Cellphone traces reveal infection-associated behavioral change

Ymir Vigfusson[a,b,1,2], Thorgeir A. Karlsson[b,1], Derek Onken[a,1], Congzheng Song[c], Atli F. Einarsson[b], Nishant Kishore[d], Rebecca M. Mitchell[a], Ellen Brooks-Pollock[e], Gudrun Sigmundsdottir[f], and Leon Danon[e,g,h]

[a]Simbiosys Lab, Department of Computer Science, Emory University, Atlanta, GA 30322; [b]School of Computer Science, Reykjavik University, 101 Reykjavik, Iceland; [c]Department of Computer Science, Cornell University, Ithaca, NY 14853; [d]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA 02115; [e]Department of Veterinary Medicine and Population Health Sciences, University of Bristol, Oakfield Grove, Bristol BS8 2BN, United Kingdom.; [f]Landspitali University Hospital and CHS-CDC, 101 Reykjavik, Iceland.; [g]University of Exeter, Exeter EX4 4QF, United Kingdom.; [h]The Alan Turing Institute, British Library, 96 Euston Rd, London NW1 2DB, United Kingdom.

Epidemic preparedness depends on our ability to predict the trajectory of an epidemic and the human behavior that drives spread in the event of an outbreak. Changes to behavior during an outbreak limit the reliability of syndromic surveillance using large scale data sources such as online social media or search behavior, which could otherwise supplement healthcare-based outbreak prediction methods. Here, we measure behavior change reflected in mobile phone call detail records (CDR), a source of passively-collected real-time behavioral information, using an anonymously linked dataset of cellphone users and their date of influenza-like illness diagnosis during the 2009 H1N1v pandemic. We demonstrate that mobile phone use during illness differs measurably from routine behavior: diagnosed individuals exhibit less movement than normal (1.1–1.4 fewer unique tower locations, $p < 3.2 \times 10^{-3}$) on average in the 2 to 4 days around diagnosis, and place fewer calls (2.3–3.3 fewer calls, $p < 5.6 \times 10^{-4}$) while spending longer on the phone (41–66 seconds average increase, $p < 4.6 \times 10^{-10}$) than usual on the day following diagnosis. The results suggest that anonymously linked CDR and health data may be sufficiently granular to augment epidemic surveillance efforts, and that infectious disease modelling efforts lacking explicit behavior change mechanisms need to be revisited.

disease | surveillance | call detail records | influenza | outbreak

Infectious disease outbreaks remain a major threat to humanity in the 21st century, as evidenced by the ongoing pandemic of Coronavirus disease 2019 (COVID-19) (1) and 5 of 10 threats to global health identified by the World Health Organization (WHO) being related to infectious disease (2). Estimating the current and future burden of disease through surveillance and predictive modelling is essential for appropriate allocation of resources aimed at reducing impact, especially in the early stages of an outbreak.

Traditional influenza healthcare-based surveillance methods rely on data gathered from symptomatic individuals seeking medical treatment from doctors. These approaches suffer from delays in reporting that differ from setting to setting and difficulty in identifying unusual activity (3). Such issues led to the development of alternative syndromic surveillance methods (4) that combine a broad range of data sources on behavioral markers; some were developed, used, and assessed during the H1N1v pandemic (5). These surveillance methods include analyzing patterns in social media such as Twitter (6, 7), search engine queries (8–10), over-the-counter medication sales (11), airport traffic patterns (12), city traffic patterns (13), cell phone surveys (14), or ensemble methods that incorporate survey data (15). Directly inferring disease incidence from these sources also assumes that the cause of behavior change is known and usually associated with influenza. Yet, studies indicate that individuals alter behavior for various reasons even when not symptomatic, *e.g.*, to avoid infection (16) or due to anxiety (17), complicating estimation of infectious disease burden (18).

Whereas data sources that depend on active, conscious user participation may produce unreliable estimates (14, 20), call detail records (CDR) can act as a passive pattern sensor (21). Mobile networks pervade most nations: in raw numbers, 2019 cellphone subscriptions in developed and developing countries exceeded 100% of their populations (22), although mobile use invariably skews away from under-resourced groups (23). CDR, collected in real-time, contain spatio-temporal information that capture mobility. Past analyses have used cell phone data to study human movement scaling (13), social network structure inference (24), poverty and wealth prediction (25), and risk and spread of multiple diseases including malaria (26, 27), cholera (28), and influenza (29). Furthermore, smartphone apps have been used to track behavior change in relation to influenza onset (30), or as contact trackers during the COVID-19 pandemic (31, 32). These methods are all limited by either unreliable health data (self-diagnosed symptoms), aggregate-level data to model the population (33), or fraught with privacy concerns (34). Until now, the link with verified health data at the individual level has been missing.

Significance Statement

Infectious disease control critically depends on surveillance and predictive modelling of outbreaks. We argue that routine mobile phone use can provide a new source of infectious disease information via the measurements of behavioral changes in call detail records (CDR) collected for billing. In anonymous CDR metadata linked with individual health information from the A(H1N1)pdm09 outbreak in Iceland, we observe that people move significantly less, and placed fewer but longer calls in the few days around diagnosis than normal. These results suggest that disease transmission models should explicitly consider behavior changes during outbreaks, and advance mobile phone traces as a potential universal data source for such efforts.

L.D conceived the study, Y.V., L.D. designed the research, G.S. contributed data, Y.V., T.K., D.O., A.F.E., N.K., R.M.M., L.D. analyzed data, T.K., D.O, C.S. devised models, Y.V., T.K., D.O., A.F.E., E.B.-P., L.D. wrote the paper.

The authors declare no conflict of interest.

[1]Y.V., T.K., and D.O. contributed equally to this work.

[2]To whom correspondence should be addressed. E-mail: ymir.vigfusson@emory.edu

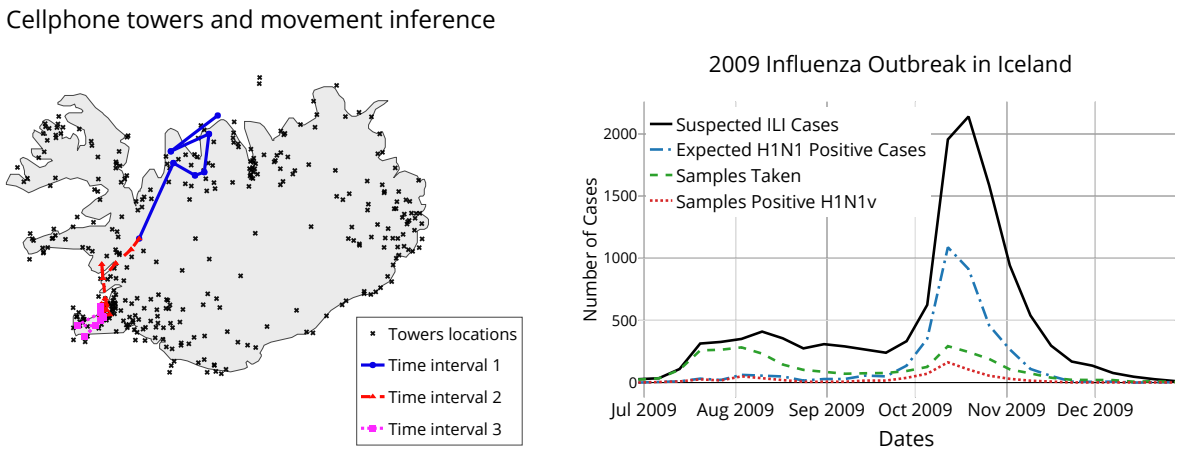


Fig. 1. Combining health records with call data records. (*left*) Cell towers act as a proxy for location which, when coupled with the timestamp, allow movement inference. Different colors show inferred movements of a typical cellphone user at different time periods over a period of three days. (*right*) The epidemic curve for the 2009 H1N1v outbreak in Iceland, showing a single pronounced peak. The green dotted line shows the number of laboratory samples taken, the red line shows the number of those testing positive for H1N1v, and the black line shows the estimate of suspected H1N1v cases per week from the recorded ILI incidence (19). The expected H1N1v positive cases (blue dotted line) are extrapolated from the suspected ILI cases and the percentage of samples found positive each week.

Here, we explicitly combine CDR with information from the 2009 H1N1v pandemic collected by the national healthcare-based surveillance system used by all health providers in Iceland through a protocol that maintains reasonable expectations of individual privacy from government surveillance. The influenza pandemic reached Iceland in May 2009 (19), with a shallow peak before the school holidays in May/June 2009, followed by a dip over the summer and a strong peak in October 2009 (Fig. 1). The outbreak started in the capital of Reykjavík, home to 37% of the population of 318,499, approximately one week ahead of the rest of the country (19). Health officials recorded the date of diagnosis (DoD) of 10,175 clinically diagnosed cases of influenza-like illness (ILI) around the country between 4 June 2009 and 11 February 2010. Of 3,011 samples taken, 700 were confirmed by a real-time polymerase chain reaction (PCR) protocol to be H1N1v influenza infections (19); we assume that other patients diagnosed with ILI were infected with the same strain, which displaced other strains until February 2010 (35).

We analyze behavioral patterns in Iceland extracted from the CDR, provided in a deidentified format by a major mobile network operator (MNO). The CDR logs span a broad time period around the 2009 outbreak. Mobile phone owners are anonymously matched to records of ILI diagnosis, yielding DoD and CDR traces for 1,434 diagnosed individuals after data processing. We measure and identify behavioral traits that show significant changes in the diagnosed group around the DoD compared to a control group.

Methods

Data Collection. The original dataset joins individual CDR that MNOs routinely gather for billing purposes with individual level ILI diagnosis data from Iceland's Centre for Health Security and Communicable Disease Control (CHS-CDC), that collects and stores all records of ILI diagnoses in Iceland. We developed and used a privacy-preserving data hand-off and merging protocol approved by the Iceland's Bioethics Committee (Vísindasiðanefnd): a large MNO sent encrypted phone IDs and national identification numbers (NINs, which are public information in Iceland), to the CHS-CDC. The CHS-CDC supplied dates of ILI diagnoses for NINs, then replaced NINs with an anonymous encrypted identifier before providing the data to us (SI:Data Linking and Privacy). The MNO provided us with CDR data (SI:Mobile Network Data) containing the encrypted IDs of the phones on either side of a call, the timestamp, the length of the call (in seconds), and the geographical coordinates of the cellphone towers that interacted with the phones (Table S1). No demographic or private data, such as age, gender, or contents of calls or texts, are included. The cell tower accessed during normal phone use provides a proxy for the device's location. The granularity of location varies with locality—regional tower density increases proportionally with regional population (Fig. 1). At the time, MNOs provided cell coverage for virtually all residences in Iceland, either directly through their network or through a roaming service. We filter out individuals with multiple subscriptions (SI:Data Preprocessing). Using phone ownership information, each phone is matched to the DoD of its owner for the subset of users that pay only for one phone. This post-processed subset, referred to as the dataset below, accounts for 25-30% of the MNO's users and encompasses all data analyzed in our paper. We define the home tower of each individual as the tower that picks up more calls and texts between midnight and 8 am than other towers. The distribution of home tower locations was strongly correlated with residential census counts for the corresponding postal codes for our dataset ($r = 0.86$, $p < 8 \times 10^{-49}$) and among those with ILI diagnosis ($r = 0.88$, $p < 2 \times 10^{-43}$). The home towers were thus spatially representative for the entire Icelandic population. We focus our analysis on the 1,434 diagnosed users who generated sufficient CDR data to establish a home tower location in the 4-week period centered on their DoD. De-identified aggregate data, code and documentation used in our analysis is available at https://github.com/SimBioSysLab/cdr-open-code.

Feature Extraction. To characterize user behavior, we extract 36 features (independent variables) from both incoming and outgoing CDR data encompassing movement, activity, and

Table 1. Feature Characteristics from the 29-day period around each individual's DoD (additional characteristics in Table S2). The final column indicates which days relative to DoD show a statistically significant difference between the control and diagnosed groups.

Feature		Control Mean	Diagnosed Mean	Anomalous Days
Unique Locations Visited		**3.04**	**2.74**	**-12, -2, -1, 0, 1, 2, 3, 4**
Number of New Location Visited		0.5	0.43	1,3
Unique Contact Count	incoming	2.25	2.02	∅
	outgoing	2.50	2.28	1,2,3,6
	both	4.04	3.67	2
New Contacts	incoming	0.61	0.50	∅
	outgoing	0.65	0.55	1, 3, 6
	both	1.19	1	0, 1, 3
Call Duration (s), total	incoming	190	480.51	0, 1, 2
	outgoing	162	435.09	0
	both	479.5	915.60	∅
Calls Count	incoming	3.10	2.84	∅
	outgoing	**3.60**	**3.37**	**1,2**
	both	6.66	6.22	1
Texts Count	incoming	2.71	2.87	-10
	outgoing	1.74	1.93	1,2
	both	4.46	4.79	-10, 11
Calls and Texts count	incoming	5.78	5.71	∅
	outgoing	5.34	5.30	1,2
	both	11.12	11.01	2
Mean Call Duration	incoming	133.35	140.74	-1, 0, 1, 2, 4, 11
	outgoing	107.96	106.74	∅
	both	**143.14**	**149.84**	**-8, -1, 0, 1, 2, 3, 4, 11, 13, 14**
Top-3 Contacts by Duration (SI:Sensitivity Analysis)	incoming	0.68	0.69	∅
	outgoing	0.70	0.69	∅
	both	1.37	1.38	1
Remaining Contacts by Duration	incoming	1.57	1.33	∅
	outgoing	1.80	1.58	1,2,4,6
	both	3.37	2.92	1,4,6
Top-3 Contacts by Frequency	incoming	0.58	0.56	∅
	outgoing	0.63	0.61	1,2
	both	1.21	1.17	1
Remaining Contacts by Frequency	incoming	1.67	1.46	0
	outgoing	1.88	1.67	-9, 2, 6
	both	3.54	3.13	∅

social network behavioral patterns (SI:Feature Extraction). Most features exhibit a right-skewed distribution (Fig. S2) and share general characteristics across control and diagnosis groups. They include the following (boldface in Table 1).

Unique locations visited measures the number of unique tower coordinates connected to by the cellphone within a time interval (bin). This feature helps describe movement during the time period, but can inflate in areas where multiple towers can provide cellular signal.

Mean call duration (incoming and outgoing) measures call activity by dividing the total duration of calls by the number of calls the user placed or received in the time interval.

Number of calls (outgoing) measures the number of calls placed by the device in the time interval.

Departure from Routine Behavior. We use $x_{fid} = E_f(i, d)$ to denote the raw feature value for a feature f, extracted from the CDR by function E, for individual i, and on day d. Extraction is performed for all features f in Table 1.

To control for the weekly behavioral routine of individual i, each feature value is detrended through linear regression over values of the same weekday in the past W weeks. Specifically, let

$$p_j = x_{fi,(d-7\cdot(W-j))} \quad \text{for } j = 0, 1, \ldots, W$$

and denote by J those indices $j \in \{0, 1, \ldots, W\}$ where p_j is defined. Then $(p_j)_{j\in J}$ is the measured behavior on the same day of the week from the previous W weeks before day d for feature f and individual i, with p_W denoting the behavior in week W. We used $W = 10$ weeks of past data to correct for seasonality in our experiments, which gave comparable results to an alternative approach to detrending based on ranking features and normalizing them (SI:Seasonality).

Based on the data, we used a linear model to capture the change in values over time $p_j = \beta j + \alpha + \varepsilon_j$ with errors ε_j for each $j \in J$, we fit parameter values for $\hat{\alpha}$ and $\hat{\beta}$ to minimize the squares regression error

$$\arg\min_{\hat{\alpha},\hat{\beta}} \sum_{j\in J} \varepsilon_j^2 = \arg\min_{\hat{\alpha},\hat{\beta}} \sum_{j\in J} (p_j - \beta j - \alpha)^2 .$$

The detrended feature value, measuring the deviation from weekly routine, is then defined as

$$z_{fid} = x_{fid} - \hat{\beta}W - \hat{\alpha}.$$

Control Group. Each diagnosed individual is matched with a control individual from the undiagnosed group, based on home location. All measurements thus far have applied to individuals diagnosed with ILI during the epidemic. To compare the diagnosed population against a control population, a subset is selected from the rest of the data—those not diagnosed for ILI are assumed to be uninfected though they may show behavior consistent with symptoms but are well, or have ILI symptoms but did not use health services. Of 74,644 people, we were able to identify home towers for 36,140. Each diagnosed person's control is selected randomly from the undiagnosed individuals among the 36,140 who share a home tower with the diagnosed individual. For this dataset, control selection exhibited no noticeable differences across three methods: selecting randomly, matching for home tower, or matching home tower and frequency of calls (36).

