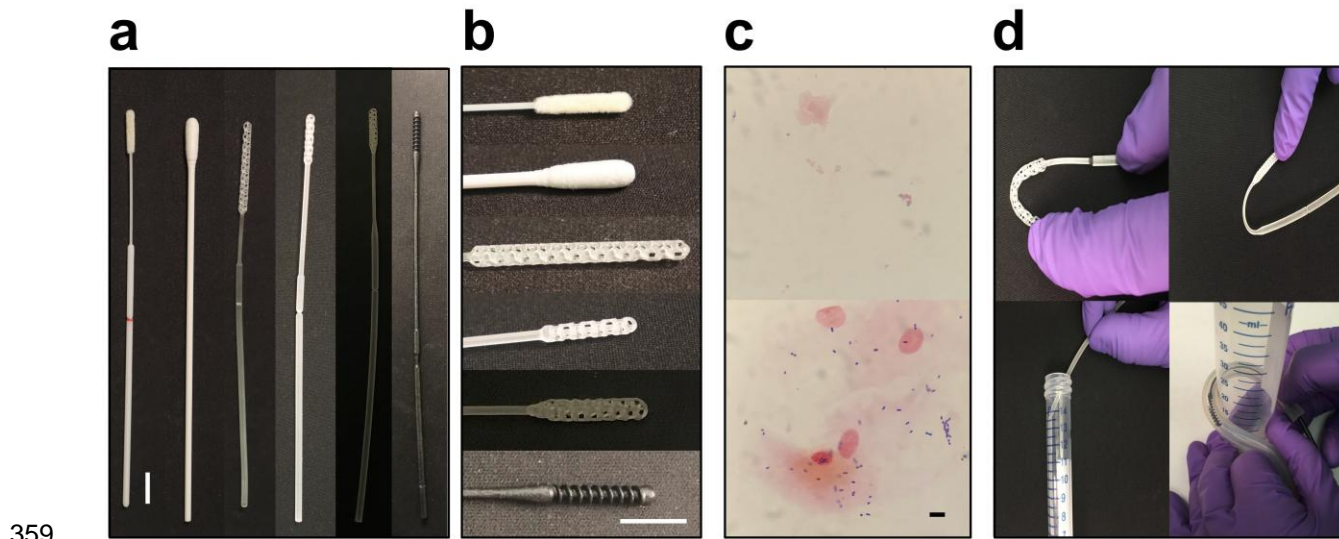
316 The content is solely the responsibility of the authors and does not necessarily represent the  
317 official views of the National Institutes of Health.

