

1 Real Time Investigation of a large Nosocomial Influenza A Outbreak Informed by Genomic 2 Epidemiology

3
4 Waleed Javaid^{1,2}, Jordan Ehni^{2#}, Ana S. Gonzalez-Reiche^{5#}, Juan Manuel Carreño³, Elena Hirsch³,
5 Jessica Tan^{3,4}, Zenab Khan⁵, Divya Kriti⁵, Thanh Ly⁶, Bethany Kranitzky⁷, Barbara Barnett^{7,12}, Freddy
6 Cera¹¹, Lenny Prespa¹¹, Marie Moss¹, Randy A. Albrecht^{3,8}, Ala Mustafa⁵, Ilka Herbison¹, Matthew M.
7 Hernandez^{3,4}, Theodore R. Pak⁵, Hala Alshammary³, Robert Sebra^{5,9,10}, Melissa Smith⁵, Florian
8 Krammer³, Melissa R. Gitman⁶, Emilia Mia Sordillo^{6*}, Viviana Simon^{1,3,8*}, Harm van Bakel^{5,10*}

9 10 Affiliations:

11 (1) Division of Infectious Diseases, Department of Medicine, Icahn School of Medicine at Mount Sinai,
12 New York, NY.

13 (2) Department of Infection Prevention, Mount Sinai Beth Israel, New York, NY

14 (3) Department of Microbiology, Icahn School of Medicine at Mount Sinai, New York, NY.

15 (4) The Graduate School of Biomedical Sciences, Icahn School of Medicine at Mount Sinai, New York,
16 NY.

17 (5) Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New
18 York, NY.

19 (6) Clinical Microbiology Laboratory, Department of Pathology, Molecular, and Cell-Based Medicine,
20 Icahn School of Medicine at Mount Sinai, New York, NY.

21 (7) Department of Medicine, Icahn School of Medicine at Mount Sinai, New York, NY.

22 (8) The Global Health and Emerging Pathogens Institute, Icahn School of Medicine at Mount Sinai,
23 New York, NY

24 (9) Icahn Institute for Data Science and Genomic Technology, Icahn School of Medicine at Mount
25 Sinai, New York, NY

26 (10) Black Family Stem Cell Institute, Icahn School of Medicine at Mount Sinai, New York, NY

27 (11) Clinical Laboratory, Mount Sinai Beth Israel, New York, NY

28 (12) Department of Emergency Medicine, Icahn School of Medicine at Mount Sinai, New York, NY

29

30

31 # Shared first authors

32 * Correspondence to waleed.javaid@mountsinai.org (WJ), viviana.simon@mssm.edu (VS),
33 harm.vanbakel@mssm.edu (HvB); Emilia.Sordillo@mountsinai.org (EMS)

34

35

36 **Abstract**

37 **Background:** Nosocomial respiratory virus outbreaks represent serious public health challenges.
38 Rapid and precise identification of cases and tracing of transmission chains is critical to end outbreaks
39 and to inform prevention measures.

40 **Methods:** We combined conventional surveillance with Influenza A virus (IAV) genome sequencing to
41 identify and contain a large IAV outbreak in a metropolitan healthcare system. A total of 381
42 individuals, including 91 inpatients and 290 health care workers (HCWs), were included in the
43 investigation.

44 **Results:** During a 12-day period in early 2019, infection preventionists identified 89 HCWs and 18
45 inpatients as cases of influenza-like illness (ILI), using an amended definition, without the requirement
46 for fever. Sequencing of IAV genomes from available nasopharyngeal (NP) specimens identified 66
47 individuals infected with a nearly identical strain of influenza A H1N1 (43 HCWs, 17 inpatients, and 6
48 with unspecified affiliation). All HCWs infected with the outbreak strain had received the seasonal
49 influenza virus vaccination. Characterization of five representative outbreak viral isolates did not show
50 antigenic drift. In conjunction with IAV genome sequencing, mining of electronic records pinpointed
51 the origin of the outbreak as a single patient and a few interactions in the emergency department that
52 occurred one day prior to the index ILI cluster.

53 **Conclusions:** We used precision surveillance to identify and control a large nosocomial IAV outbreak,
54 mapping the source of the outbreak to a single patient rather than HCWs as initially assumed based
55 on conventional epidemiology. These findings have important ramifications for more effective
56 prevention strategies to curb nosocomial respiratory virus outbreaks.

57

58 **Key words:** Precision surveillance, respiratory viruses, influenza A virus, nosocomial outbreak, next
59 generation pathogen sequencing

60

61

62 **Introduction**

63 Nosocomial outbreaks of pathogens represent major challenges for health care providers and
64 institutions. It is critical for hospitals and health systems to not only quickly identify infected cases but
65 also determine the source of the outbreak in order to mitigate the threat to patients and health care
66 workers (HCWs). Nosocomial influenza virus outbreaks have been described worldwide (1-3);
67 children, the elderly, institutionalized and immuno-compromised patients are particularly vulnerable. In
68 some instances, nosocomial outbreaks have been caused by HCWs who work while ill (4).

69 Influenza virus is a single-stranded, negative sense, segmented RNA virus that causes an acute
70 infection of the upper respiratory tract. Two main disease-causing influenza virus types, influenza A
71 viruses (IAV) and influenza B viruses (IBV), circulate in human populations. IAV is further divided into
72 subtypes based on the hemagglutinin (H) and neuraminidase (N) surface proteins. Both IAV subtypes
73 H1N1 and H3N2 circulate in humans and were prevalent in the winter/spring of 2018-2019 in New
74 York City (NYC).

75 Co-circulation and ongoing transmission of multiple virus subtypes in the surrounding community can
76 complicate the determination of whether there is a clonal outbreak and make the identification of the
77 source of the outbreak challenging. In order to develop effective strategies to control and prevent
78 nosocomial outbreaks of influenza, approaches that enable rapid detection of transmission of a clonal
79 influenza virus strain and identification of the origin of that transmission are needed. Complementing
80 traditional epidemiological investigations, precision surveillance utilizing viral genome sequencing has
81 been used previously to trace nosocomial influenza outbreaks (2, 5, 6).