We analyze the pattern differences between the means of the detrended feature values (z_{fid}) of the individuals in the two groups. The 29-day range (2 weeks either side of DoD) centered around every diagnosed individual's DoD range $[-14, 14]$, with DoD mapping to 0. Controls use the same days of data as their diagnosed match. The average deviation from weekly routine on all days in the range are compared (Fig. S9) with original feature values (x_{fid}) shown in Figs. S2, S3 and S8.

Statistical Comparison. We compare the behavior of the diagnosed and control groups across each detrended feature value z_{fid} and each day using the Kolmogorov-Smirnov statistic. To counteract the increase in Type I errors caused by running multiple significance tests, we use the Benjamini-Hochberg (BH) procedure to control the False Discovery Rate (FDR), as it presents the most conservative FDR correction for this mix; the adjusted p-values can then be used to assess the evidence for or against the null hypothesis. The BH procedure assumes independent tests. Some tests act on dependent, interacting samples—*e.g.*, a value on a specific day is ranked against values from the same day of the week for several weeks prior—whereas others are independent tests. Confidence bands for the KS-test are computed and plotted for each day of the primary three features deemed significant based on the p-values with the FDR correction (Fig. S9). The significance test and the confidence interval calculations use $\alpha = 0.05$.

Results

Several features show significant change between the routine behavior of the control and diagnosed populations around their DoD. The actual time period and magnitude of the behavioral change varies by feature (Table 1, rightmost column), but the

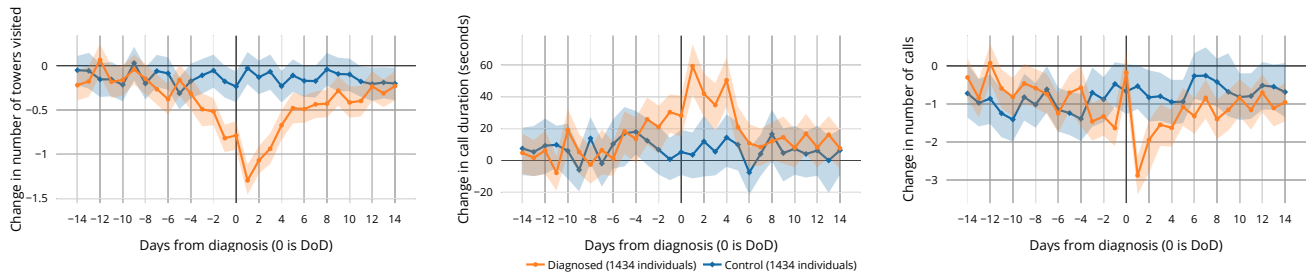


Fig. 2. Changes in average phone use behavior associated with diagnosis. (*left*) Users visit fewer locations on days around diagnosis. (*middle*) They make and receive longer phone calls on days near diagnosis. (*right*) They initiate fewer calls on the days after diagnosis, with the exception of the day of diagnosis itself. Graphs display the mean deviation from "normal" routine behavior (z_{fid}) for each group on the relative day of illness determined by date of diagnosis (day 0). Confidence intervals (2.5% to 97.5%) are calculated using bootstrapping (SI:Visualizations).

number of *unique locations visited*, *mean call duration*, and the *number of outgoing calls* show the most pronounced signals of behavior change.

Less movement. The ***unique locations visited*** feature indicates that the diagnosed group tends to travel less than usual, even before diagnosis. Such lower travel patterns coincide with the typical symptomatic period of influenza (37). The maximum effect is observed on the day following diagnosis, when diagnosed individuals travel to 1.1–1.4 fewer locations than normal. Differences are observed between the diagnosed and control groups from two days prior to the DoD until four days after DoD ($KS > 0.084$, $p < 3.2 \times 10^{-3}$, Figs. 2 and S10). Other days in the 4-week period display the diagnosed and control groups acting similarly.

Longer calls. ***Mean call duration*** shows that people tend to make longer calls on average on the day after the DoD (Fig. 2), when significant differences are observed between the diagnosed and control groups ($KS = 0.155$, $p < 4.6 \times 10^{-10}$, Fig. S22). On the day following diagnosis, diagnosed individuals spend an average of 41–66 seconds longer on the phone than usual.

Fewer calls placed. ***Number of outgoing calls*** gives another perspective of behavior following diagnosis. Although call duration increases around DoD, the number of outgoing calls decreases on the day after the DoD with an average of 2.3–3.3 fewer calls than is routine ($KS = 0.102$, $p = 5.6 \times 10^{-4}$, Fig. S18). On the day of diagnosis, diagnosed individuals increase outgoing calls relative to their routine compared to the days before and after.

Statistical significance through FDR-corrected p-values is supported by KS confidence intervals for nearly all comparisons (Figs. 2 and S9). Notably, the diagnosed group displays significant changes in mobility even prior to seeking healthcare and receiving a diagnosis (SI:Visualizations).

Limitations. The results depend on the metadata arising from mobile phone use, presenting both advantages and drawbacks (21, 33). The increased data bandwidth provided by MNOs and rapid device and app development over the past decade has altered user behavior patterns to communicate more via internet-based applications and less via calls and text. In our dataset, cellular internet data access (denoted GPRS) provided additional location information to CDR records of calls and texts, a situation that has likely shifted since the H1N1 outbreak (SI:Comparing CDR and GPRS Data). Three years following the epidemic, the Icelandic CDR and GPRS data contained a stronger location proxy than in 2009 due to more smartphone apps periodically connecting to cellular towers for Internet access but poorer information for features pertaining to call duration, frequency, and top contacts. Since many nations experience limited Internet access (53.6% of world population in 2019 (22)) and smartphone availability (39.4% worldwide (38)), it would be reasonable to assume that call and text usage in these locations may follow similar patterns as in our dataset, but we caution against assuming all cell phone behavior to be universal (33). Further, mobile phone ownership may bias against those in greatest need of public health intervention. The results report aggregate behavior change which are likely to include patterns caused by other illnesses or injuries. Our approach depends on maintaining individual-level behavioral histories since the signal we identified concerns departure from routine behavior rather than the actual behavior itself, as seen by comparing the raw and detrended distributions six days prior to diagnosis (Fig. S2) with the day following the DoD (Fig. S3). Finally, Iceland contains a small, mostly homogeneous, and generally affluent population bound to an island, with idiosyncratic behavior, including unusually high mobile phone usage. Seasonal effects may be exaggerated in Icelanders compared to other populations due to Iceland's proximity to the Arctic.

Discussion

The combination of mobile phone traces with health records reveals behavior change associated with symptom onset for H1N1v in unprecedented detail. Observations of behavior in CDR are consistent with our knowledge of influenza pathology: individuals become infected and begin showing symptoms which their behavior reflects; they then access health care, receive a diagnosis, and display activity patterns different from normal for a period of time, after which they return to normality. This picture depicts a group trend; however, in an effort to avoid ecological inference fallacy (39), we observe that individuals' changed behavior varied widely within a group. The variability of individuals' behavioral responses suggests that CDR data is best suited for aggregate analysis of symptomatic behavior.

Although we cannot know the exact cause in each individual case, collectively, the duration of anomalies is consistent with estimates of influenza symptom duration (40). The use case in

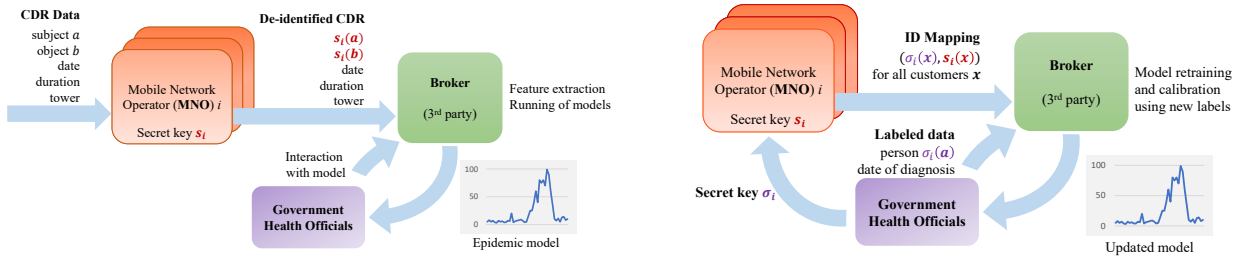


Fig. 3. Privacy-preserving data sharing protocol. Privacy-preserving architecture for syndromic surveillance using CDR data for future experimental design. An independent 3rd party broker is provided with real-time de-identified CDR data, extracts features and runs the prediction models to generate an epidemic curve (left, (**O1**)). The broker could also be provided labeled anonymous health information to join with the CDR data to calibrate or retrain the classifiers (right, (**O2**)). The design accommodates mutual distrust, ensuring that health officials cannot monitor behavior or track mobility of individuals, that MNOs are not provided with any health information of customers, and that the broker only operates on de-identified data.

Iceland demonstrates that disease monitoring systems could be expanded with CDR, already passively collected by local mobile operators, that can discern behavior consistent with ILI symptoms while following a protocol to preserve user privacy, and our approach provides a complementary way of estimating the duration of symptoms, and therefore an important component for estimating the economic impact of an outbreak.

The results presented here have important implications for modelling disease dynamics. As individuals change behavior due to symptom onset, their potential to transmit is modified, yet, modelling efforts that have been central to mitigation measures for novel pathogens tend to ignore behavioral effects due largely to a dearth of quantitative information. Such limitation is evident in the case of modeling of SARS-CoV-2 transmission, for instance, where different groups vary in their ability to alter their behavior in response to exposure or illness (31, 41). Here, we quantify the direction and magnitude of the behavioral change effect for H1N1v on an atypical population that exhibits fewer sources of variability than most. Other pathogens and populations will have different properties that will require a context-specific investigation. Our work provides a methodology for capturing and quantifying behavior change that can be used to improve the predictive power of models in future outbreaks. We argue that such an approach would have an important part to play in outbreak response for novel pathogens.

A separation of access to private data is vital for ensuring public trust. While aggregation helps protect privacy (31), enabling health officials to interact with the data increases the risk to individual or group privacy. Concerns have being raised over government responses to COVID-19, where contacts of those infected are traced from historical CDR data (34). Our data sharing protocol (Fig. 3 and SI:Privacy-Preserving Data Sharing) mitigates risk by ensuring that: (a) mobile operators that hold cell phone metadata do not have access to any new health information for their customers held by health officials, and (b), health officials do not access cell phone metadata. To further strengthen the separation, differential privacy methods can be used to introduce controlled noise to the data in such a manner that aggregate statistics remain unchanged while provably protecting the privacy of individuals and small groups (42, 43). At the same time, communicating the collective benefit of studies such as this one, and the effort taken to protect data, is necessary to help the public decide when the public health value of the information provided is worth the risk to their privacy.

Our results suggest that CDR metadata may allow surveillance of symptomatic diseases whose symptom intervals are sufficiently long and behavioral changes sufficiently pronounced that they produce a signal that is visible at the resolution afforded by the data. The granularity of these data is rapidly refining, both spatially with denser tower infrastructure being built in response to population growth and newer generations of devices (*e.g.*, 5G), and temporally as mobile phones become increasingly used for Internet applications. Greater data resolution may help offset the relatively small effect sizes in our results, which are confounded by other brief interruptions to people's routines, and allow the approach to extend beyond a large-scale epidemic of a transmissible pathogen. Environments lacking health monitoring infrastructure but where mobile phone use is prevalent and consistent (33) have the greatest potential gains from CDR-based epidemic surveillance. In particular, establishing the nature of symptomatic behavior provides an opportunity to use artificial intelligence to identify patterns suggesting an individual or a group is symptomatic, and thus estimate the numbers of cases. We are optimistic that further study could establish the full generality and versatility of infectious disease surveillance using call data records on their own.

ACKNOWLEDGMENTS. The work was partially supported by the Icelandic Centre for Research (RANNIS) Award 152620-051, an Emory University Research Council Award, National Science Foundation (NSF) Faculty Early Career Development (CAREER) Grant 1553579 and a hardware donation from NVIDIA Corporation. LD gratefully acknowledges the Leverhulme Trust Early Career Fellowship and is supported by The Alan Turing Institute EPSRC EP/N510129/1. LD and EBP are also supported by Medical Research Council (MRC) (MC/PC/19067).

1. World Health Organization, Coronavirus disease 2019 (COVID-19): situation report, 72 (2020).
2. World Health Organisation, Ten threats to global health in 2019 (2019) Accessed: 2020-03-16.
3. JR Ortiz, et al., Strategy to enhance influenza surveillance worldwide. *Emerg. infectious diseases* **15**, 1271 (2009).
4. Triple S Project, Assessment of syndromic surveillance in Europe. *The Lancet* **378**, 1833–1834 (2011).
5. M Lipsitch, FG Hayden, BJ Cowling, GM Leung, How to maintain surveillance for novel influenza A H1N1 when there are too many cases to count. *The Lancet* **374**, 1209–1211 (2009).
6. A Sadilek, HA Kautz, V Silenzio, Predicting disease transmission from geo-tagged micro-blog data in *Proc. of 26th AAAI Conference on Artificial Intelligence (AAAI)*. (2012).
7. A Signorini, AM Segre, PM Polgreen, The use of twitter to track levels of disease activity and public concern in the us during the influenza A H1N1 pandemic. *PloS one* **6**, e19467 (2011).
8. PM Polgreen, Y Chen, DM Pennock, FD Nelson, RA Weinstein, Using internet searches for influenza surveillance. *Clin. infectious diseases* **47**, 1443–1448 (2008).
9. J Ginsberg, et al., Detecting influenza epidemics using search engine query data. *Nature* **457**, 1012–1014 (2009).
10. C Li, et al., Retrospective analysis of the possibility of predicting the COVID-19 outbreak from internet searches and social media data, China, 2020. *Eurosurveillance* **25**, 2000199 (2020).