318 **REFERENCES**

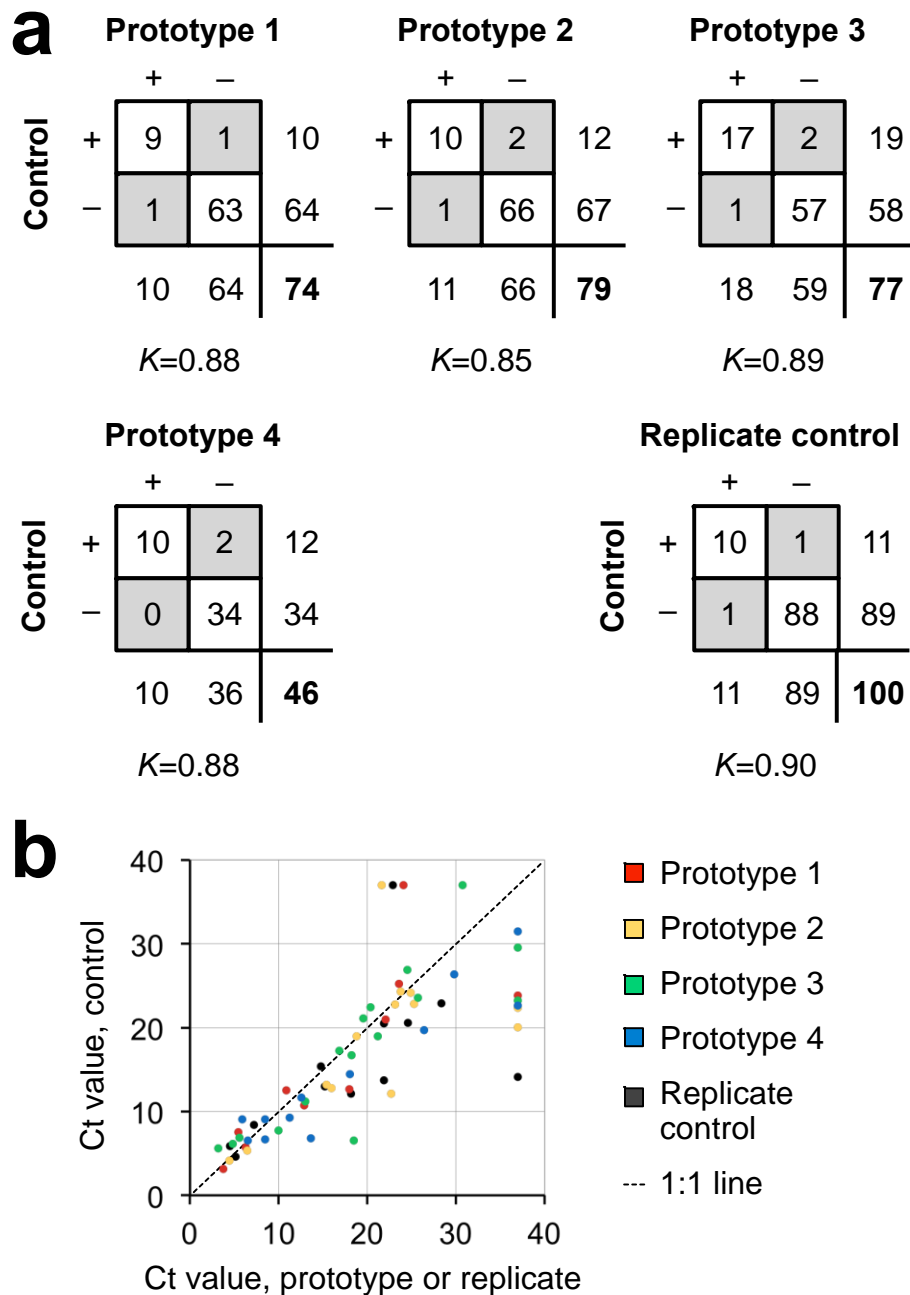
- 319 1. COVID-19 Map - Johns Hopkins Coronavirus Resource Center.  
320 <https://coronavirus.jhu.edu/map.html>. Accessed April 21, 2020.
- 321 2. Thomas K. Coronavirus Test Obstacles: A Shortage of Face Masks and Swabs. *The New*  
322 *York Times*. [https://www.nytimes.com/2020/03/18/health/coronavirus-test-shortages-face-](https://www.nytimes.com/2020/03/18/health/coronavirus-test-shortages-face-masks-swabs.html)  
323 [masks-swabs.html](https://www.nytimes.com/2020/03/18/health/coronavirus-test-shortages-face-masks-swabs.html). Published March 18, 2020. Accessed April 13, 2020.
- 324 3. Kodzius R, Xiao K, Wu J, et al. Inhibitory effect of common microfluidic materials on PCR  
325 outcome. *Sensors and Actuators B: Chemical*. 2012;161(1):349-358.  
326 doi:10.1016/j.snb.2011.10.044
- 327 4. Bessetti J. An Introduction to PCR Inhibitors. 2007. [https://www.promega.es/-](https://www.promega.es/-/media/files/resources/profiles-in-dna/1001/an-introduction-to-pcr-inhibitors.pdf?la=es-es)  
328 [/media/files/resources/profiles-in-dna/1001/an-introduction-to-pcr-inhibitors.pdf?la=es-es.](https://www.promega.es/-/media/files/resources/profiles-in-dna/1001/an-introduction-to-pcr-inhibitors.pdf?la=es-es)  
329 Accessed April 13, 2020.
- 330 5. Tack P, Victor J, Gemmel P, Annemans L. 3D-printing techniques in a medical setting: a  
331 systematic literature review. *BioMed Eng OnLine*. 2016;15(1):115. doi:10.1186/s12938-  
332 016-0236-4
- 333 6. rarnaout. *Rarnaout/Covidswab*.; 2020. <https://github.com/rarnaout/Covidswab>. Accessed  
334 April 13, 2020.
- 335 7. Lakhani KR, Boudreau KJ, Loh P-R, et al. Prize-based contests can provide solutions to  
336 computational biology problems. *Nat Biotechnol*. 2013;31(2):108-111. doi:10.1038/nbt.2495
- 337 8. Lee J, Min J, Lee H. The Effect of Organizational Structure on Open Innovation: A  
338 Quadratic Equation. *Procedia Computer Science*. 2016;91:492-501.  
339 doi:10.1016/j.procs.2016.07.128
- 340 9. Mak RH, Endres MG, Paik JH, et al. Use of Crowd Innovation to Develop an Artificial  
341 Intelligence–Based Solution for Radiation Therapy Targeting. *JAMA Oncol*. 2019;5(5):654.  
342 doi:10.1001/jamaoncol.2019.0159
- 343 10. BD Gram Stain Kits and Reagents. September 2017.  
344 <https://legacy.bd.com/resource.aspx?IDX=19184>. Accessed April 13, 2020.