82 Here we report the integration of our conventional infection prevention measures with precision
83 surveillance, including sequencing the genome of IAV from clinical biospecimens and data mining of
84 electronic medical records (EMRs), to successfully identify and control a large nosocomial IAV
85 outbreak affecting both inpatients and HCWs.

86

87 **Results**

88 **Epidemiology of the nosocomial influenza outbreak.** In early 2019, symptoms suggestive of
89 influenza like illness (ILI) were first observed in several HCWs as well as in inpatients receiving critical
90 care at Hospital A in NYC. At the direction of the Infection Prevention Department, the hospital's
91 incident command system was activated. An extensive outbreak investigation was started, which
92 included mandatory staff symptom checks and testing of all inpatients with any respiratory symptoms,
93 regardless of fever status. Enhanced cleaning of patient care areas and clinical staff workspaces was
94 also performed.

95 Over the course of the hospital-wide outbreak investigation, a total of 381 individuals (91 inpatients
96 and 290 HCWs) were screened by regular body temperature checks, symptom surveys and/or
97 molecular diagnostic testing for IAV, IBV and respiratory syncytial virus (RSV). A total of 18 inpatients
98 (19.8%) and 89 HCWs (29.7%) included in the epidemiological investigation tested positive for IAV
99 (**Figure 1A**).

100 Subtyping of IAV from the nasopharyngeal (NP) samples collected during the epidemiological
101 investigation (N=104), the routine influenza surveillance at Hospital A (N=150) and Hospital B
102 (N=231), revealed a stark increase of IAV/H1 at day 4, 5 and 6 of the investigation (**Figure 1B**). Of
103 note, all the samples from inpatients and HCWs included in the investigation that we successfully
104 subtyped harbored IAV/H1N1, suggesting a single transmission chain.

105 The 89 positive HCWs were distributed across 29 different work assignment categories (**Figure 1C**),
106 predominantly front line care providers, including 24 resident physicians (residents, fellows, or interns),
107 16 registered nurses, 8 patient care assistants, and 6 attending physicians. Eighty-seven of these 89
108 HCWs (>90%) had been vaccinated with the quadrivalent seasonal influenza virus vaccine two to five
109 months (average: 108 days) prior to being tested positive for IAV (**Figure 1D**). Importantly, these
110 infected HCWs presented various, and mostly minor, clinical symptoms, and most individuals would
111 not have been classified as having “influenza-like illness” given the lack of fever (**Figure 1E**). Because
112 of this altered influenza disease manifestation in vaccinated HCWs, the case definition was amended
113 early in the context of our investigation.

114

115 **Genomics of the nosocomial influenza virus outbreak.** In order to determine whether there was
116 transmission of a single IAV strain or there were several independent introductions into the hospital
117 system, we performed next generation sequencing (NGS) of IAV from the NP specimens that were
118 banked following the initial diagnostic testing. As part of our Institution’s Pathogen Surveillance
119 Program, we routinely sequence influenza virus from a subset of the patients seeking care at our
120 hospitals (termed “surveillance”). Thus, in addition to cases identified in the outbreak investigation, we
121 included surveillance samples obtained from the general patient population seeking care at our
122 hospitals as part of our Institution’s Pathogen Surveillance Program in order to determine potential
123 community circulating strains.

124 Complete genomic sequences were obtained from 214 IAV isolates (**Figure 2A**), including 126 from
125 Hospital A (investigation and surveillance) and 88 from Hospital B (surveillance only). Pairwise
126 comparison of these viral genomic sequences showed a large cluster of 66 viral isolates that differed
127 by no more than 3 single-nucleotide variants (SNVs), indicating that a single virus clone was
128 responsible for a large portion of the nosocomial outbreak (**Figure 2B**). Additionally, our analyses

129 indicated that other independent introductions of IAV H1N1 strains, with limited forward transmissions,
130 had caused smaller clusters of ILI at both Hospital A and Hospital B. We also noted two small
131 independent clusters due to transmission of IAV H3N2 viruses.

132 Correlating virus genomic sequences with the timing and the source of these isolates showed that all
133 of the virus isolates obtained on Day 0 and most of virus isolates on Day 1 of the infection prevention
134 investigation were distinct from the viral isolate that caused the large outbreak. Although two HCWs
135 were tested positive for IAV on Day 0, their viruses were different from the one that caused the
136 outbreak, and not associated with any nosocomial transmission. All HCWs infected with the outbreak
137 virus had received the seasonal influenza virus vaccine. The first isolate that clustered with the
138 outbreak virus strain was obtained from a patient identified on Day 1 of the investigation (**Figure 2C**).

139 Altogether, the genomic analyses of available clinical influenza isolates showed that cases identified
140 by the conventional epidemiological investigation encompassed patients and HCWs who together
141 harbored 12 different IAV strains and that only one specific strain of these 12 caused widespread
142 nosocomial transmission.

143

144 **Phylogenetic and functional properties of the IAV outbreak strains.** Phylogenetic analysis of the
145 sequenced IAV genomes showed that the outbreak strain tightly clustered within a specific H1N1
146 6b1A subclade. Other IAV H1N1 isolates obtained in the infection prevention investigation and
147 routine surveillance were mapped throughout the H1N1 6b1A clade (**Figure 3A**, compare green to red
148 dots) and likely reflected the predominant seasonal spread in the community. A detailed analysis of
149 genomic sequences of all outbreak IAV strains showed that they were highly conserved; most of
150 nucleotide variations occurred in the HA and the NA segments encoding surface proteins and the
151 majority of predicted amino acid changes occurred only in the NA segment (**Figure 3B**). We selected
152 five outbreak virus strains with representative variants in to be propagated in cell culture for functional
153 characterization (**Figure 3C**). Hemagglutination inhibition assays performed with these outbreak virus
154 strains confirmed that none had drifted as compared to the H1N1 vaccine strain used in that season
155 (**Figure 3D**).

156

157 **Reconstruction of the transmission chain in the early days of the outbreak.** In order to
158 understand the origin of the outbreak and the factors that facilitated its rapid spread, we first focused
159 on the early stages of the outbreak between days 0-3. We obtained IAV genomes for 10 of the 12
160 positive cases identified in the outbreak investigation (83%), as well as for 34 of 46 other positive
161 surveillance samples obtained at hospitals A and B (74%). Of these, eight cases and one surveillance

162 sample matched the outbreak strain, and we used the *PathoSPOT* framework (<https://pathospot.org>)
163 to query various electronic hospital records in order to create a timeline (**Figure 4A**).