11. S Todd, PJ Diggle, PJ White, A Fearne, JM Read, The spatiotemporal association of non-prescription retail sales with cases during the 2009 influenza pandemic in Great Britain. *BMJ open* **4**, e004869 (2014).
12. D Balcan, et al., Multiscale mobility networks and the spatial spreading of infectious diseases. *Proc. Natl. Acad. Sci.* **106**, 21484–21489 (2009).
13. MC Gonzalez, CA Hidalgo, AL Barabasi, Understanding individual human mobility patterns. *Nature* **453**, 779–782 (2008).
14. M Lajous, et al., Mobile messaging as surveillance tool during pandemic (H1N1) 2009, Mexico. *Emerg. Infect. Dis.* **16**, 1488–1489 (2010).
15. FS Lu, et al., Accurate influenza monitoring and forecasting using novel internet data streams: A case study in the boston metropolis. *JMIR public health surveillance* **4** (2018).
16. CE Mills, JM Robins, M Lipsitch, Transmissibility of 1918 pandemic influenza. *Nature* **432**, 904–906 (2004).
17. GJ Rubin, R Amlôt, L Page, S Wessely, Public perceptions, anxiety, and behaviour change in relation to the swine flu outbreak: cross sectional telephone survey. *BMJ* **339** (2009).
18. M Lipsitch, , et al., Managing and reducing uncertainty in an emerging influenza pandemic. *New Engl. J. Medicine* **361**, 112–115 (2009).
19. G Sigmundsdottir, et al., Surveillance of influenza in iceland during the 2009 pandemic. *Eurosurveillance* **15**, 19742 (2010).
20. BM Althouse, et al., Enhancing disease surveillance with novel data streams: challenges and opportunities. *EPJ Data Sci.* **4**, 17 (2015).
21. Office of National Statistics, UK, Ons methodology working paper series no. 8- statistical uses for mobile phone data: literature review (2019) Accessed: 2019-03-07.
22. International Telecommunication Union, Global and regional ICT data (2005-2019) (2019) Accessed: 2020-03-16.
23. S Asongu, The impact of mobile phone penetration on African inequality. *Int. J. Soc. Econ.* (2015).
24. N Eagle, AS Pentland, D Lazer, Inferring friendship network structure by using mobile phone data. *Proc. Natl. Acad. Sci.* **106**, 15274–15278 (2009).
25. J Blumenstock, G Cadamuro, R On, Predicting poverty and wealth from mobile phone metadata. *Science* **350**, 1073–1076 (2015).
26. A Wesolowski, et al., Quantifying the impact of human mobility on malaria. *Science* **338**, 267–270 (2012).
27. CO Buckee, A Wesolowski, NN Eagle, E Hansen, RW Snow, Mobile phones and malaria: modeling human and parasite travel. *Travel. medicine infectious disease* **11**, 15–22 (2013).
28. L Bengtsson, et al., Using mobile phone data to predict the spatial spread of cholera. *Sci. reports* **5**, 8923 (2015).
29. M Tizzoni, et al., On the use of human mobility proxies for modeling epidemics. *PLoS Comput. Biol* **10**, e1003716 (2014).
30. CC Freifeld, et al., Participatory epidemiology: use of mobile phones for community-based health reporting. *PLoS medicine* **7**, e1000376 (2010).
31. SY Chang, et al., Mobility network modeling explains higher SARS-CoV-2 infection rates among disadvantaged groups and informs reopening strategies. *Nature* (2020).
32. N Ahmed, et al., A survey of COVID-19 contact tracing apps. *IEEE Access* **8**, 134577–134601 (2020).
33. SL Erikson, Cell phones $\neq$ self and other problems with big data detection and containment during epidemics. *Med. Anthropol. Q.* **32**, 315–339 (2018).
34. DM Halbfinger, I Kershner, R Bergman, To track Coronavirus, Israel moves to tap secret trove of cellphone data. *The New York Times* (2020).
35. A Amato-Gauci, et al., Surveillance trends of the 2009 influenza A (H1N1) pandemic in Europe. *Eurosurveillance* **16**, 19903 (2011).
36. N Kishore, et al., Flying, phones and flu: Anonymized call records suggest that Keflavik International Airport introduced pandemic H1N1 into Iceland in 2009. *Influ. other respiratory viruses* **14**, 37–45 (2020).
37. DKM Ip, et al., The dynamic relationship between clinical symptomatology and viral shedding in naturally acquired seasonal and pandemic influenza virus infections. *Clin. Infect. Dis.* **62**, 431–437 (2016).
38. NewZoo, Global mobile market report (2018) Accessed: 2019-03-16.
39. G King, et al., *Ecological Inference: New Methodological Strategies*. (Cambridge University Press, New York), (2004).
40. LL Lau, et al., Viral shedding and clinical illness in naturally acquired influenza virus infections. *The J. infectious diseases* **201**, 1509–1516 (2010).
41. J Patel, et al., Poverty, inequality and COVID-19: the forgotten vulnerable. *Public Heal.* **183**, 110 (2020).
42. Y Cao, M Yoshikawa, Y Xiao, L Xiong, Quantifying differential privacy under temporal correlations in *33rd Intl. Conf. on Data Engineering (ICDE)*. (IEEE), pp. 821–832 (2017).
43. DJ Mir, S Isaacman, R Cáceres, M Martonosi, RN Wright, Dp-where: Differentially private modeling of human mobility in *2013 Intl. Conf. on Big Data*. (IEEE), pp. 580–588 (2013).
44. Statistics Iceland, Telecommunication, penetration of cellular subscriptions per 1,000 inhabitants (Retrieved June 18, 2020) (2020).
45. Statistics Iceland, Population - inhabitants (Retrieved June 18, 2020) (2020).
46. S Altizer, et al., Seasonality and the dynamics of infectious diseases. *Ecol. letters* **9**, 467–484 (2006).
47. N Rojansky, A Brzezinski, JG Schenker, Seasonality in human reproduction: an update. *Hum. reproduction* **7**, 735–745 (1992).

Table S1. The CDR calls table contains 123,924,260 call-detail records between June 4, 2009 and February 11, 2010 in the following format.

Subject	Object	Timestamp	Incoming	Call/Text	Length (s)	Tower (Latitude, Longitude)
4197	89504	2010-05-06 16:07:24	True	Call	7	(65.23456, -18.0101)
98937	52674	2010-09-17 10:34:46	False	Text	0	(65.34567, -18.09876)
51993	607	2010-09-29 01:47:50	False	Call	25	(65.12345, -18.54321)

The first record (for example) reads: phone with ID 4197 is called by the phone denoted 89504 at 16:07 on May 6, 2010. The call lasted 7 seconds. The subject phone connected with the tower located at (65.23456,-18.0101) to receive the call.

Supplemental Information

1. Data Linking and Privacy

The Iceland National Bioethics Committee approved the following protocol of linking CDR and ILI-diagnosis data to ensure three privacy concerns are met: (i) that the MNO does not obtain health diagnosic information for its subscribers, (ii) that the CHS-CDC does not obtain CDR data, and (iii) that the researchers neither learn of original unencrypted phone numbers nor the NINs of individuals contained in the dataset. The process began with the MNO encrypting all the phone numbers with a secret, private key. Next, the MNO sent the encrypted numbers and the associated bill-paying NIN (recall that Icelandic NINs are public information, locally called *kennitala*) to the CHS-CDC. Third, the CHS-CDC merged the encrypted phone number and NIN data with their health data by joining on the NIN. Fourth, the CHS-CDC cryptographically hashed the NINs with a nonce (a disposable private key) to deidentify them. Finally, the CHS-CDC provided us with health data joined to the encrypted phone IDs, and the MNO provided us with CDR data containing encrypted phone IDs.

2. Mobile Network Data

CDR Metadata. Call-detail records were provided by an MNO in three installments as a database table in the aforementioned format (Fig. 1 and Table S1), containing $1,517,276,930$ records and spanning more than a three year period from February 1, 2009 to June 30, 2012. Each record in the table contains anonymous identifiers (subject and object) for the two phones involved in the call. The incoming field gives the direction of the call, *i.e.*, whether it was initiated by the subject. For every call where both phone numbers are MNO customers, the calls table contains two rows: one per constituent of the call. The two records will have the subject and object fields switched because the tower and incoming information are presented from the perspective of the subject phone ID. For legacy reasons, all CDR pertaining to text messages have an object ID corresponding to a fixed destination number (the short-message service center (SMSC) gateway) rather than the destination phone number of the text message. The features extracted from the text metadata thus cannot be used to make statements about text recipients.

GPRS Metadata. When mobile phones access the Internet, they use protocols that differ from voice and text (SMS). These accesses produce metadata logs that differ from CDR metadata. We were provided with General Packet Radio Service (GPRS) logs from the MNO that had been deidentified by the same process as the CDR data. These data capture only a timestamp, the encrypted subject phone number, tower location, and data transfer volume statistics. The GPRS logs spanned the same time period and contained $101,489,436$ records. No destination is recorded for Internet data usage entries.

3. Data Preprocessing

The cell tower coordinates pinpoint a unique GPS position (latitude/longitude); however, every cell tower base station contains three directional antennas that each span a horizontal sector of 120°. In the dataset, the MNO specified which of the three antennas was used for each interaction. Our early experiments suggested that the antenna orientation alone in the CDR data was inadequate to determine more granular positioning or trajectory of a subject. Among the contributing factors, we note that orientation, terrain, altitude, geography, and transmission power all affect the signal strength between a cellular device and a particular tower, and that the likelihood of interaction with a given cell antenna further depends on protocol (2G/3G/4G, voice call vs. GPRS data) and the load of nearby cell towers. We thus discard the orientation information and combine the three antennas into a single tower GPS coordinate during feature extraction.

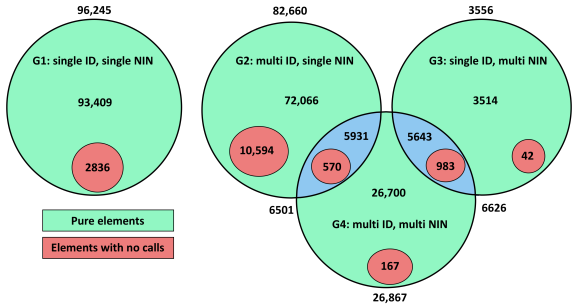


Fig. S1. *Data Preprocessing.* Venn diagram of the division of cell phone number groups, highlighting that some cell numbers have been paid for by a succession of multiple accounts, and that some subscribers pay for multiple cell phone numbers.

Cell phone numbers are inherently not bound to a single individual; they can change ownership or a single individual can own several phone numbers. Since the health data links the CDR data via the NINs, the question remains whether the person billed for a phone number is the same as the one diagnosed with ILI. Grouping based on de-identified NINs and phone IDs yields four different scenarios (Fig. S1). Two of the four groups allow us to match a specific diagnosis date with a NIN and the phone (or phones) specifically used by that one person (excluding any dependents).

Using (phone ID, NIN) pairs, the groups are:

- **G1** (single ID, single NIN): Every phone ID is paid for by one NIN, and every NIN in this group only pays for one phone number during the time period in the data. Thus, the phone numbers never change ownership in our period, and the person only pays for that one phone. Here, we assume that the person billed for the phone and the person using it day-by-day are one and the same.
- **G2** (multi ID, single NIN): Every phone ID is paid by one NIN, and every NIN in this group pays for more than one phone number. Families, small companies, and people with multiple phones for themselves (*e.g.*, personal and work) comprise this group, as well as those whose phone number changes during the time period.
- **G3** (single ID, multi NIN): Every phone ID is paid by multiple NINs during the time period, but each NIN only pays for one phone number. The tariff billing codes that accompany the CDR data indicate that the majority of these phones are company phones that switch between NINs regularly.
- **G4** (multi ID, multi NIN): Every phone ID is paid by multiple NINs during the time period, and each NIN pays for multiple phones. This group has large crossover with G2 and G3.

G1, the largest group, contains over 93,000 one-to-one pairs of phone numbers and NINs, which we assert with high confidence represent the same individual. G1 also is disjoint from the other

groups which themselves overlap (Fig. S1). The members of G1 who were diagnosed with ILI comprise our dataset. Future work could incorporate the progressively larger and more representative datasets: the entire G1 set, G1 combined with those phone numbers in G2 that can be inferred to belong to individuals, and all groups. We refer to G1 as the "dataset" in the paper.

Characterizing Bias. We considered the scope of sampling biases in the data sets to ascertain whether the dataset represents the population. There were three major mobile network operators in Iceland in 2009, each surpassing 25% market share. That year, the penetration rate of cell phone subscriptions was 1,039 per 1,000 people (44). The Icelandic population in 2009 was homogeneous with no identified socioeconomic factors that would significantly affect cell phone ownership. There was minimal vendor lock-in for two reasons: (1) binding contracts or termination penalties were uncommon, (2) the porting of cellular numbers between the Icelandic mobile phone providers was simple. All three MNO providers provided 2G or better cell coverage in both rural and urban areas, covering virtually all residences in Iceland through their own tower infrastructure or roaming partners. Residential cell coverage was thus not a driving factor in the dispersion of subscribers across mobile network providers. We compared population data per postal code in Iceland according to the 2009 census (45) with the inferred home tower location, and corresponding postal code, of subscribers in our dataset. We found the correlation to be strong for both the dataset ($r = 0.86$, $p < 8 \times 10^{-49}$) and among those with ILI diagnosis ($r = 0.88$, $p < 2 \times 10^{-43}$), suggesting that the data set is spatially representative of the entire adult Icelandic population. However, a source of bias in our dataset is the disproportionately small number of children and teenagers (under 18 years of age) who own and pay for their cell phone subscription. For privacy reasons, we were not permitted to use age information of either subscribers or those diagnoses with ILI as part of this study, and we are therefore unable to quantify this bias.

4. Feature Extraction

The choice and representation of features from the data can substantially impact how well behavioral changes can be quantified. We extract a range of features encompassing several aspects of human behavior (Table S3) that span three categories:

- **Mobility:** Conventional wisdom suggests that individuals tend to stay home when they exhibit symptoms of flu. To study this intuition, we define three features that focus on human mobility: *Unique Locations Visited*, *Distance Traveled*, and *New Locations Visited* all use the spatiotemporal attributes of the data to estimate movement of individuals.
- **Social:** Sickness can affect the number of interactions with other people. We calculated *Unique Contacts* and *New Contacts* to identify such variability. Social features use the notion of "top" contacts, which are the K highest ranked contacts of an individual when sorted by either their total call duration or call frequency over the past 30 days. Symptomatic people might prefer to communicate with their closer contacts than others while exhibiting symptoms. The features *Calls to Top Contacts* and *Calls to Non-Top Contacts* help evaluate this tendency.
- **Activity:** Symptomatic individuals may also alter their phone usage. The activity features extract the number of interactions (calls or text) along with the call duration to determine if individuals' usage pattern statistics change when sick. The features are *text and call count*, *total call duration*, and *mean call duration*.

Many of these features have both incoming and outgoing variants, resulting in a total of 36 extracted features (Table S2). All feature extraction functions return a single decimal number for a given time interval. The functions take in a user, day, and bin to compute a feature value for the time interval defined by the bin.

Features were selected to be simple and interpretable. Investigating further feature representations remains future work and includes considering the contact relationship strength, the proximity to popular locations, and analysis of the GPRS (and corresponding 3G/4G/5G protocols) data upload and download patterns.

5. Sensitivity Analysis

The cut-off of $K = 3$ for "top" contacts by call duration and number of calls (interactions) was chosen as follows. For each user in a 100-person sample of diagnosed individuals, we aggregated the total duration of calls and number of calls with each of their contact over and 8-week period. We found that on average, the percentage of calls and call time to the three highest ranked contacts was high, or over 40% of interactions as shown in Figure S4.

6. Seasonality

Alternate Approach for Detrending Weekly Routine Behavior. We considered a different approach for removing the weekly seasonality in our data, which we call *normalized fractional rankings*.

As before, we use $x_{fid} = E_f(i, d)$ to denote the raw feature value for a feature f, extracted from the CDR by function E, for individual i, and on day d.

To control for the weekly behavioral routine of individual i, each feature value is ranked among those of the same weekday over the past 10 weeks, then normalized to the range $[0, 1]$. Specifically, we rank x_{fid} for feature f on day d within the bag (multi-set) PREV_{fid} where

$$\text{PREV}_{fid} = \{x_{fi,(d-7j)} \mid j = 0, 1, \ldots, 10\}$$

is the behavior on the same day of the week from the previous 10 weeks. The resultant rank r_{fid} lies between 1 (lowest value on d compared to past weeks) and 11 (highest), breaking ties with the average of spanned ranks. Finally, the rank is linearly mapped to the normalized fractional rank

$$v_{fid} = \frac{r_{fid} - 1}{|\text{PREV}_{fid}| - 1} \in [0, 1].$$

For example, suppose user i received $x_{fid} = 300$ seconds of calls (f) on a given Tuesday d and had $\text{PREV}_{fid} = \{300, 160, 300, 700\}$ seconds of incoming calls on this and the previous three Tuesdays. The rank of 300 in the sorted list $(160, 300, 300, 700)$ is 2.5, hence the fractional rank v_{fid} is

$$v_{fid} = \frac{2.5 - 1}{|\{300, 160, 300, 700\}| - 1} = \frac{1.5}{4 - 1} = 0.5,$$

bearing in mind that multi-sets can contain multiple copies of an element.