- 345 11. Preparation of Viral Transport Medium SOP# DSR-052-01. March 2020.  
346 <https://www.cdc.gov/coronavirus/2019-ncov/downloads/Viral-Transport-Medium.pdf>.
- 347 12. Abbott RealTime SARS-CoV-2. March 2020. [https://www.molecular.abbott/sal/9N77-](https://www.molecular.abbott/sal/9N77-095_SARS-CoV-2_US_EUA_Amp_PI.pdf)  
348 [095\\_SARS-CoV-2\\_US\\_EUA\\_Amp\\_PI.pdf](https://www.molecular.abbott/sal/9N77-095_SARS-CoV-2_US_EUA_Amp_PI.pdf). Accessed April 13, 2020.
- 349 13. Miller JM, Campbell S, Loeffelholz M. Changing Swabs: To Validate or Not To Validate? *J*  
350 *Clin Microbiol*. 2013;51(11):3910-3910. doi:10.1128/JCM.02023-13
- 351 14. Kwiecien R, Kopp-Schneider A, Blettner M. Concordance Analysis. *Dtsch Arztebl Int*.  
352 2011;108(30):515-521. doi:10.3238/arztebl.2011.0515
- 353 15. Mann HB, Whitney DR. On a Test of Whether one of Two Random Variables is  
354 Stochastically Larger than the Other. *Ann Math Statist*. 1947;18(1):50-60.  
355 doi:10.1214/aoms/1177730491
- 356 16. Fehr E, Fischbacher U. The nature of human altruism. *Nature*. 2003;425(6960):785-791.  
357 doi:10.1038/nature02043

358 **FIGURES AND FIGURE LEGENDS**

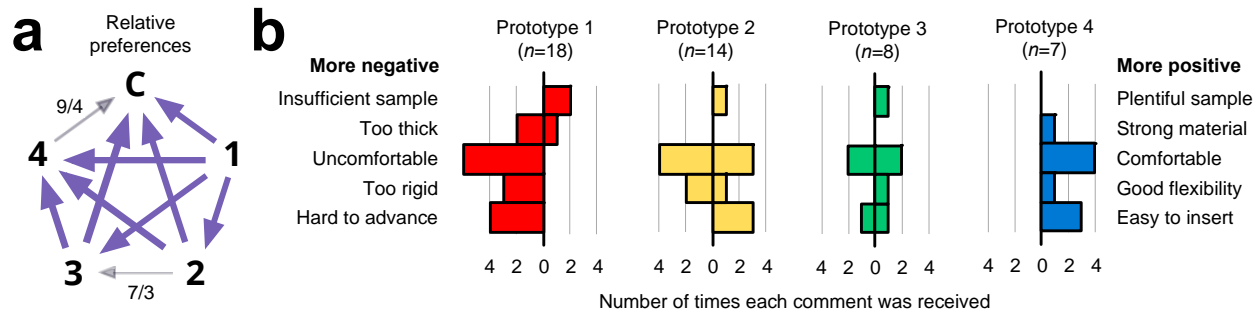


360 **Figure 1:** Control and prototype swabs. **(a)** From left to right: the control swab (Copan  
361 501CS01), a repurposed urogenital cleaning swab approved for NP testing through our process,  
362 Prototype 1 (Resolution Medical), Prototype 2 (EnvisionTec), Prototype 3 (Origin.io), and  
363 Prototype 4 (Hewlett Packard). **(b)** From top to bottom, closeups of the heads of the swabs in  
364 (a). Scale bars, 1cm. **(c)** Examples of Gram stain of cheek swab using control (top) and  
365 prototype swabs. Scale bar, 10 $\mu$ m. **(d)** Examples of materials testing. Clockwise from top left:  
366 head flexibility and robustness to fracture, neck flexibility and robustness to fracture, robustness  
367 to repeat insertion into and removal from a tortuous canal (diameter 3cm), and breakpoint  
368 evaluation.



369

370 **Figure 2: Concordance results. (a)** 2x2 tables giving counts for each prototype vs. the control  
 371 swab (first three panels) and for control vs. replicate control obtained within 24 hours on the  
 372 same individual. Discordant results in gray; totals for each swab below and to the right of each  
 373 box; total number of pairs in bold;  $K$ =Cohen's kappa. **(b)** Scatterplot of Ct values for pairs of  
 374 swabs for which at least one swab was SARS-CoV-2 positive. For discordant pairs, the negative  
 375 swab was assigned a Ct value of 37 (the maximum number of cycles run).



376

377 **Figure 3: Subjective feedback. (a)** Round-robin A-B testing of net preferences among  
 378 Prototypes 1-3 (large bold numbers) and the control (“C”). Each arrow points from the less  
 379 preferred to the more preferred swab. Arrow weight indicates strength of relative preference.  
 380 Preferences were unanimous except where noted with numbers separated by a slash: the first  
 381 number denotes the number of responses for the direction indicated by the arrowhead, while the  
 382 second number denotes the number of responses that had the opposite preference. The weight  
 383 of the arrow is proportional to the difference (e.g. 7-3=a net preference of 4). Unless noted,  
 384 each arrow denotes 12-15 separate responses. **(b)** Number of positive and negative comments  
 385 received from study staff who administered the swabs, tabulated by category. In each plot,  
 386 negative feedback is to the left of the zero, while positive feedback is to the right. Bars on both  
 387 the positive and negative sides of zero reflect differing opinions among study staff. *n*, total  
 388 number of comments received about each prototype.