164 The data showed that 4 of the 9 initial nosocomial IAV cases were seen (patients p1, p2, and p3) or
165 worked (HCW1) in the emergency department (ED) on the same day (Day -1), during overlapping
166 time periods. The 3 patients were admitted from the ED to different wards and had no other shared
167 interactions with HCWs, indicating that the common exposure most likely occurred in the ED.
168 Similarly, because one of the patients who acquired nosocomial IAV did not have direct contact with
169 HCW1 and had already been transferred out prior to the time HCW1 was present in the ED, the
170 evidence indicates HCW1 was exposed during that work shift rather than being the primary case.

171 In contrast, patient p3, the putative primary case for this outbreak, was brought to the ED in the
172 morning of Day -1, several hours before p1 and p2, and interacted with HCW1. P3 was admitted to a
173 medical unit the same day (**Figure 4A**, grey) with fever, shaking chills, dyspnea, and abdominal pain,
174 but developed systemic inflammatory response syndrome and was transferred to the ICU where the
175 patient was intubated, a procedure that can generate significant aerosols (7) Because blood cultures
176 of p3 grew Gram-negative bacteria, the diagnosis of IAV was delayed. However, the patient remained
177 febrile despite antibiotics prompting diagnostic testing for influenza virus on day 3.

178 The next three early cases (HCW2, p4 and p5) most likely acquired infection in the ICU from p3,
179 although p4 and p5 had further overlapping stays following transfer to the same medical/surgical
180 inpatient unit from the ICU. Our data suggests that cases p6 and p7, who were admitted through the
181 ED several days after the start of the outbreak, acquired IAV infection from ED HCWs who had been
182 exposed to p3 and became IAV-test positive in the days thereafter.

183 The interaction network based on available contact records (**Figure 4B**) indicates that almost all later
184 cases can be traced back in the some way to the initial nine cases shown in the timeline (**Figure 4A**),
185 with transfers of patients acting as the major vectors for spread to other hospital units. A direct
186 transmission link could not be documented for two patients cared at and two HCWs assigned to the
187 same closed unit psychiatric ward, suggesting that indirect interactions may have occurred elsewhere
188 in the hospital.

189

190 **Discussion**

191 In this report we describe using a combination of conventional infection prevention approaches and
192 precision surveillance to identify and control a nosocomial IAV outbreak in the spring of 2019. Swift
193 interventions such as widespread molecular influenza virus testing and modification of the case

194 definition to include mild respiratory disease presentations brought the outbreak under control within
195 two weeks.

196 Almost a third of the 290 HCWs included in the epidemiological investigation tested positive for IAV
197 (**Figure 1**). The amended case definition for ILI was based on the realization that most HCWs infected
198 with IAV were afebrile. The vast majority of infected HCWs had received the seasonal influenza virus
199 vaccination. Among the infected hospital employees for which viral influenza genotypes were
200 available (49/89), 41 harbored the outbreak strain whereas eight were infected with unrelated H1N1
201 viruses (**Figure 2**). In the absence of our infection prevention intervention, many of these cases would
202 have gone undiagnosed, pointing to the fact that influenza virus infections in vaccinated HCWs remain
203 largely under-diagnosed due to the milder disease presentation. Thus, diagnostic testing of
204 vaccinated HCWs with mild symptoms should be considered, especially when the seasonal influenza
205 vaccine is well matched to the circulating viruses.

206 By sequencing influenza virus genomes from the infected patients and HCWs, we could focus the
207 investigation into the source of the outbreak on only those cases that actually were infected with the
208 identified outbreak virus strain. Of note, the two HCWs who were the first to be diagnosed with ILI and
209 initially considered the likely source(s) of the outbreak were infected with viruses distinct from the
210 outbreak strain (**Figure 2**). It is also important to note that the time of diagnosis may not be a good
211 indicator for the actual chronology of an outbreak, which may be true especially for vaccinated HCWs
212 who may ignore or downplay symptoms. Indeed, we observed that almost three days transpired
213 between the onset of symptoms and positive IAV test in HCWs compared to the average of one-half
214 day for inpatients. Integrating data from various hospital electronic records with molecular
215 confirmation of which patients and HCWs were infected with the outbreak strain enabled
216 reconstruction of the dynamics of the outbreak, identification of the likely primary case, and therefore
217 ensured reassessment of transmission and heightened remediation for areas where transmission
218 occurred.

219 Since we routinely sequence influenza virus isolates from patients receiving care at our health system
220 as part of our Pathogen Surveillance Program, we could compare the strains from the outbreak
221 investigation conducted at Hospital A to the strains found in the surveillance of Hospital A as well as
222 Hospital B (**Figure 3**). These additional data allowed us to not only identify previously unrecognized
223 smaller transmission events (four inpatients and two HCWs at Hospital A and six patients at Hospital
224 B) but also ascertain that there was a large (in number) but limited (in time) outbreak of H1N1 in our
225 health system (**Figures 2 and 3**). Importantly, this specific H1N1 virus outbreak strain did not spread
226 further in the community.

227 A limitation of our study is that we did not have access to biospecimens from 22 HCWs whose tests
228 were performed at laboratories outside our health system. Additionally, only partial viral genomes
229 could be retrieved from two of the available biospecimens linked to the epidemiological outbreak
230 investigation. However, we were able to obtain viral genomes for all the patients identified during the
231 first three days of the outbreak, providing a solid foundation for the reconstruction of the transmission
232 chain.

233 Our data suggest that the outbreak began in the ED most likely through introduction of the virus by a
234 single patient, who had received aggressive resuscitative care and was subsequently transferred to
235 the intensive care unit of Hospital A (**Figure 4**). A possible solution to mitigate such risks to HCWs
236 and patients in the future is to enhance screening and isolation of patients coming into the ED with
237 any respiratory symptoms, even when an alternative diagnosis seems to be the predominant
238 complaint. Additionally, recognition by hospital leadership of the potential for transmission even from
239 HCWs with mild influenza illness resulted in administrative support for intensified education of staff to
240 avoid working while ill, extended sick leave when needed, and a move away from the HCW culture of
241 “presenteeism” which can contribute to nosocomial transmission of influenza (4) and other respiratory
242 viruses.