In a similar fashion as detrending using simple linear regression, our ranked feature values normalize for a cyclical weekly pattern. Accordingly, the ranked feature values for the control group, normalized to the range $[0, 1]$, could be expected to hover around the median value of 0.5 each day in the limit of large groups. Since the values are normalized based on 10 weeks of activity, however, and four weeks are displayed, it is possible for latent longer term trends to drive the average away from 0.5 (Fig. S10-S27). An advantage of normalized fractional ranking over linear regression is the universal limited range across all feature types (between 0 and 1), making the metric useful as input into machine learning applications, such as the epidemic curve generated below (SI:Privacy-Preserving Data Sharing). Our results also show that normalized fractional ranking distills the deviation signal more cleanly than deviation from weekly average, as evidenced below by the side-by-side comparisons of raw feature values, normalized fractional ranking, and the deviation from weekly average through linear regression (Fig. S10-S27). A disadvantage of normalized fractional ranking, however, is the loss of magnitude dimension for the feature which can impact interpretability.

Other Seasonality. Seasons affect people's habits (46, 47); we observe seasonal behavior and cyclical weekly pattern in the extracted feature values (Fig. S6, the colder 10 weeks average fewer *unique locations visited* than the previous warmer 10 weeks), as well as in the corresponding fractionally ranked and normalized values (Fig. 2, S5).

Major holidays such as Christmas, New Year's or Easter can drastically impact a model. The behavior of a large sample shows major correlated fluctuations in people's behavior during the 2009 winter holidays (Fig. S6).

Table S2. ***Full Feature Characteristics*** **from the 29-day period around each individual's DoD and corresponding control. Most features exhibit a right-skewed distribution (Fig. S2, left column) and share general characteristics across control and diagnosis groups. The rightmost column indicates which days around diagnosis have a statistically significant p-value from performing the Kolmogorov–Smirnov test with a Benjamini-Hochberg correction on the control and diagnosed groups ($\alpha = 0.05$).**

Feature		Control					Diagnosed					Significant Days
		Min	Max	Mean	Median	Mode	Min	Max	Mean	Median	Mode	
Unique Locations Visited		0	27	3.04	2	1	0	33	2.74	2	1	-12, -2 to 4
Distance Traveled		0	1011	22.57	5.47	0	0	1515	19.6	3.69	0	-2 to 4, 13
Number of New Location Visited		0	18	0.5	0	0	0	26	0.43	0	0	-2, -1, 1, 2, 3
Unique Contacts	incoming	0	27	2.25	2	0	0	39	2.021	2	1	∅
	outgoing	0	52	2.50	2	0	0	29	2.277	2	1	1, 2, 3, 6
	both	0	53	4.04	3	0	0	57	3.67	3	1	2
New Contacts	incoming	0	20	0.61	0	0	0	13	0.50	0	0	∅
	outgoing	0	45	0.65	0	0	0	20	0.55	0	0	1, 3, 6
	both	0	46	1.19	1	0	0	24	1	1	0	∅
Call Duration	incoming	0	18558	475	190	0	0	23461	480.51	174	0	0, 1, 2
	outgoing	0	11564	460	162	0	0	26553	435.09	135.5	0	0
	both	0	22687	935	479.5	0	0	32989	915.60	461	0	0
Calls Count	incoming	0	45	3.10	2	0	0	56	2.84	2	1	∅
	outgoing	0	152	3.60	2	0	0	181	3.37	2	0	1, 2
	both	0	152	6.66	5	0	0	188	6.22	4	1	1
Texts Count	incoming	0	253	2.71	1	0	0	123	2.87	1	0	-10
	outgoing	0	127	1.74	0	0	0	121	1.93	0	0	∅
	both	0	380	4.456	1	0	0	237	4.79	1	0	-10, 11
Calls & Texts count	incoming	0	254	5.779	4	0	0	130	5.71	3	1	∅
	outgoing	0	153	5.341	3	0	0	226	5.30	3	1	1, 2
	both	1	381	11.120	7	0	1	259	11.008	7	1	2
Mean Call Duration	incoming	0	6270	133.35	72.33	0	0	5140	140.74	70.38	0	-1, 0, 1, 2, 4, 11
	outgoing	0	4370	107.96	57.53	0	0	5980	106.74	50.50	0	∅
	both	0	4692	143.14	93.85	0	0	5095	149.84	94.25	0	-8, -1 to 4, 11, 13, 14
Top-3 Contacts by Duration (SI:Sensitivity Analysis)	incoming	0	3	0.68	1	0	0	3	0.69	1	0	∅
	outgoing	0	3	0.70	1	0	0	3	0.69	1	0	1
	both	0	6	1.37	1	0	0	6	1.38	1	0	∅
Remaining Contacts by Duration	incoming	0	26	1.57	1	0	0	37	1.33	1	0	∅
	outgoing	0	51	1.80	1	0	0	27	1.58	1	0	1, 2, 4, 6
	both	0	52	3.37	2	0	0	59	2.92	2	1	∅
Top-3 Contacts by Frequency	incoming	0	3	0.58	0	0	0	3	0.56	0	0	∅
	outgoing	0	3	0.63	0.5	0	0	3	0.61	0	0	1, 2
	both	0	6	1.21	1	0	0	6	1.17	1	0	1
Remaining Contacts by Frequency	incoming	0	26	1.67	1	0	0	37	1.46	1	0	∅
	outgoing	0	51	1.88	1	0	0	26	1.67	1	0	-9, 2, 6
	both	0	52	3.54	2	0	0	59	3.13	2	1	∅

Table S3. ***Feature Descriptions.*** **Details of how each feature is extracted. The function producing each feature outputs a single decimal number.**

Type	Name	Description
Mobility	Unique Location Visited	The number of unique tower locations to which the phone connects.
	Distance Traveled	The total distance (km) between the towers to which the phone connects chronologically.
	New Locations Visited	The number of unique tower locations to which the phone connects and which were not seen in the previous 30 days.
Social	Unique Contacts	The number of unique contacts who interacted with this phone.
	New Contacts	The number of unique contacts who interacted with this phone, but had not in the past 30 days.
	Calls to Top Contacts	The number of call interactions with the top $K = 3$ contacts.
	Calls to Non-Top (Remaining) Contacts	The number of call interactions for contacts besides the top K.
Activity	Text and Call Count	Number of texts and calls.
	Total Call Duration	Total duration of all calls.
	Mean Call Duration	Average duration of all calls.

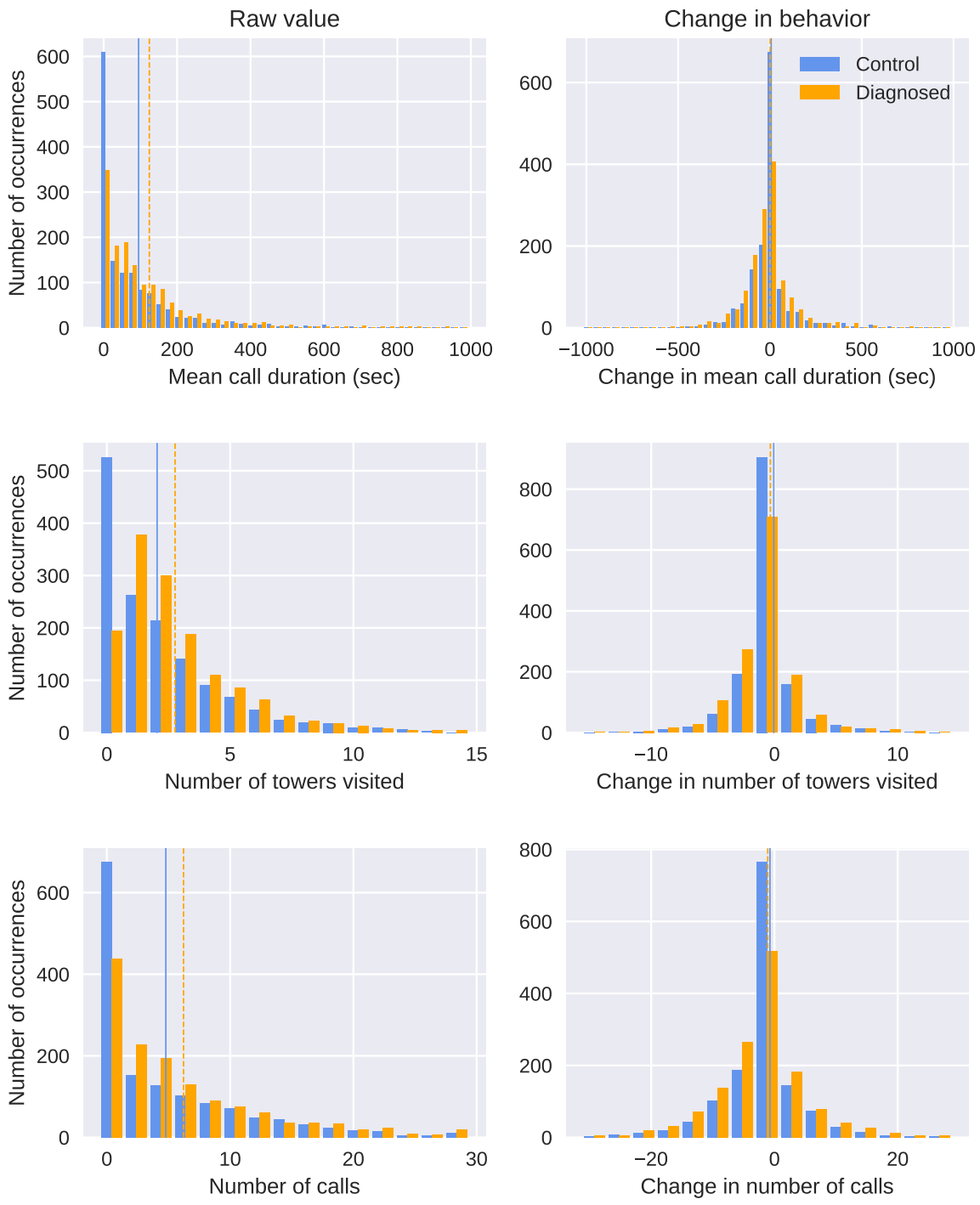


Fig. S2. Full distributions of three features six days before diagnosis (DoD-6). **Left:** Raw feature values (x_{fid}) for the three most significant features for the diagnosed (orange) and control groups (blue). Features predominantly follow right-skewed and heavy-tailed distributions (characteristics in Table S2). **Right:** The detrended change in behavior is shown for comparison. The averages for the diagnosed and control groups (dashed vertical lines) are closer than for the raw values.

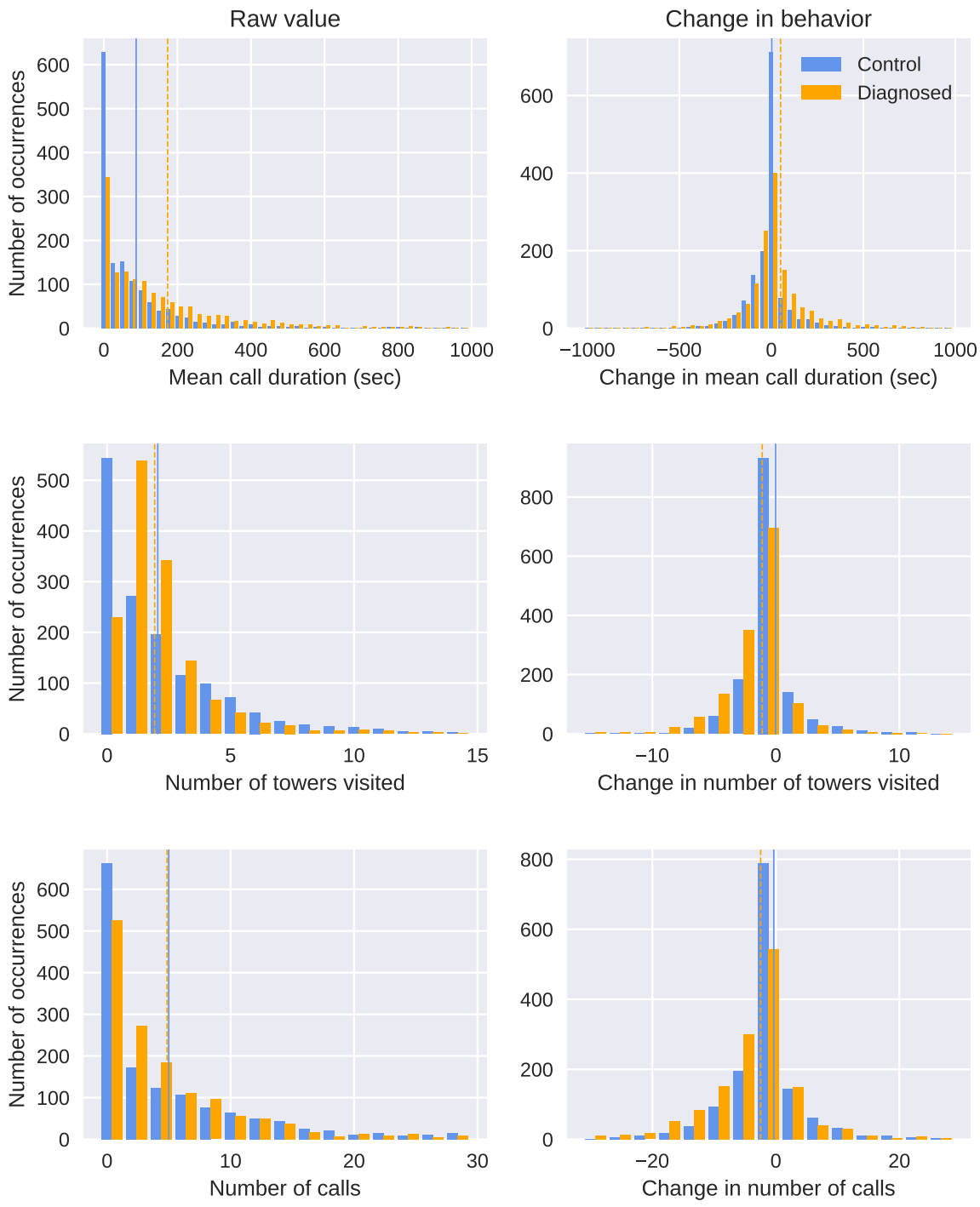


Fig. S3. Full distributions of three features the day after diagnosis (DoD+1). **Left:** Raw feature values (x_{fid}) for the three most significant features for the diagnosed (orange) and control groups (blue). **Right:** The detrended change in behavior is shown for comparison. The averages for the diagnosed and control groups (dashed vertical lines) are now further apart than on six days prior to diagnosis (S2).

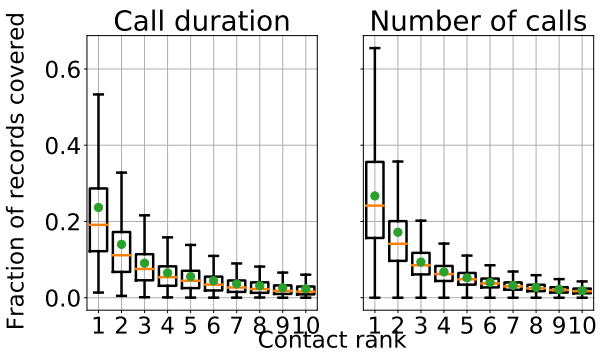


Fig. S4. Percentage of calls with each top contact by rank, measured by total duration of calls (left) and total number of calls (right), over an 8-week period for a random sample of 100 diagnosed individuals.

7. Comparing CDR and GPRS Data

Both the CDR and GPRS data provides a proxy for location about the subject through the tower location. The who-calls-whom relationship in CDR data encapsulates a deidentified social network of users whereas GPRS data acts as a location proxy. On the other hand, GPRS entries are made whenever the mobile device makes an Internet connection, for example when an application periodically checks for new e-mail in the background, providing location information even when the user is not actively using the device.

As mobile users increasingly make voice, video calls and exchange messaging via Internet applications such as WhatsApp, Skype, or Facebook Messenger using the mobile data services (GPRS for short) for rather than the via the native mobile network voice and text (SMS) protocols, CDR data becomes sparser (fewer calls) whereas the GPRS data becomes richer (more data connections). We measured the extent to which these two data sources complement one another as follows.