243 It is critical for patient care that any healthcare organization quickly detects the occurrence of hospital
244 acquired infections and limits their spread through swift identification of their origins. Conventional
245 infection prevention approaches, however, are challenged if the hospital outbreak occurs in the
246 context of widespread community acquired infections (e.g., during the peak of the influenza season as
247 in this study). Our findings are applicable to a wide range of highly transmissible respiratory viral
248 pathogens including Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2). Indeed,
249 emerging evidence suggest that a high percentage of HCWs became infected with SARS-CoV-2
250 before screening for acute infection became more available. Implementation of precision surveillance
251 measures as outlined here will be of critical importance to mitigate the risks of nosocomial
252 transmission for patients and HCWs alike.

253 **Methods:**

254 **Ethics statement.** The Pathogen Surveillance Program and the in depth analysis of the outbreak
255 were approved by the Institutional Review Board of the Icahn School of Medicine at Mount Sinai.

256 **Summary of infection prevention measures and investigation.** When we detected ILI cases in
257 inpatients in early 2019, infection prevention measures, including extensive screening of inpatients
258 and HCWs for ILI and collection of nasopharyngeal (NP) swab specimens were implemented. Once
259 the outbreak was identified, a case definition was developed. This specific case definition included
260 any HCWs or patient with cough, rhinorrhoea, sore throat, body aches, with or without fever, and
261 positive diagnostic test for influenza virus by molecular polymerase chain reaction (PCR) testing
262 (Xpert® Xpress Flu/RSV test, Cepheid) of a NP swab specimen in universal transport medium (NP-
263 UTM).

264 During this initial investigation, a total of 381 individuals from hospital A, including 91 inpatients and
265 290 HCWs, were screened based on symptoms and laboratory detection of respiratory viruses.
266 Among these 381 individuals, a total of 89 HCWs and 18 patients were found to be infected with IAV.
267 A high incidence of cases was found among HCWs, including interns, residents, fellows, and rotators,
268 as well as nursing and associated fields. Clusters were found in several areas inside the hospital
269 including intensive care units and several medical/surgical floors and an inpatient psychiatric unit.
270 Prophylaxis with oseltamivir was offered to exposed HCWs and inpatients. Treatment with oseltamivir
271 was provided to infected inpatients and HCWs, in addition to extended sick leave for HCWs who
272 remained afebrile but symptomatic. There was no mortality reported in association with the outbreak.
273 One patient required readmission secondary to ILI. We educated staff throughout the hospital
274 including the trainees on symptoms associate with influenza and encouraged all to report to their
275 supervisors if they had any symptoms. Mandatory symptom checks at beginning of each shift were
276 also implemented. Anyone symptomatic was encouraged to get tested. We were able to contain this
277 outbreak from detection to eradication within 10 days, with incredible collaboration between many
278 groups.

279

280 **Collection of NP-UTM for influenza A subtyping and sequencing.** A total of 486 IAV positive NP-
281 UTM specimens, including 104 samples from the epidemiological outbreak investigation, 150 routine
282 influenza surveillance samples at Hospital A, and 231 routine influenza surveillance samples at
283 Hospital B, were collected and submitted for IAV subtyping and next generation sequencing (NGS).
284 The time frame from which surveillance and investigation samples were included covered a total of 27
285 days, starting from six days before the investigation to seven days after the 12 days long outbreak
286 investigation. Viral RNA was extracted from 280µL of NP-UTM using the QIAamp Viral RNA Minikit

287 (QIAGEN, cat. 52904), as per the manufacturer's instruction. To distinguish between IAV H1 and H3,
288 we used a modified version of the WHO One-Step Real-time RT-PCR protocol and viral RNA isolated
289 from NP-UTM specimens. The primers and probes are multiplexed so that one can distinguish
290 influenza A subtypes (e.g., IAV/H1pdm09 and IAV/H3) in the same reaction. All reactions were run in
291 duplicates using the QuantiFast Pathogen RT-PCR +IC Kit (QIAGEN, cat. 211454) on the Roche
292 LightCycler 480 Instrument II (Roche Molecular Systems, 05015243001) in 384-well plates. Every run
293 included two positive controls and nuclease-free water served as a non-template control. The
294 following temperature profiles was used: 50°C for 20 min, 95°C for 1 sec, 95°C for 5 min, followed by
295 40 cycles of 95°C for 15 sec and 60°C for 45 sec, during which quantitation of products occurred. For
296 a result to be called positive for influenza A(H1)pdm09 or A(H3) subtype, replicates were required to
297 have Ct values < 35. A result was deemed negative if both reactions displayed Ct values ≥ 35. A
298 result was deemed inconclusive if one of the reactions was negative (Ct ≥ 35).

299 **Next generation sequencing of the genome of influenza virus.** RNA from the NP samples and
300 viral isolates were used for whole genome amplification of the IAVs by performing a multisegment RT-
301 PCR for whole genome amplification with Opti1 primers (8). Multi-segment PCR amplicons were
302 cleaned by 0.45X of Agencourt AMPure XP magnetic beads (Beckman Coulter) according to
303 manufacturer's protocol. The concentration of purified amplicons was measured using the Qubit High
304 Sensitivity dsDNA kit. After samples were normalized to a concentration of 0.2 ng/μl, adapters added
305 by tagmentation using the Nextera XT DNA library preparation kit (Illumina). Samples were purified
306 using 0.7X of Agencourt AMPure XP Magnetic Beads and fragment size distributions were analyzed
307 on a Bioanalyzer using the High Sensitivity DNA kit (Agilent). After bead-based normalization
308 (Illumina) according to the manufacturer's protocols, sequence-ready libraries were sequenced in a
309 paired-end run using the MiSeq v2, 300cycle reagent kit (Illumina). Complete genomes were
310 assembled from next generation Illumina reads using a custom genome assembly pipeline (8).
311 Custom codes and pipelines are also available at <https://bitbucket.org/bakellab/flugap>.