Define the Lin's divergence of two probability distributions $P, Q : \mathcal{X} \to [0, 1]$ as

$$D_L\left(P \parallel Q\right) = \sum_{x \in \mathcal{X}} P(x) \log\left(\frac{2P(x)}{P(x) + Q(x)}\right).$$

The metric takes values in $[0, 1]$, with $D_L\left(P \parallel Q\right) = 0$ when $P \equiv Q$ almost surely. This last property is an advantage over the traditional Kullback-Liebler (KL) divergence, which requires absolute continuity ($Q(x) \neq 0$ whenever $P(x) \neq 0$),

We considered the interval of data between 1 January and 15 February 2010. Let $a_u(t, x)$ denote the probability that user u interacts with cell tower at location x at time of day t. Here, towers are grouped by the unique GPS location of their GSM masts as before. Time is measured in 1-hour intervals. We further subdivide a_u into c_u and d_u where $c_u(t, x)$ is the probability of interaction according to CDR call data, and $d_u(t, x)$ is the same probability according to GPRS data-connection records for that user.

We calculated the average $D_L(c_u \parallel a_u)$ and $D_L(d_u \parallel a_u)$ for all users and plot the results in Figure S7. GPRS logs provided mostly redundant location information except during the night hours (4-6am) relative to CDR call data in our dataset. Advances in smartphone technology coupled with more affordable and available cellular data access are likely to affect this result.

8. Privacy-Preserving Data Sharing

Model and Assumptions. Syndromic surveillance based on models and information from diverse sources requires a careful data collection architecture to ensure maximal privacy for the subjects. An experimental validation or a later practical implementation for monitoring epidemics using CDR data could hypothetically support two modes of operation:

O1 the standard monitoring mode for health officials, producing interactive charts of the epidemic curve in real-time using already trained models, and

O2 the ability to calibrate the models based on new labeled data, for instance when supporting or monitoring diseases with different profile of symptoms.

Whether as part of an experimental design or a full-fledged implementation, both of these modes should maintain two invariants to ensure privacy. First, while we assume mobile operators already have access to cell phone data and metadata, they should not be privy to any new health information of their customers or small groups of users. Second, health officials should not be able to glean spatiotemporal information of individuals or small groups, including from any aggregate call data statistics or epidemic charts.

Protocol Design. At a high-level, our architecture consists of three kinds of entities: MNOs, health officials, and an independent 3rd party broker (Fig. 3). We assume multiple MNOs may intend to contribute streams of CDR data to the trusted 3rd party broker. The broker is then responsible for computing features from the CDR data and running our models, and then generating interactive epidemic curves, either for experimental analysis and validation, or directly for use by the health officials. In practice, the broker could be operated by a university or other non-profit, or via an administratively separate joint venture between the disease control and MNOs.

Before streaming CDR data to the broker, the MNOs take a one-way hash of identities (subject and object numbers) in a similar fashion as was done when preparing the dataset in this paper. Specifically, each MNO i maps a customer phone number a by first combining it with a secret key s_i and then a strong cryptographic hash function (such as SHA-3, which is not susceptible to length-extension attacks). The customer identified by phone number a is thereafter only known to the broker by the identifier $\texttt{sha3}(s_i \| a)$ which we will write $s_i(a)$. Note that CDR from when customers of two different MNOs call one another cannot be compared, as the records will have incompatible identifiers under this design (unless the MNOs all share the same secret key). Thus, the social network over the clients will be partitioned between MNOs; this is unlikely to be a problem for outbreak surveillance.

To enable retraining of the model in light of new data (**O2**), we must provide a mechanism for assigning labels based on health information to phone numbers without violating our privacy invariants. In our design, the health officials also provide each MNO i with a designed secret key σ_i (Fig. 3, right). These keys are not known by the broker. Each MNO i regularly provides the broker with an up-to-date mapping $(\sigma_i(a), s_i(a))$ between the two hashes of the phone number a. The hashing functions are non-invertible and so the broker does not learn anything about any phone number a from these maps. When a health official now wants to create a label for the person with phone number a, which is known through a separate look-up mechanism to have cell phone service with MNO i, the health officials submit the identifier $\sigma_i(a)$ to the broker with accompanying health information (such as date of diagnosis) for the label. The broker is now able to use the labeled data to retrain and calibrate the epidemic curve models as needed. The MNO has not been exposed to any health information about their customers; the broker has not learned any identities or phone numbers, and the scientists or health officials have not been provided spatiotemporal or other personal information about any user.

Differential Privacy. In this set-up, the scientists or health officials are only provided aggregate statistics and thus learn nothing about individual or small numbers of subjects. Aggregation is not sufficient: if health officials are allowed to interact with the data, such as to study particular demographics or areas, the small number of individuals studied by such queries might compromise their identities. For example, a particular rural area might have only one person in the specific age range, which would give officials unjustified insights into their movements and private affairs. The study of differential privacy mitigates precisely these kinds of vulnerabilities, allowing queries to be run for smaller groups of people without unduly compromising user privacy. The underlying idea is to add appropriate noise to the data being queried based on the specificity of the query while retaining the aggregate statistics, even in spatiotemporal settings (1, 2). The implementation of differential privacy for CDR data should account for combination and correlation of aggregate information and statistics with other data

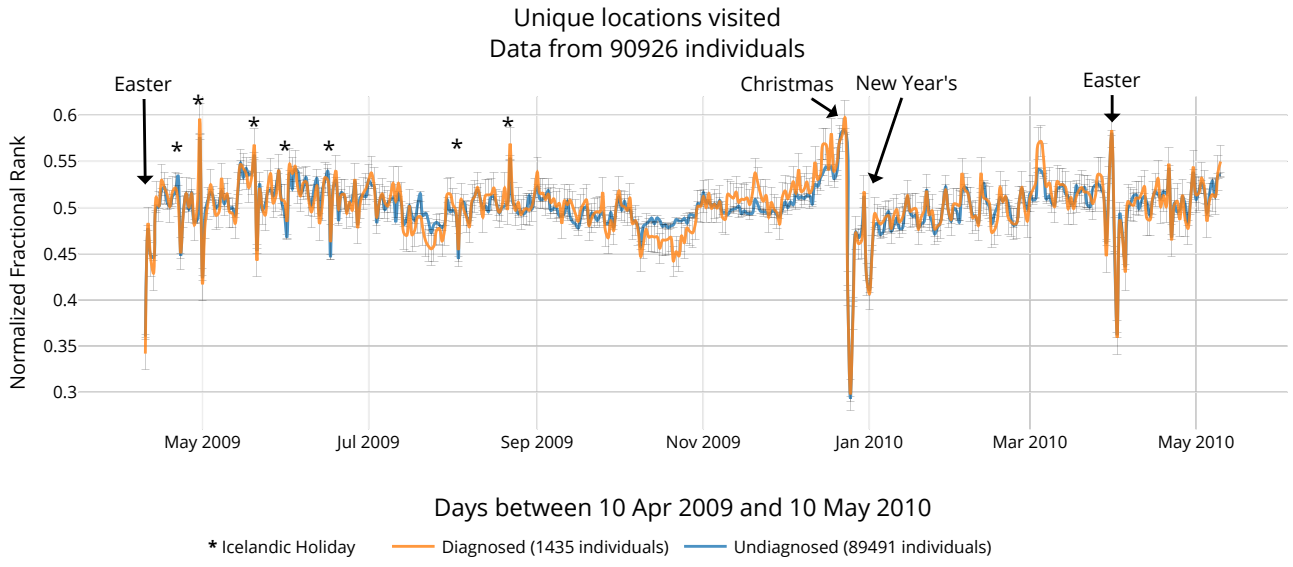


Fig. S5. Correlated anomalies over time for the diagnosed and entire undiagnosed group (G1). Holidays impact people's regular routines: movement inferred from distinct locations visited drops drastically during both winter holidays (last week of December 2009) and Easter (first week of April 2010). Surveillance systems need be robust to large-scale correlated changes of patterns, such as holidays; our model demonstrates robustness in this sense. The diagnosed and control (here, undiagnosed) groups move in tandem, except during the peak of the H1N1v influenza epidemic in October 2009 where they display a pronounced separation.

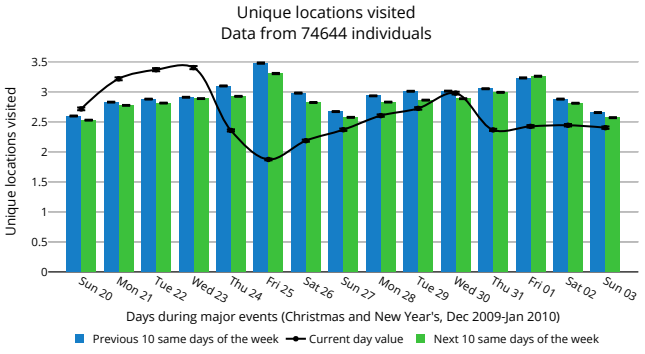


Fig. S6. Mean number of unique locations (towers) visited for 74,644 people before, during, and after winter holidays of 2009, a major vacation time in Iceland. The graph shows mean feature values (x_{fid}) for the entire group before individual ranking and normalization. The averages exhibit weekly periodicity, with Fridays tending to have highest activity. Each day's value is compared with averages from the same weekday in the prior (blue bars) and following (green bars) 10 weeks, showing fewer distinct locations being visited in the colder months. Christmas Eve (Dec 24) and Christmas Day (Dec 25) have lower cellular activity and movement than typical, whereas Dec 23 has greater activity than normal for a Wednesday. Error bars show the 99.9% confidence intervals.

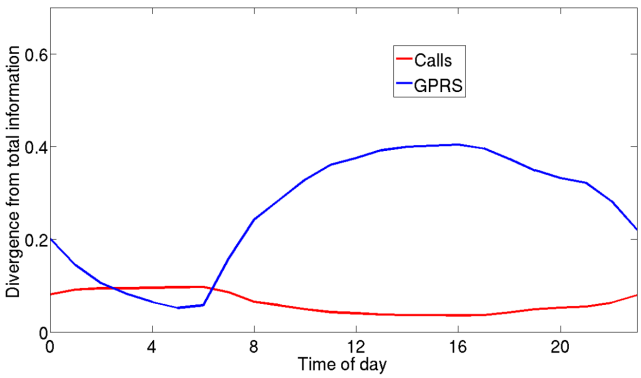


Fig. S7. Average Lin divergence $D_L(c_u \parallel a_u)$ and $D_L(d_u \parallel a_u)$ for per-hour location information in CDR call data (red) and GPRS data-access records (blue), respectively, from all available location data across all users between 1 January 2010 and 15 February 2010. The GPRS data provides useful location information during the night, whereas CDR data are more informative at other hours.

sources, including recent attacks such as trajectory recovery (3) and membership inference (4).

9. Visualizations

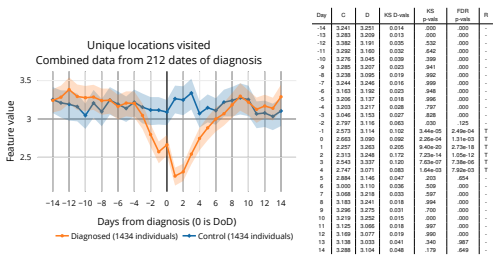
The behavioral changes are visualized through two families of graphs: the extracted feature plots (Fig. S6) and deviation from routine plots (Fig. S10-S27). The former illustrate the mean trend for the entire population sample while the latter is normalized on an individual basis and instead shows the average normalized values. We observe people's behavioral changes depend on the day of the week (Fig. S6); for instance, Sundays are on average considerably less active in terms of movement and call activity than Fridays. Given the focus on change in routine behavior, all the visualizations use the preceding 10 weeks (on the same day of the week) to either normalize by (Fig. S10-S27) or to plot as a reference (Fig. S6). Some plots display raw feature values immediately after extraction (Figs. S8, S2 and S3) while others compare values after either detrending through linear regression or fractional ranking and normalization (Figs. S10–S27).

For most features, the underlying data do not follow a normal distribution. Accordingly, the error bars show confidence intervals computed through bias-corrected and accelerated (BCa) bootstrapping with 10,000 samples.

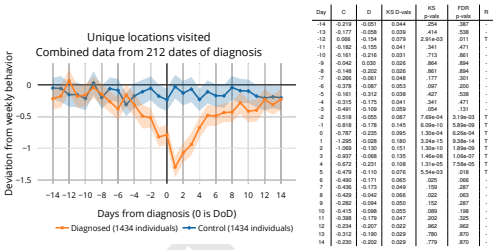
Statistical Significance. Because of the variability in feature distribution types, we choose the non-parametric two-sample KS test for comparing the diagnosed and control groups after fractional ranking and normalization. For Fig. S10-S27, an accompanying significance table is depicted alongside the line plots (Table S4 is a more verbose example). Here, *Day* is the day relative to the DoD and corresponds to the horizontal axes of the line plots. Recall that in these comparisons, each control is aligned at the same DoD as their diagnosed counterpart. The control (C) and diagnosed (D) columns show the average value of the feature in the respective group, either raw, detrended, or normalized after fractional ranking. Each *KS* test returns a D-value representing the test statistic and a p-value for the the given feature and *Day*. FDR shows the corrected p-values after the BH multiple test correction. *Reject* (R) is a Boolean variable that indicates whether the null hypothesis of the KS-test should be rejected based on a significance level of α (set as default value of $\alpha = 0.05$).

Table S4. Template for the significance tables used in Figs. S10-S27.

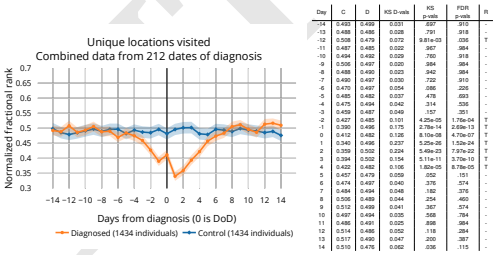
Day	Control	Diagnosed	KS D-vals	KS p-vals	FDR p-vals	Reject
⋮	⋮	⋮	⋮	⋮	⋮	⋮



(a) Raw feature value.



(b) Deviation from weekly average.

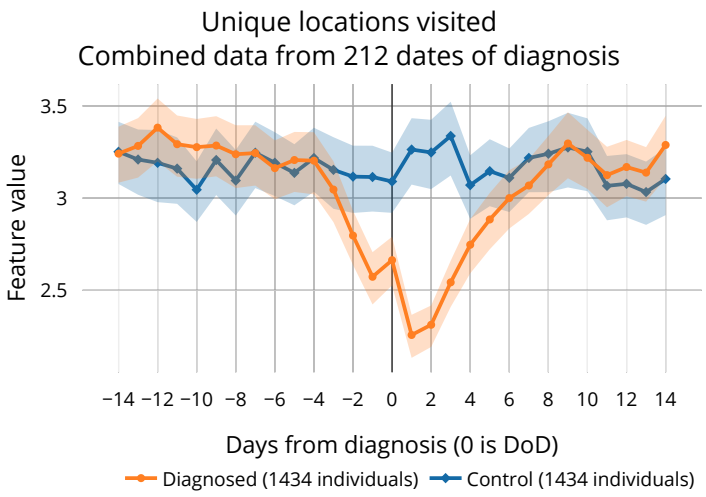


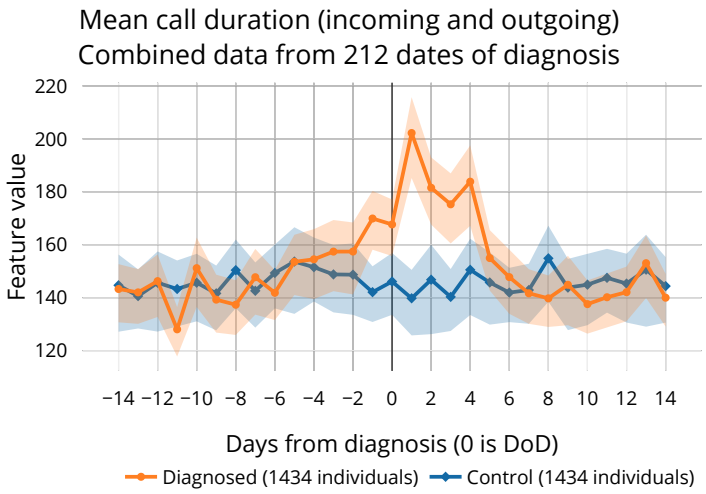
(c) Normalized fractional ranking.