312 **Identification of clonal outbreak isolates.** To detect clusters of highly related outbreak isolates we
313 used the open-source *PathoSPOT* (Pathogen Sequencing Phylogenomic Outbreak Toolkit) software
314 (<https://pathospot.org>), which we developed to aid detection and visualization of transmissions in
315 nosocomial settings. The *PathoSPOT compare* pipeline was used to first cluster influenza genomes
316 based on their Mash distance using default distance threshold parameters. Segment sequences were
317 concatenated together into a single genome sequence per isolate, and the concatenated sequences
318 were then used to construct a multi-genome alignment of isolates in each cluster using *parsnp*.
319 (9) Pairwise distances between genomes were then calculated as the number of single nucleotide
320 variants (SNVs) between genome alignments in each cluster for further analysis. To identify
321 transmission events we used the *PathoSPOT heatmap view* (Fig. 2). We set a threshold of ≤3 SNVs

322 to identify potential transmissions. Next, the *dendro-timeline view* was used to perform phylogenetic
323 analysis of outbreak isolates, and to reconstruct the early outbreak timeline based on the full
324 admission/transfer/discharge (ADT) history for each patient obtained from the electronic medical
325 record. Finally, the ADT history was combined with patient-HCW interaction data to reconstruct a
326 network of all known contacts in *Cytoscape*.(10).

327 **Phylogenetic analysis.** Hemagglutinin (HA) H1 segment sequences representing the global diversity
328 of H1N1 viruses circulating during the 2018-19 influenza season (September 1, 2018 - April 30, 2019)
329 as of August 20, 2019 were obtained from the Global Initiative on Sharing All Influenza Data (GISAID)
330 database. Only complete HA coding sequences (1701 nt) were included. In order to avoid potential
331 duplicates, for records with identical strain names but different accession numbers, only one
332 sequence was retained. When available, sequences derived from the original samples or from
333 isolates with the lowest number of passages were included. After applying these filters, a total of
334 8,314 HA H1 sequences were retained. In addition, to inform lineage annotation, 6 reference strains
335 from 2015 to 2018, including the vaccine strains A/Michigan/45/2015 (H1N1) and A/Brisbane/02/2018
336 (H1N1) were included. To enable comparison, 163 H1N1 sequences from isolates recovered from
337 Hospital B patients and HCWs during the outbreak time-window were also included. A maximum-
338 likelihood (ML) phylogeny was inferred with RAxML (11) performing 20 ML searches under the
339 GTRCAT model of nucleotide substitution and assessed with 100 bootstrap replicates.

340 In order to confirm phylogenetic grouping, we inferred a second ML analysis at the whole genome
341 level for all strains that belonged to the highest supported clade that contained the outbreak isolates.
342 For this analysis, the coding regions for the 10 main IAV proteins (PB2, PB1, PA, HA, NP, NA, M1,
343 M2, NS1, and NEP) were concatenated. In addition to the outbreak isolates, this clade included 17
344 isolates from locations outside New York City that were retained as an outgroup.

345 **Propagation of clinical viral isolates.** NP-UTM that tested positive for influenza virus were aliquoted
346 and frozen at -80°C. For viral isolation and growth, samples were pre-diluted in infection media
347 consisting on Minimum Essential Media (Gibco) supplemented with 1µg/ml of tosyl sulfonyl
348 phenylalanyl chloromethyl ketone (TPCK) treated-trypsin (Gibco) and added to 90% confluent MDCK
349 cells monolayer. Infections were let to proceed for 1 hour with intermittent shaking. The inoculum was
350 removed and infection media was added. Plates were incubated at 37°C with 5% CO₂ and
351 supernatants were collected at 24-48 hours after infection when cytopathic effect was observed.
352 Samples were cleared by centrifugation at 4300g for 5 min, supernatants were filtered to avoid
353 bacterial contamination then stored at -80°C.

354 **Hemagglutination inhibition (HI) assay.** These assays were performed as described before (12, 13).
355 Briefly, serum samples from eight healthy study participants receiving the seasonal influenza virus

356 vaccine (four weeks post-vaccination) were treated with receptor destroying enzyme (RDE) (14) and
357 serially diluted (2 fold dilutions) in 96-V-well microtiter plates, followed by addition of 8
358 hemagglutination units (HAU) from five selected viral isolates from HCW (HCW20, HCW18, HCW36)
359 as well as inpatients (p2 and p3) per 50 μ l of PBS. The mixtures above were incubated for 30 min at
360 room temperature, then a suspension of 0.5% turkey red blood cells (RBC, Lampire Biological) was
361 added, and plates were further incubated for 1 hour at 4°C. A/Michigan/45/2015 (matching with
362 seasonal vaccine) was used as the reference strain to assess antigenic drift in the clinical viral
363 isolates. Likewise, immune sera from ferrets exposed to A/Michigan/45/2015 were used for this
364 purpose. Hemagglutination titers were calculated as the reciprocal of the last serum dilution at which
365 the antibodies present in the serum inhibited agglutination of the viruses with the red blood cells.

366

367 **Acknowledgments**

368 We thank the numerous team members of the Clinical Microbiology laboratories at the Mount Sinai
369 Health System and the Department of Microbiology for expert support and their willingness to go the
370 extra mile. We thank Ms. M.C. Bermudez and the team of the Personalized Virology Initiative for
371 expert processing of human samples. We also extend our gratitude to Dr. A. Garcia-Sastre and Dr. P.
372 Palese for guidance and many thought provoking discussions. We are indebted to Dr. C. Cordon-
373 Cardo for continued support of the Pathogen Surveillance Program.

374 We would like to thank the staff, the Nursing leadership, including Christine Mahoney and Maria
375 Latrace, the Residency Program Leadership including Daniel Steinberg, the Hospital Leadership
376 including Dr. J. Boal at Mount Sinai Beth Israel, and the New York State Department of Health,
377 especially Dr. K. Southwick and Rafael Fernandez, for their extraordinary help and support during the
378 outbreak.

379 We gratefully acknowledge the authors, originating and submitting laboratories of sequences from
380 GISAID's EpiFlu (www.gisaid.org) that were used as background for our phylogenetic inferences. The
381 list of authors and submitting laboratories is shown in **Table S1**.

382 This study was supported in parts by the National Institute of Allergy and Infectious Disease (NIAID)
383 Centers of Excellence for Influenza Research and Surveillance (CEIRS) contract
384 HHSN272201400008C, the Office of Research Infrastructure of the National Institutes of Health
385 (S10OD018522 and S10OD026880) and institutional seed funds.

386 This manuscript was edited by Life Science Editors.

387

388

389 Authors contributions:

390 J.W, J.E, B.K., B.B., M.M, and I.M. were involved in the outbreak investigation. J.W, J.E, M.G., and
391 E.M.S. provided clinical evaluations. J.W, J.E, M.G., and E.M.S. collected and verified clinical data.
392 E.H., J.T, T.L., F.C., L.P., M.M.H., M.G., E.M.S., V.S., and H.B. performed clinical sample
393 accessioning. E.H., M.M.H., and H.A. performed RNA extraction and viral subtyping. J.M.C, and V.S.
394 selected and grew viral isolates . R.A.A., and V.S. provided banked serum samples. J.M.C performed
395 HAI assays . A.S.GR., Z.K., and L.M. performed NGS experiments. A.S.GR., R.S., and M.S. provided
396 NGS services. A.S.GR., D.K., T.R.P., and H.B. performed genome assembly and/or
397 comparative genome analyses. J.E, T.R.P., E.M.S., and H.B. performed mining of electronic medical
398 records. J.W, J.E, A.S.GR., M.M.H., F.K., E.M.S., V.S., and H.B. analyzed, interpreted, and/or
399 discussed data. A.S.GR., E.M.S., V.S., and H.B. wrote the manuscript. E.M.S., V.S., and H.B.
400 conceived the study. E.M.S., V.S., and H.B. supervised the study. E.M.S., V.S., and H.B. raised
401 financial support

402

403

404 **Figure legends:**

405

406 **Fig. 1: Epidemiology of the nosocomial IAV outbreak**

407 **A:** Time line of the nosocomial IAV outbreak at a metropolitan hospital. Day 0: initiation of the
408 infection prevention investigation. IAV, influenza A virus; HCW, healthcare workers.

409 **B:** Distribution of IAV subtypes detected in individuals identified in the outbreak investigation and
410 patients seeking care at the hospital and testing positive for IAV between 6 days before (Day -6) and
411 19 days after (Day 19) the initiation (Day 0) of the infection prevention investigation. The red
412 background highlights the period of the outbreak investigation.

413 **C:** The distribution of professions (29 job categories) of the 89 healthcare workers who tested IAV
414 positive. PCA: Patient care assistant.

415 **D:** The distribution of days between receiving seasonal influenza virus vaccination and being tested
416 positive for IAV among HCWs. Of note, 87 of 89 HCWs tested positive for IAV were vaccinated.

417 **E:** Clinical signs and symptoms reported by HCWs who were tested positive for IAV. The data
418 available for each symptom differs with respect to the number of employees (N listed in parenthesis
419 provides the absolute numbers). Note that 63% of HCWs were afebrile.

420

421

422 **Fig. 2: Genomics of the nosocomial outbreak**

423 **A:** Venn diagram illustrating the number of sequenced outbreak confirmed H1N1 strains (N=66) and
424 outbreak excluded IAV, including the unrelated H1N1 strains (N=113) and H3N2 strains (N= 36),
425 identified in Hospital A (investigation and surveillance) and Hospital B (surveillance). Epidemiology
426 describes the cases identified by infection prevention.

427 **B:** Pairwise comparison of the complete viral genomes. Note the tight cluster of the outbreak H1N1
428 strains (N=66, red cluster) at the center and the two small H3N2 clusters at the top left of the pairwise
429 comparison.

430 **C:** Dynamics of the case numbers and IAV strains during the investigation period. All color codes
431 used in this panel are same as those used in the panel B. The outbreak H1N1 strain (shown in red)
432 was first detected on the day after the initiation of the investigation in a hospitalized patient. The first
433 two employees who were tested positive for IAV on the day of the initiation of the investigation
434 harbored unrelated H1N1 strains.

435

436

437 **Fig. 3: Phylogenetic and functional properties of the IAV outbreak strains**

438 **A:** Analysis of IAV H1N1 genomic sequence diversity. Note that the outbreak strains form a tight
439 cluster that maps to one subclade of the H1N1 6b1A clade whereas those not responsible for the
440 outbreak are diverse and map to various subclades that are different from the outbreak one. IAV
441 H1N1 vaccine strains A/Michigan/2015 and A/Brisbane/2018 are included in the analysis as
442 references.

443 **B:** Phylogenetic relationships based on single nucleotide variant (SNV) distance among the outbreak
444 virus strains. All eight viral segments are shown, with grey indicating synonymous changes and dark
445 blue indicating non-synonymous substitutions.

446 **C:** Genotype of the five representative outbreak strains that were grown in cell culture. Grey,
447 synonymous mutations; dark blue: non-synonymous changes.

448 **D:** Hemagglutination inhibition titers of five representative outbreak strains with the sera of six recently
449 vaccinated individuals. Sera from ferrets infected with the vaccine strain A/Michigan/45/2015 served
450 as a positive control for antisera and the A/Michigan/45/2015 virus is the antigen positive control.

451

452 **Fig. 4: Reconstruction of the transmission chain in the early days of the outbreak.**

453 **A:** Timeline and locations inside the hospital of the first nine cases harboring the IAV/H1N1 outbreak
454 strain. Day 0 is the day the investigation started. Dots on the line point to the day on which
455 patient/health care worker tested positive for IAV. The dashed lines indicate presumed contact
456 between infected individuals. The black box highlights the three patients that received care in the
457 emergency department a day prior to the start of the infection control investigation.

458 **B:** Interaction network of health care workers (HCW) and patients (p) who were tested positive for the
459 IAV H1N1 outbreak strain. The cases identified by thick black outlines were critical in seeding the
460 transmission within the hospital.