Fig. S10. Feature graph and significance test results for ***unique locations visited*** using raw, detrended, and ranked data.

SI References

1. Chen, R., Fung, B., Desai, B. C., and Sossou, N. M. (2012) Differentially private transit data publication: a case study on the Montreal transportation system. In *18th ACM SIGKDD Intl. Conf. on Knowledge Discovery and Data Mining (KDD)* ACM pp. 213–221.
2. Xiao, Y. and Xiong, L. (2015) Protecting locations with differential privacy under temporal correlations. In *22nd ACM SIGSAC Conf. on Computer and Communications Security* ACM pp. 1298–1309.
3. Xu, F., Tu, Z., Li, Y., Zhang, P., Fu, X., and Jin, D. (2017) Trajectory Recovery From Ash: User Privacy Is NOT Preserved in Aggregated Mobility Data. In *Proceedings of the 26th International Conference on World Wide Web* Republic and Canton of Geneva, Switzerland: International World Wide Web Conferences Steering Committee WWW '17 pp. 1241–1250.
4. Pyrgelis, A., Troncoso, C., and Cristofaro, E. D. (2018) Knock Knock, Who's There? Membership Inference on Aggregate Location Data. In *25th Annual Network and Distributed System Security Symposium, NDSS 2018, San Diego, California, USA, February 18-21, 2018* The Internet Society.





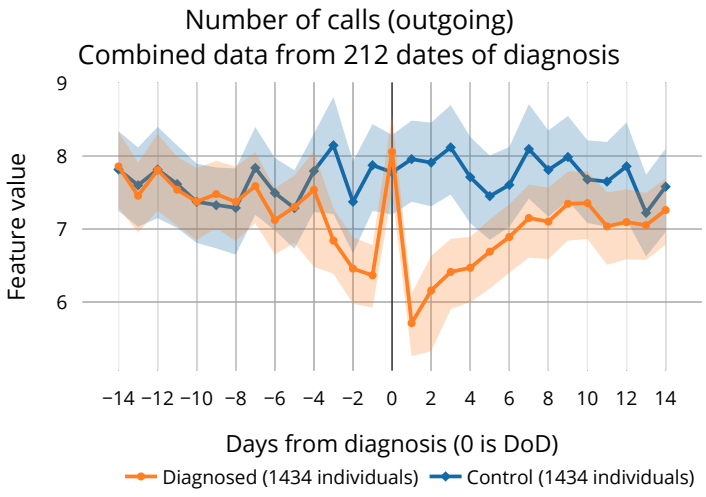


Fig. S8. Raw feature values (x_{fid}) for the three most significant features. This data is transformed into the detrended data (Fig. 2) upon which the statistical testing is performed. These raw values further defend the statement that the diagnosed and control groups display different patterns around DoD by showing the effect sizes without ranking and normalized transformation. Error bars (2.5% to 97.5%) are generated via bootstrapping.

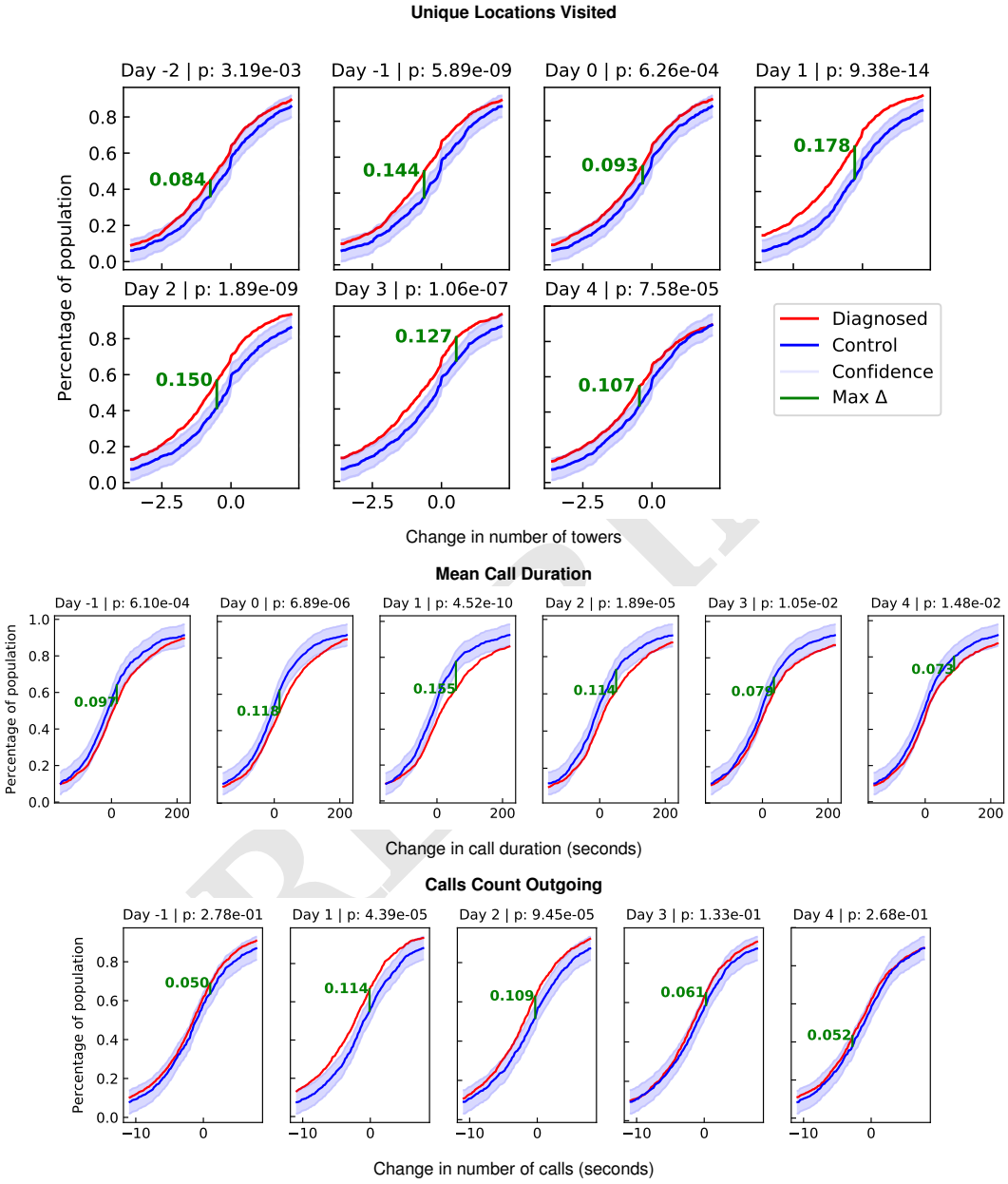
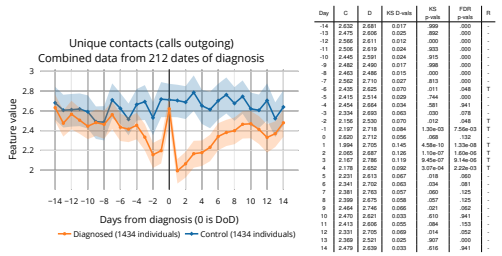
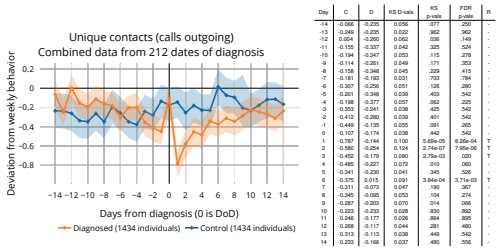


Fig. S9. Effect sizes for the main features. KS confidence bands for the cumulative distribution functions of the changes in unique locations visited (top), mean call duration (center), and number of outgoing calls (bottom) on each significant day relative to DoD, comparing the diagnosed group with a control group. Each day around DoD determined to be significant by the KS test is shown (*cf.* Table S2) and FDR-corrected p-value is listed in each graph title. A critical value D_α was calculated using the formula $D_\alpha = \sqrt{\frac{1}{2} \cdot \ln \frac{2}{\alpha} \frac{n+m}{nm}}$ where n and m are the sample sizes for the control and diagnosed groups and $\alpha = 0.05$. The confidence band around each control group is obtained by adding $\pm D_\alpha$ to the control group line.



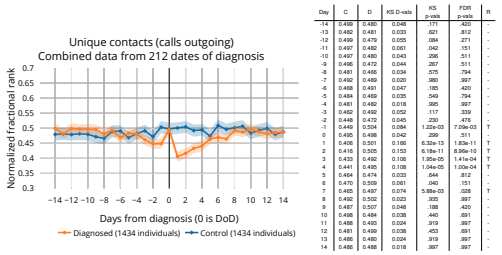
Day	C	D	KS D-vals	KS p-vals	FDR p-vals	R
-14	2.632	2.681	0.017	.999	.000	-
-13	2.473	2.606	0.025	.892	.000	-
-12	2.566	2.611	0.012	.000	.000	-
-11	2.506	2.619	0.024	.933	.000	-
-10	2.445	2.591	0.024	.915	.000	-
-9	2.482	2.490	0.017	.998	.000	-
-8	2.463	2.486	0.015	.000	.000	-
-7	2.562	2.710	0.027	.813	.000	-
-6	2.435	2.625	0.070	.011	.048	T
-5	2.415	2.514	0.029	.744	.000	-
-4	2.454	2.664	0.034	.581	.941	-
-3	2.234	2.693	0.063	.030	.078	T
-2	2.156	2.530	0.070	.012	.048	T
-1	2.137	2.718	0.094	1.30e-03	7.58e-03	T
0	2.630	2.712	0.056	.068	.132	-
1	1.994	2.705	0.145	4.58e-10	1.33e-08	T
2	2.065	2.687	0.126	1.10e-07	1.60e-06	T
3	2.167	2.786	0.119	5.45e-07	9.14e-06	T
4	2.178	2.652	0.092	3.07e-04	2.22e-03	T
5	2.231	2.613	0.067	.018	.060	-
6	2.341	2.702	0.063	.034	.081	-
7	2.381	2.763	0.057	.060	.125	-
8	2.398	2.675	0.058	.057	.125	-
9	2.464	2.746	0.066	.021	.062	-
10	2.476	2.631	0.033	.610	.941	-
11	2.413	2.606	0.055	.094	.153	-
12	2.331	2.705	0.069	.014	.052	-
13	2.369	2.521	0.025	.907	.000	-
14	2.479	2.639	0.033	.616	.941	-

(a) Raw feature value.



Day	C	D	KS D-vals	KS p-vals	FDR p-vals	R
-14	-0.166	-0.235	0.059	.077	.259	-
-13	-0.249	-0.235	0.022	.962	.962	-
-12	0.004	-0.260	0.062	.026	.149	-
-11	-0.155	-0.337	0.042	.325	.524	-
-10	-0.194	-0.347	0.053	.115	.278	-
-9	-0.114	-0.261	0.049	.171	.353	-
-8	-0.158	-0.348	0.045	.229	.415	-
-7	-0.181	-0.193	0.031	.703	.784	-
-6	-0.307	-0.256	0.051	.126	.280	-
-5	-0.201	-0.348	0.039	.403	.542	-
-4	-0.198	-0.377	0.057	.062	.225	-
-3	-0.353	-0.241	0.038	.425	.542	-
-2	-0.412	-0.280	0.039	.401	.542	-
-1	-0.449	-0.135	0.055	.091	.265	-
0	-0.107	-0.174	0.038	.442	.542	-
1	-0.787	-0.144	0.100	5.69e-05	8.26e-04	T
2	-0.580	-0.254	0.124	2.74e-07	7.95e-06	T
3	-0.452	-0.179	0.080	2.79e-03	.020	T
4	-0.465	-0.227	0.072	.010	.060	-
5	-0.341	-0.230	0.041	.345	.526	-
6	-0.375	-0.015	0.091	3.84e-04	3.71e-03	T
7	-0.311	-0.073	0.047	.190	.387	-
8	-0.345	-0.095	0.053	.104	.274	-
9	-0.287	-0.203	0.070	.014	.066	-
10	-0.223	-0.233	0.028	.830	.892	-
11	-0.246	-0.177	0.026	.864	.895	-
12	-0.268	-0.117	0.044	.281	.480	-
13	-0.313	-0.113	0.038	.449	.542	-
14	-0.233	-0.166	0.037	.480	.556	-

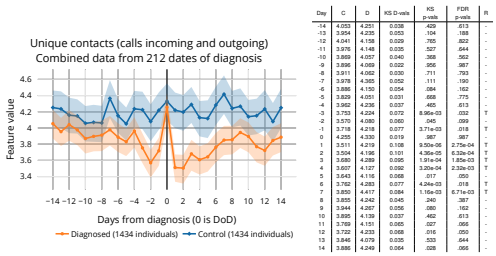
(b) Deviation from weekly average.



Day	C	D	KS D-vals	KS p-vals	FDR p-vals	R
-14	0.499	0.480	0.048	.171	.420	-
-13	0.482	0.481	0.033	.621	.812	-
-12	0.499	0.479	0.055	.094	.271	-
-11	0.497	0.482	0.061	.042	.151	-
-10	0.497	0.480	0.043	.296	.511	-
-9	0.496	0.472	0.044	.267	.511	-
-8	0.481	0.485	0.034	.575	.794	-
-7	0.492	0.489	0.020	.980	.997	-
-6	0.488	0.491	0.047	.185	.420	-
-5	0.494	0.469	0.035	.549	.794	-
-4	0.483	0.482	0.018	.995	.997	-
-3	0.462	0.492	0.052	.117	.339	-
-2	0.448	0.472	0.045	.239	.476	-
-1	0.440	0.504	0.084	1.22e-03	7.09e-03	T
0	0.495	0.498	0.042	.299	.511	-
1	0.408	0.501	0.166	6.32e-13	1.83e-11	T
2	0.416	0.505	0.153	6.18e-11	8.96e-10	T
3	0.433	0.492	0.106	1.95e-05	1.41e-04	T
4	0.441	0.495	0.108	1.04e-05	1.00e-04	T
5	0.464	0.474	0.033	.644	.812	-
6	0.470	0.509	0.061	.049	.151	-
7	0.465	0.497	0.074	5.88e-03	.028	T
8	0.492	0.502	0.023	.935	.997	-
9	0.487	0.507	0.048	.188	.420	-
10	0.498	0.494	0.038	.440	.691	-
11	0.488	0.493	0.024	.919	.997	-
12	0.481	0.499	0.038	.453	.691	-
13	0.486	0.493	0.024	.919	.997	-
14	0.486	0.488	0.018	.997	.997	-

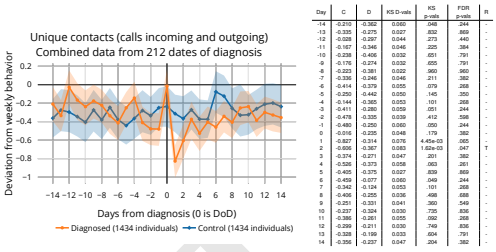
(c) Normalized fractional ranking.

Fig. S11. Feature graph and significance test results for ***unique contact count (outgoing)*** using raw, detrended, and ranked data.