461

462

463

464 **Literature cited:**

- 465 1. **Blackburn RM, Frampton D, Smith CM, Fragaszy EB, Watson SJ, Ferns RB, Binter S,**
466 **Coen PG, Grant P, Shallcross LJ, Kozlakidis Z, Pillay D, Kellam P, Hue S, Nastouli E,**
467 **Hayward AC, group I.** 2019. Nosocomial transmission of influenza: A retrospective cross-
468 sectional study using next generation sequencing at a hospital in England (2012-2014).
469 *Influenza Other Respir Viruses* **13**:556-563.
- 470 2. **Godoy P, Torner N, Soldevila N, Rius C, Jane M, Martinez A, Cayla JA, Dominguez A,**
471 **Working Group on the Surveillance of Severe Influenza Hospitalized Cases in C.**
472 2020. Hospital-acquired influenza infections detected by a surveillance system over six
473 seasons, from 2010/2011 to 2015/2016. *BMC Infect Dis* **20**:80.
- 474 3. **Huzly D, Kurz S, Ebner W, Dettenkofer M, Panning M.** 2015. Characterisation of
475 nosocomial and community-acquired influenza in a large university hospital during two
476 consecutive influenza seasons. *J Clin Virol* **73**:47-51.
- 477 4. **Wilson KE, Wood SM, Schaecher KE, Cromwell KB, Godich J, Knapp MH, Sklar MJ,**
478 **Ewing D, Raviprakash K, Defang G, Whitman TJ.** 2019. Nosocomial outbreak of
479 influenza A H3N2 in an inpatient oncology unit related to health care workers presenting
480 to work while ill. *Am J Infect Control* **47**:683-687.
- 481 5. **Cummings MJ, Tokarz R, Bakamutumaho B, Kayiwa J, Byaruhanga T, Owor N,**
482 **Namagambo B, Wolf A, Mathema B, Lutwama JJ, Schluger NW, Lipkin WI, O'Donnell**
483 **MR.** 2019. Precision Surveillance for Viral Respiratory Pathogens: Virome Capture
484 Sequencing for the Detection and Genomic Characterization of Severe Acute Respiratory
485 Infection in Uganda. *Clin Infect Dis* **68**:1118-1125.
- 486 6. **Houlihan CF, Frampton D, Ferns RB, Raffle J, Grant P, Reidy M, Hail L, Thomson K,**
487 **Mattes F, Kozlakidis Z, Pillay D, Hayward A, Nastouli E.** 2018. Use of Whole-Genome
488 Sequencing in the Investigation of a Nosocomial Influenza Virus Outbreak. *J Infect Dis*
489 **218**:1485-1489.
- 490 7. **Tran K, Cimon K, Severn M, Pessoa-Silva CL, Conly J.** 2012. Aerosol generating
491 procedures and risk of transmission of acute respiratory infections to healthcare workers:
492 a systematic review. *PLoS One* **7**:e35797.
- 493 8. **Mena I, Nelson MI, Quezada-Monroy F, Dutta J, Cortes-Fernandez R, Lara-Puente JH,**
494 **Castro-Peralta F, Cunha LF, Trovao NS, Lozano-Dubernard B, Rambaut A, van Bakel**
495 **H, Garcia-Sastre A.** 2016. Origins of the 2009 H1N1 influenza pandemic in swine in
496 Mexico. *Elife* **5**.
- 497 9. **Treangen TJ, Ondov BD, Koren S, Phillippy AM.** 2014. The Harvest suite for rapid core-
498 genome alignment and visualization of thousands of intraspecific microbial genomes.
499 *Genome Biol* **15**:524.
- 500 10. **Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski**
501 **B, Ideker T.** 2003. Cytoscape: a software environment for integrated models of
502 biomolecular interaction networks. *Genome Res* **13**:2498-2504.
- 503 11. **Stamatakis A.** 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis
504 of large phylogenies. *Bioinformatics* **30**:1312-1313.
- 505 12. **Liu STH, Behzadi MA, Sun W, Freyn AW, Liu WC, Broecker F, Albrecht RA, Bouvier**
506 **NM, Simon V, Nachbagauer R, Krammer F, Palese P.** 2018. Antigenic sites in influenza
507 H1 hemagglutinin display species-specific immunodominance. *J Clin Invest* **128**:4992-
508 4996.

- 509 13. **Tan GS, Krammer F, Eggink D, Kongchanagul A, Moran TM, Palese P.** 2012. A pan-H1
510 anti-hemagglutinin monoclonal antibody with potent broad-spectrum efficacy in vivo. *J*
511 *Virol* **86**:6179-6188.
- 512 14. **Nachbagauer R, Wohlbold TJ, Hirsh A, Hai R, Sjursen H, Palese P, Cox RJ, Krammer F.**
513 2014. Induction of broadly reactive anti-hemagglutinin stalk antibodies by an H5N1
514 vaccine in humans. *J Virol* **88**:13260-13268.
- 515