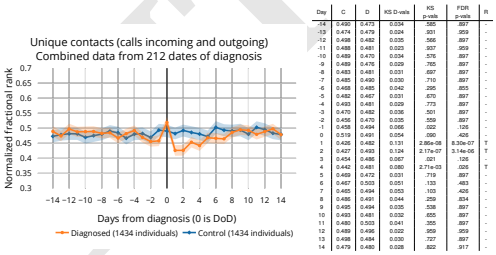
Day	C	D	KS D-vals	KS p-vals	FDR p-vals	R
-14	4.053	4.251	0.038	.429	.613	-
-13	3.904	4.235	0.053	.104	.188	-
-12	4.041	4.158	0.029	.765	.822	-
-11	3.976	4.148	0.035	.527	.644	-
-10	3.869	4.057	0.040	.368	.562	-
-9	3.896	4.069	0.022	.956	.987	-
-8	3.911	4.062	0.030	.711	.793	-
-7	3.978	4.365	0.052	.111	.190	-
-6	3.886	4.150	0.054	.094	.162	-
-5	3.829	4.051	0.031	.668	.775	-
-4	3.902	4.236	0.037	.455	.613	-
-3	3.753	4.224	0.072	8.96e-03	.032	T
-2	3.570	4.080	0.060	.045	.099	T
-1	3.718	4.218	0.077	3.71e-03	.018	T
0	4.255	4.330	0.019	.987	.987	-
1	3.511	4.219	0.108	9.50e-06	2.75e-04	T
2	3.504	4.196	0.101	4.36e-05	6.32e-04	T
3	3.680	4.289	0.095	1.91e-04	1.85e-03	T
4	3.607	4.127	0.092	3.20e-04	2.32e-03	T
5	3.643	4.116	0.068	.017	.050	-
6	3.782	4.283	0.077	4.24e-03	.018	T
7	3.850	4.417	0.084	1.16e-03	6.71e-03	T
8	3.855	4.242	0.045	.240	.397	-
9	3.944	4.267	0.056	.080	.162	-
10	3.895	4.139	0.037	.462	.613	-
11	3.769	4.151	0.065	.027	.066	-
12	3.722	4.233	0.068	.016	.050	-
13	3.846	4.079	0.035	.533	.644	-
14	3.886	4.249	0.064	.028	.066	-

(a) Raw feature value.



Day	C	D	KS D-vals	KS p-vals	FDR p-vals	R
-14	-0.210	-0.362	0.060	.048	.244	-
-13	-0.335	-0.275	0.027	.832	.869	-
-12	-0.028	-0.297	0.044	.273	.440	-
-11	-0.167	-0.346	0.046	.225	.394	-
-10	-0.238	-0.406	0.032	.651	.791	-
-9	-0.176	-0.274	0.032	.655	.791	-
-8	-0.233	-0.381	0.022	.960	.960	-
-7	-0.336	-0.246	0.046	.211	.382	-
-6	-0.414	-0.379	0.055	.079	.268	-
-5	-0.255	-0.442	0.050	.145	.330	-
-4	-0.144	-0.365	0.053	.101	.268	-
-3	-0.411	-0.285	0.059	.051	.244	-
-2	-0.478	-0.335	0.039	.412	.598	-
-1	-0.480	-0.255	0.060	.050	.244	-
0	-0.016	-0.235	0.048	.179	.382	-
1	-0.827	-0.314	0.076	4.45e-03	.065	T
2	-0.606	-0.367	0.083	1.62e-03	.047	T
3	-0.374	-0.271	0.047	.201	.382	-
4	-0.526	-0.373	0.058	.063	.261	-
5	-0.405	-0.375	0.027	.839	.869	-
6	-0.459	-0.077	0.060	.049	.244	-
7	-0.342	-0.124	0.053	.101	.268	-
8	-0.406	-0.255	0.036	.498	.698	-
9	-0.251	-0.331	0.041	.360	.549	-
10	-0.237	-0.324	0.030	.735	.836	-
11	-0.366	-0.261	0.055	.092	.268	-
12	-0.299	-0.211	0.030	.749	.836	-
13	-0.328	-0.199	0.033	.604	.791	-
14	-0.359	-0.237	0.047	.204	.382	-

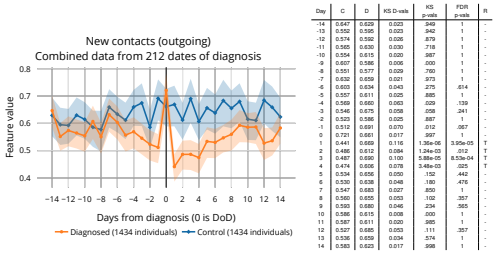
(b) Deviation from weekly average.



Day	C	D	KS D-vals	KS p-vals	FDR p-vals	R
-14	0.490	0.473	0.034	.585	.897	-
-13	0.474	0.479	0.024	.931	.959	-
-12	0.498	0.482	0.035	.566	.897	-
-11	0.488	0.481	0.023	.957	.959	-
-10	0.489	0.470	0.034	.576	.897	-
-9	0.489	0.476	0.029	.765	.897	-
-8	0.483	0.481	0.031	.697	.897	-
-7	0.485	0.490	0.030	.710	.897	-
-6	0.488	0.485	0.042	.295	.853	-
-5	0.482	0.467	0.031	.670	.897	-
-4	0.483	0.481	0.029	.773	.897	-
-3	0.470	0.482	0.036	.501	.897	-
-2	0.456	0.470	0.035	.559	.897	-
-1	0.459	0.494	0.066	.022	.126	-
0	0.519	0.491	0.054	.090	.426	-
1	0.426	0.482	0.131	2.86e-08	8.30e-07	T
2	0.427	0.493	0.124	2.17e-07	3.14e-06	T
3	0.454	0.486	0.067	.021	.126	-
4	0.442	0.481	0.080	2.71e-03	.026	T
5	0.469	0.472	0.031	.719	.897	-
6	0.467	0.503	0.051	.133	.483	-
7	0.485	0.494	0.053	.103	.426	-
8	0.486	0.491	0.044	.259	.834	-
9	0.495	0.494	0.035	.538	.897	-
10	0.493	0.481	0.038	.455	.897	-
11	0.480	0.503	0.041	.355	.897	-
12	0.489	0.496	0.022	.959	.959	-
13	0.488	0.484	0.032	.727	.897	-
14	0.479	0.480	0.029	.822	.917	-

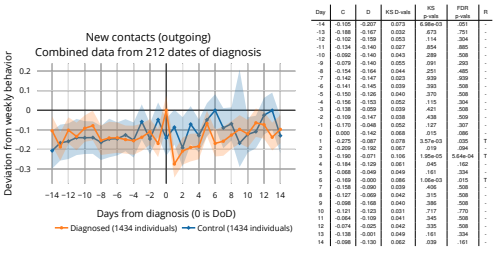
(c) Normalized fractional ranking.

Fig. S12. Feature graph and significance test results for ***unique contact count (both incoming and outgoing)*** using raw, detrended, and ranked data.



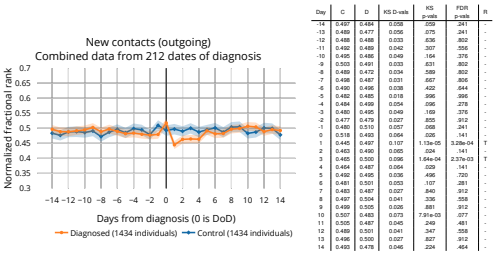
Day	C	D	KS D-vals	KS p-vals	FDR p-vals	R
-14	0.647	0.629	0.023	.949	1	-
-13	0.552	0.595	0.023	.942	1	-
-12	0.574	0.592	0.026	.879	1	-
-11	0.565	0.630	0.030	.718	1	-
-10	0.554	0.615	0.030	.687	1	-
-9	0.607	0.586	0.006	.000	1	-
-8	0.551	0.577	0.029	.760	1	-
-7	0.632	0.659	0.021	.973	1	-
-6	0.603	0.634	0.043	.275	.614	-
-5	0.557	0.611	0.025	.885	1	-
-4	0.569	0.660	0.063	.029	.139	-
-3	0.546	0.575	0.058	.058	.241	-
-2	0.523	0.586	0.025	.887	1	-
-1	0.512	0.591	0.070	.012	.067	-
0	0.721	0.661	0.017	.997	1	-
1	0.441	0.669	0.116	1.36e-06	3.95e-05	T
2	0.486	0.612	0.084	1.24e-03	.012	T
3	0.487	0.657	0.100	5.85e-05	8.53e-04	T
4	0.474	0.606	0.078	3.48e-03	.025	T
5	0.534	0.556	0.050	.152	.442	-
6	0.530	0.638	0.048	.180	.476	-
7	0.547	0.683	0.027	.850	1	-
8	0.569	0.655	0.053	.102	.357	-
9	0.593	0.680	0.046	.234	.565	-
10	0.586	0.615	0.008	.000	1	-
11	0.587	0.611	0.020	.985	1	-
12	0.527	0.685	0.053	.111	.357	-
13	0.536	0.659	0.034	.574	1	-
14	0.583	0.623	0.017	.998	1	-

(a) Raw feature value.



Day	C	D	KS D-vals	KS p-vals	FDR p-vals	R
-14	-0.105	-0.207	0.073	6.98e-03	.051	-
-13	-0.188	-0.167	0.032	.673	.751	-
-12	-0.102	-0.159	0.053	.114	.304	-
-11	-0.134	-0.140	0.027	.854	.999	-
-10	-0.092	-0.140	0.043	.289	.508	-
-9	-0.079	-0.140	0.055	.091	.293	-
-8	-0.154	-0.164	0.044	.251	.485	-
-7	-0.143	-0.147	0.023	.939	.939	-
-6	-0.141	-0.145	0.039	.393	.508	-
-5	-0.155	-0.135	0.040	.370	.508	-
-4	-0.156	-0.153	0.052	.115	.304	-
-3	-0.138	-0.059	0.039	.421	.508	-
-2	-0.109	-0.147	0.038	.438	.509	-
-1	-0.170	-0.048	0.052	.127	.307	-
0	0.000	-0.142	0.068	.015	.086	-
1	-0.275	-0.087	0.078	3.57e-03	.035	T
2	-0.209	-0.192	0.067	.019	.094	-
3	-0.190	-0.071	0.106	1.95e-05	5.64e-04	T
4	-0.184	-0.129	0.061	.045	.162	-
5	-0.068	-0.049	0.049	.161	.334	-
6	-0.169	-0.000	0.086	1.06e-03	.015	T
7	-0.158	-0.090	0.039	.406	.508	-
8	-0.127	-0.069	0.042	.315	.508	-
9	-0.098	-0.168	0.040	.386	.508	-
10	-0.121	-0.123	0.031	.717	.770	-
11	-0.064	-0.109	0.041	.345	.508	-
12	-0.074	-0.025	0.042	.335	.508	-
13	-0.138	-0.001	0.049	.161	.334	-
14	-0.098	-0.130	0.062	.039	.161	-

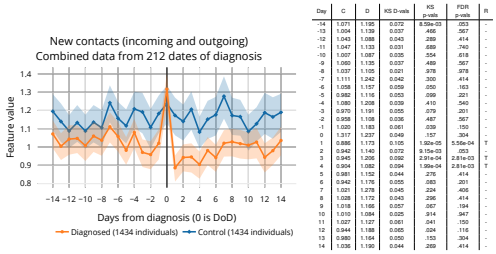
(b) Deviation from weekly average.



Day	C	D	KS D-vals	KS p-vals	FDR p-vals	R
-14	0.497	0.484	0.058	.059	.241	-
-13	0.489	0.477	0.056	.075	.241	-
-12	0.488	0.488	0.033	.636	.802	-
-11	0.492	0.489	0.042	.307	.556	-
-10	0.495	0.486	0.049	.164	.376	-
-9	0.503	0.491	0.033	.631	.802	-
-8	0.489	0.472	0.034	.589	.802	-
-7	0.498	0.487	0.031	.667	.806	-
-6	0.490	0.496	0.038	.432	.644	-
-5	0.482	0.485	0.018	.996	.996	-
-4	0.494	0.499	0.054	.096	.278	-
-3	0.480	0.495	0.049	.169	.376	-
-2	0.477	0.479	0.027	.855	.912	-
-1	0.480	0.510	0.057	.068	.241	-
0	0.518	0.493	0.064	.026	.141	-
1	0.445	0.497	0.107	1.13e-05	3.28e-04	T
2	0.463	0.490	0.065	.024	.141	-
3	0.485	0.500	0.096	1.64e-04	2.37e-03	T
4	0.464	0.487	0.064	.029	.141	-
5	0.492	0.495	0.036	.496	.720	-
6	0.481	0.501	0.053	.107	.281	-
7	0.483	0.487	0.027	.840	.912	-
8	0.504	0.504	0.041	.336	.558	-
9	0.499	0.505	0.026	.881	.912	-
10	0.507	0.483	0.073	7.91e-03	.077	-
11	0.505	0.487	0.045	.249	.481	-
12	0.489	0.501	0.041	.347	.558	-
13	0.496	0.500	0.027	.827	.912	-
14	0.493	0.478	0.046	.224	.464	-

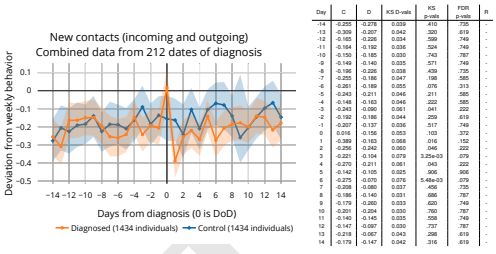
(c) Normalized fractional ranking.

Fig. S13. Feature graph and significance test results for ***new contacts (outgoing)*** using raw, detrended, and ranked data.



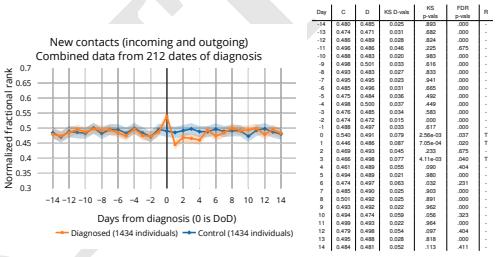
Day	C	D	KS D-vals	KS p-vals	FDR p-vals	R
-14	1.071	1.195	0.072	8.59e-03	.053	-
-13	1.004	1.139	0.037	.466	.567	-
-12	1.043	1.088	0.043	.289	.414	-
-11	1.047	1.133	0.031	.669	.740	-
-10	1.007	1.087	0.035	.554	.618	-
-9	1.060	1.135	0.037	.489	.567	-
-8	1.037	1.105	0.031	.678	.678	-
-7	1.111	1.242	0.042	.300	.414	-
-6	1.058	1.157	0.058	.050	.163	-
-5	0.982	1.116	0.053	.099	.221	-
-4	1.080	1.208	0.039	.410	.540	-
-3	0.970	1.191	0.055	.079	.201	-
-2	0.958	1.108	0.036	.487	.567	-
-1	1.030	1.183	0.061	.039	.150	-
0	1.317	1.237	0.049	.157	.304	-
1	0.886	1.173	0.105	1.92e-05	5.59e-04	T
2	0.942	1.140	0.072	9.15e-03	.053	-
3	0.945	1.206	0.092	2.91e-04	2.81e-03	T
4	0.904	1.082	0.094	1.99e-04	2.81e-03	T
5	0.981	1.152	0.044	.276	.414	-
6	0.942	1.176	0.055	.083	.261	-
7	1.021	1.278	0.045	.224	.406	-
8	1.028	1.172	0.043	.296	.414	-
9	1.018	1.166	0.057	.067	.194	-
10	1.010	1.094	0.025	.914	.947	-
11	1.027	1.127	0.061	.041	.150	-
12	0.944	1.188	0.065	.024	.116	-
13	0.980	1.164	0.050	.153	.304	-
14	1.036	1.190	0.044	.269	.414	-

(a) Raw feature value.