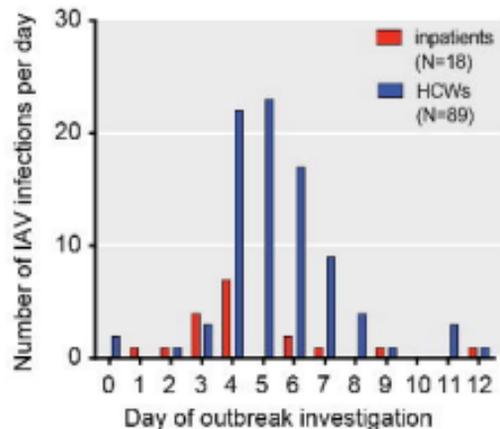
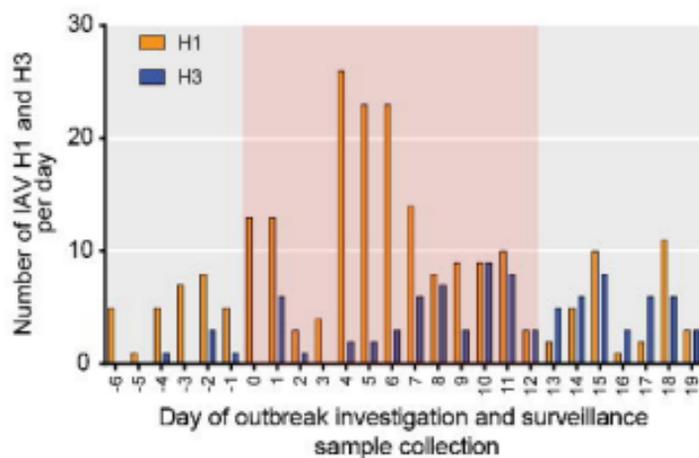
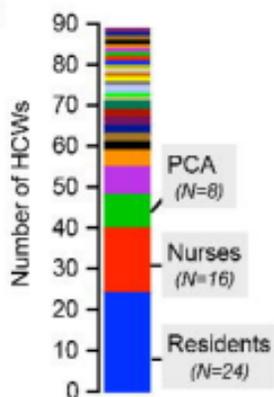
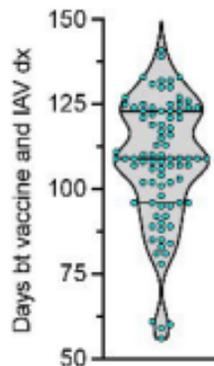
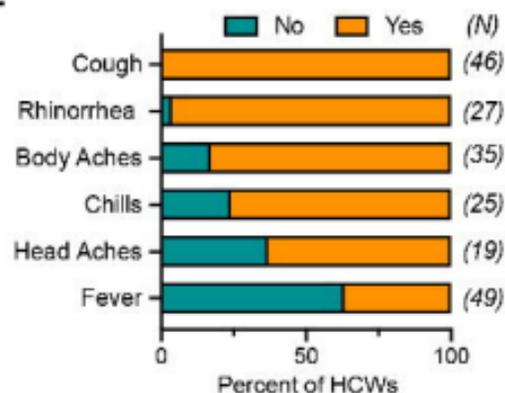
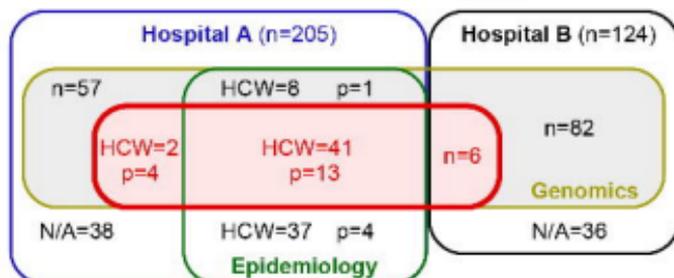
Figure 1**A****B****C****D****E**

Figure 2**A**

Positive IAV tests on day 0-12



Information from viral genomes

○ Confirmed outbreak ○ Excluded outbreak ○ N/A (No viral genome)

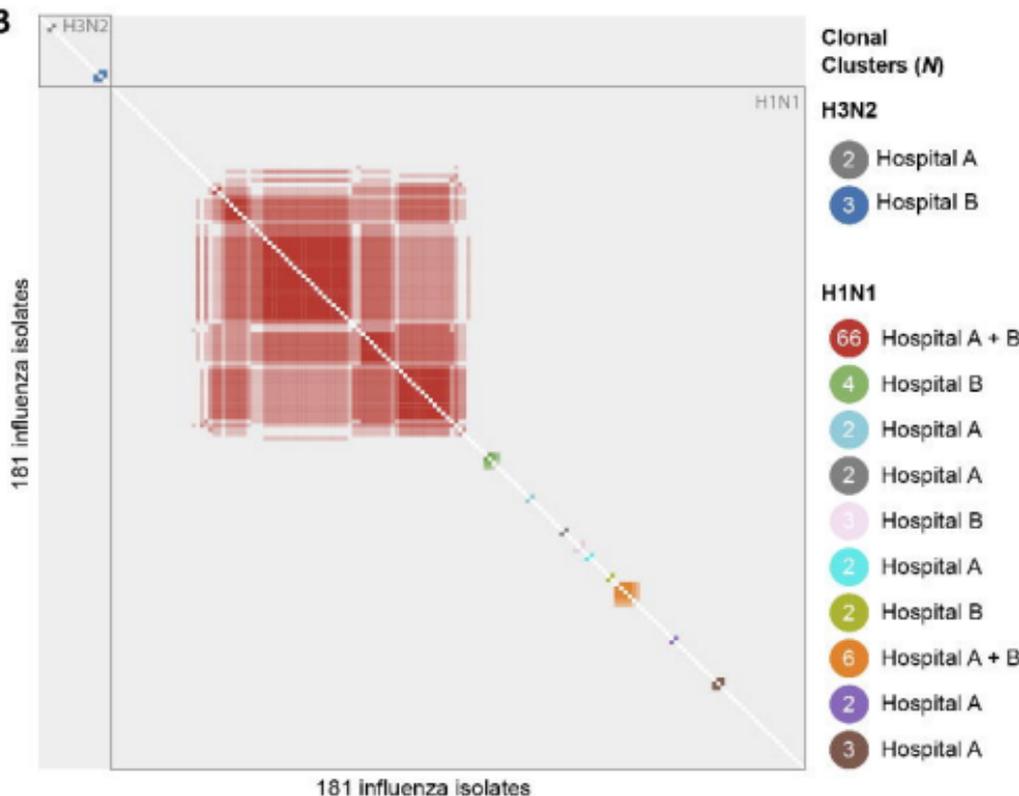
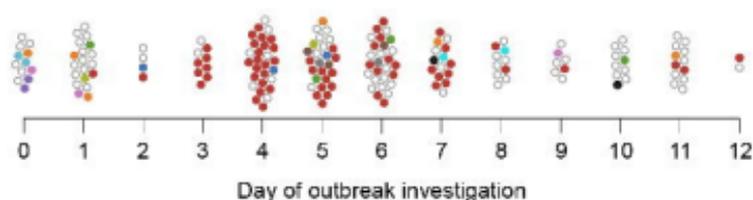
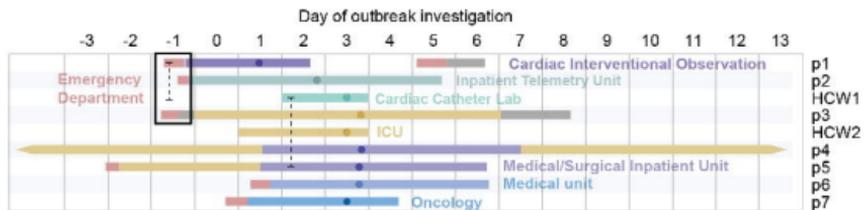
B**C**

Figure 4**A****B**