Day	C	D	KS D-vals	KS p-vals	FDR p-vals	R
-14	-0.255	-0.278	0.039	.410	.735	-
-13	-0.309	-0.207	0.042	.320	.619	-
-12	-0.165	-0.228	0.034	.529	.749	-
-11	-0.164	-0.192	0.036	.524	.749	-
-10	-0.150	-0.185	0.030	.743	.787	-
-9	-0.149	-0.140	0.035	.571	.749	-
-8	-0.196	-0.226	0.038	.439	.735	-
-7	-0.255	-0.186	0.047	.198	.585	-
-6	-0.261	-0.189	0.055	.076	.313	-
-5	-0.243	-0.211	0.046	.211	.585	-
-4	-0.148	-0.163	0.046	.222	.585	-
-3	-0.243	-0.092	0.061	.041	.222	-
-2	-0.192	-0.186	0.044	.259	.619	-
-1	-0.207	-0.137	0.036	.517	.749	-
0	0.014	-0.156	0.053	.103	.372	-
1	-0.389	-0.163	0.068	.016	.152	-
2	-0.256	-0.242	0.060	.046	.222	-
3	-0.221	-0.104	0.079	3.25e-03	.079	-
4	-0.270	-0.211	0.061	.043	.222	-
5	-0.142	-0.105	0.025	.906	.906	-
6	-0.275	-0.070	0.076	5.48e-03	.079	-
7	-0.208	-0.080	0.037	.456	.735	-
8	-0.186	-0.140	0.031	.686	.787	-
9	-0.179	-0.260	0.033	.620	.749	-
10	-0.201	-0.204	0.030	.760	.787	-
11	-0.140	-0.145	0.035	.558	.749	-
12	-0.147	-0.097	0.030	.737	.787	-
13	-0.218	-0.067	0.043	.298	.619	-
14	-0.179	-0.147	0.042	.316	.619	-

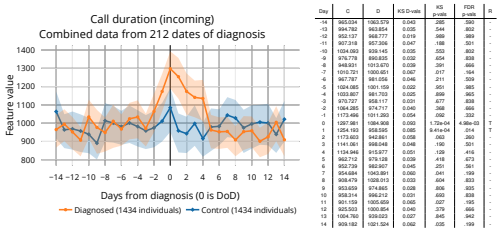
(b) Deviation from weekly average.



Day	C	D	KS D-vals	KS p-vals	FDR p-vals	R
-14	0.480	0.485	0.025	.893	.000	-
-13	0.474	0.471	0.031	.682	.000	-
-12	0.486	0.489	0.028	.824	.000	-
-11	0.492	0.486	0.046	.225	.675	-
-10	0.488	0.483	0.020	.983	.000	-
-9	0.498	0.501	0.033	.616	.000	-
-8	0.493	0.483	0.027	.833	.000	-
-7	0.495	0.495	0.023	.941	.000	-
-6	0.485	0.496	0.031	.685	.000	-
-5	0.475	0.484	0.036	.492	.000	-
-4	0.498	0.500	0.037	.449	.000	-
-3	0.476	0.485	0.034	.583	.000	-
-2	0.474	0.472	0.015	.000	.000	-
-1	0.488	0.497	0.033	.617	.000	-
0	0.540	0.491	0.079	2.56e-03	.037	T
1	0.446	0.486	0.087	7.05e-04	.020	T
2	0.469	0.493	0.045	.233	.675	-
3	0.498	0.498	0.077	4.11e-03	.040	T
4	0.461	0.489	0.055	.090	.404	-
5	0.494	0.489	0.021	.980	.000	-
6	0.474	0.497	0.063	.032	.231	-
7	0.485	0.490	0.025	.903	.000	-
8	0.501	0.492	0.025	.891	.000	-
9	0.493	0.492	0.020	.962	.000	-
10	0.494	0.474	0.059	.056	.323	-
11	0.499	0.493	0.022	.964	.000	-
12	0.479	0.498	0.054	.097	.404	-
13	0.492	0.488	0.028	.818	.000	-
14	0.484	0.481	0.052	.113	.411	-

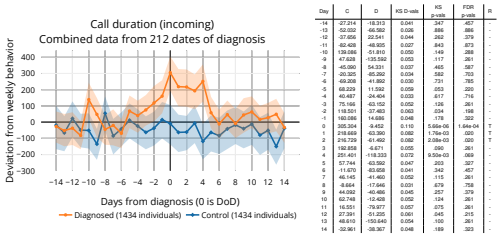
(c) Normalized fractional ranking.

Fig. S14. Feature graph and significance test results for ***new contacts (both incoming and outgoing)*** using raw, detrended, and ranked data.



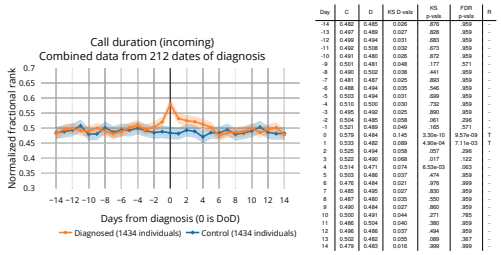
Day	C	D	KS D-vals	KS p-vals	FDR p-vals	R
-14	965.034	1063.579	0.043	.295	.590	-
-13	994.760	963.854	0.035	.544	.802	-
-12	952.137	968.777	0.019	.989	.989	-
-11	967.318	957.396	0.047	.188	.501	-
-10	1034.093	939.145	0.035	.553	.802	-
-9	976.778	899.835	0.032	.654	.838	-
-8	948.931	1013.670	0.039	.391	.666	-
-7	1010.721	1000.451	0.067	.017	.164	-
-6	967.797	981.656	0.046	.211	.509	-
-5	1024.085	1001.159	0.022	.951	.985	-
-4	1033.807	981.702	0.025	.899	.965	-
-3	970.727	958.117	0.031	.677	.838	-
-2	1064.265	974.717	0.040	.368	.666	-
-1	1173.496	1011.293	0.054	.092	.332	-
0	1297.381	1084.908	0.093	1.73e-04	4.66e-03	T
1	1254.193	958.595	0.085	9.41e-04	.014	T
2	1173.823	942.881	0.058	.060	.260	-
3	1141.061	998.048	0.048	.190	.501	-
4	1134.946	915.677	0.051	.129	.416	-
5	962.712	979.128	0.039	.418	.673	-
6	952.739	982.907	0.045	.251	.561	-
7	954.684	1043.891	0.060	.041	.199	-
8	908.479	1009.213	0.033	.604	.833	-
9	953.659	974.865	0.038	.806	.935	-
10	958.314	996.212	0.031	.693	.838	-
11	901.159	1005.459	0.065	.027	.195	-
12	925.503	1000.854	0.040	.379	.666	-
13	1004.769	939.023	0.027	.845	.942	-
14	909.182	1021.324	0.062	.035	.199	-

(a) Raw feature value.



Day	C	D	KS D-vals	KS p-vals	FDR p-vals	R
-14	-27.214	-18.313	0.041	.347	.457	-
-13	-52.052	-66.562	0.026	.886	.886	-
-12	-57.656	-22.541	0.044	.242	.538	-
-11	-62.428	-48.935	0.027	.843	.873	-
-10	-139.846	-51.810	0.050	.149	.388	-
-9	-47.628	-135.592	0.053	.117	.261	-
-8	-45.060	54.331	0.037	.465	.587	-
-7	-20.325	-85.292	0.034	.592	.703	-
-6	-69.209	-41.852	0.030	.731	.785	-
-5	68.229	11.592	0.059	.053	.220	-
-4	-60.487	-24.404	0.033	.617	.716	-
-3	75.166	-63.152	0.052	.126	.261	-
-2	118.531	-37.463	0.063	.034	.198	-
-1	160.086	14.686	0.048	.178	.322	-
0	305.324	-9.452	0.110	5.66e-06	1.66e-04	T
1	218.669	-63.390	0.092	1.76e-03	.020	T
2	216.729	-61.492	0.092	2.09e-03	.020	T
3	192.858	-6.671	0.055	.090	.261	-
4	251.401	-118.333	0.072	9.50e-03	.069	-
5	57.744	-63.592	0.047	.203	.327	-
6	-11.670	-83.858	0.041	.342	.457	-
7	46.145	-41.469	0.052	.115	.261	-
8	-8.664	-17.646	0.031	.679	.758	-
9	44.092	-40.469	0.045	.257	.379	-
10	62.748	-12.428	0.052	.124	.261	-
11	16.551	-79.877	0.057	.075	.261	-
12	27.391	-51.235	0.061	.045	.215	-
13	48.610	-100.645	0.054	.100	.261	-
14	-32.961	-38.367	0.048	.189	.323	-

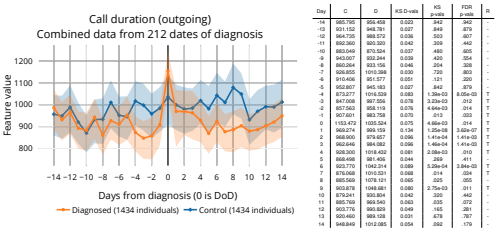
(b) Deviation from weekly average.



Day	C	D	KS D-vals	KS p-vals	FDR p-vals	R
-14	0.482	0.485	0.026	.876	.959	-
-13	0.497	0.489	0.027	.828	.959	-
-12	0.499	0.494	0.031	.693	.959	-
-11	0.492	0.508	0.032	.673	.959	-
-10	0.491	0.480	0.026	.872	.959	-
-9	0.501	0.481	0.048	.177	.571	-
-8	0.490	0.502	0.038	.441	.959	-
-7	0.493	0.487	0.025	.893	.959	-
-6	0.488	0.494	0.035	.546	.959	-
-5	0.503	0.494	0.031	.699	.959	-
-4	0.510	0.500	0.030	.732	.959	-
-3	0.495	0.492	0.025	.890	.959	-
-2	0.504	0.485	0.058	.061	.296	-
-1	0.521	0.489	0.049	.165	.571	-
0	0.578	0.484	0.145	3.30e-10	9.57e-09	T
1	0.533	0.482	0.089	4.90e-04	7.11e-03	T
2	0.525	0.494	0.058	.057	.296	-
3	0.522	0.490	0.058	.017	.122	-
4	0.514	0.471	0.074	6.53e-03	.063	-
5	0.503	0.486	0.037	.478	.959	-
6	0.476	0.484	0.021	.976	.999	-
7	0.485	0.495	0.027	.835	.959	-
8	0.487	0.480	0.035	.550	.959	-
9	0.490	0.484	0.027	.860	.959	-
10	0.500	0.491	0.044	.271	.785	-
11	0.486	0.504	0.040	.380	.959	-
12	0.496	0.486	0.037	.474	.959	-
13	0.502	0.482	0.055	.089	.367	-
14	0.479	0.483	0.016	.999	.999	-

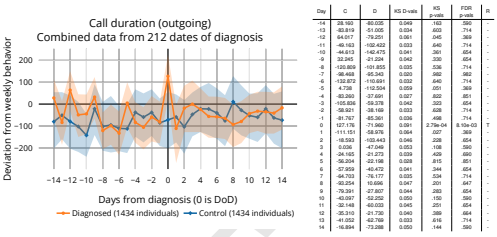
(c) Normalized fractional ranking.

Fig. S15. Feature graph and significance test results for ***call duration (incoming)*** using raw, detrended, and ranked data.



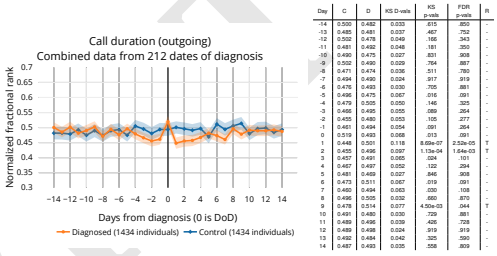
Day	C	D	KS D-vals	KS p-vals	FDR p-vals	R
-14	985.795	956.458	0.033	.942	.942	-
-13	931.152	946.781	0.027	.849	.879	-
-12	964.795	988.572	0.036	.553	.607	-
-11	992.360	930.393	0.042	.309	.442	-
-10	883.049	870.524	0.037	.480	.605	-
-9	943.007	932.944	0.039	.429	.554	-
-8	960.264	933.156	0.046	.254	.328	-
-7	926.855	1010.398	0.039	.720	.802	-
-6	910.406	951.577	0.051	.191	.290	-
-5	952.607	945.183	0.027	.842	.879	-
-4	873.277	1016.539	0.083	1.39e-03	6.05e-03	T
-3	947.008	997.556	0.078	3.33e-03	.012	T
-2	857.563	958.119	0.076	4.94e-03	.014	T
-1	907.601	983.758	0.070	.013	.033	T
0	1153.472	1035.324	0.075	4.66e-03	.014	T
1	969.274	999.159	0.134	1.25e-08	3.62e-07	T
2	968.900	979.657	0.096	1.41e-04	1.41e-03	T
3	962.646	984.092	0.096	1.46e-04	1.41e-03	T
4	928.300	1018.432	0.081	2.38e-03	.010	T
5	868.498	981.406	0.044	.269	.411	-
6	923.770	1042.314	0.089	5.29e-04	3.84e-03	T
7	876.068	1010.531	0.068	.014	.034	T
8	885.569	1078.121	0.065	.025	.055	-
9	903.878	1048.681	0.080	2.75e-03	.011	T
10	879.241	993.804	0.042	.330	.442	-
11	885.769	969.540	0.043	.035	.072	-
12	903.775	990.929	0.049	.165	.381	-
13	920.460	989.128	0.031	.678	.787	-
14	948.849	1012.085	0.054	.092	.179	-

(a) Raw feature value.



Day	C	D	KS D-vals	KS p-vals	FDR p-vals	R
-14	-38.160	-60.035	0.049	.163	.590	-
-13	-93.819	-51.005	0.054	.603	.714	-
-12	-54.037	-79.251	0.061	.045	.369	-
-11	-99.163	-102.422	0.033	.640	.714	-
-10	-44.613	-140.475	0.041	.361	.654	-
-9	32.245	-21.224	0.042	.330	.654	-
-8	-120.809	-101.855	0.035	.536	.714	-
-7	-98.468	-95.343	0.020	.982	.982	-
-6	-132.872	-110.691	0.032	.640	.714	-
-5	4.738	-112.504	0.059	.051	.369	-
-4	-83.260	-37.491	0.027	.822	.851	-
-3	-155.836	-59.378	0.042	.323	.654	-
-2	-56.921	-38.169	0.033	.628	.714	-
-1	-91.797	-85.361	0.036	.498	.714	-
0	127.176	-71.960	0.091	2.79e-04	8.10e-03	T
1	-111.151	-58.976	0.064	.027	.369	-
2	-18.593	-102.443	0.046	.228	.654	-
3	0.036	-47.049	0.053	.108	.590	-
4	-24.165	-21.273	0.039	.429	.690	-
5	-56.204	-22.198	0.028	.815	.851	-
6	-57.959	-49.472	0.041	.344	.654	-
7	-94.703	-76.177	0.035	.534	.714	-
8	-93.254	19.696	0.047	.201	.647	-
9	-79.391	-27.807	0.044	.283	.654	-
10	-93.097	-52.252	0.050	.150	.590	-
11	-32.148	-60.030	0.045	.251	.654	-
12	-35.310	-21.730	0.040	.389	.684	-
13	-41.652	-62.769	0.033	.616	.714	-
14	-16.894	-73.288	0.050	.144	.590	-

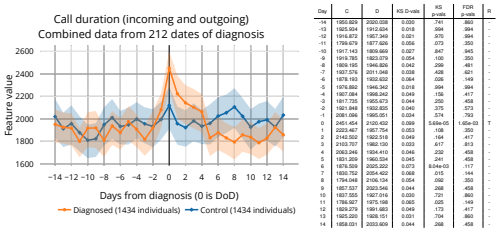
(b) Deviation from weekly average.



Day	C	D	KS D-vals	KS p-vals	FDR p-vals	R
-14	0.500	0.482	0.033	.615	.850	-
-13	0.485	0.481	0.037	.467	.752	-
-12	0.502	0.478	0.049	.166	.343	-
-11	0.481	0.492	0.048	.181	.350	-
-10	0.490	0.475	0.027	.831	.908	-
-9	0.502	0.492	0.035	.764	.887	-
-8	0.471	0.474	0.036	.511	.780	-
-7	0.494	0.492	0.024	.917	.919	-
-6	0.476	0.493	0.030	.705	.881	-
-5	0.496	0.475	0.067	.016	.091	-
-4	0.479	0.505	0.050	.146	.325	-
-3	0.468	0.495	0.055	.089	.264	-
-2	0.495	0.480	0.053	.105	.277	-
-1	0.461	0.494	0.054	.091	.264	-
0	0.516	0.493	0.068	.013	.091	-
1	0.448	0.501	0.118	8.69e-07	2.52e-05	T
2	0.455	0.496	0.097	1.13e-04	1.64e-03	T
3	0.457	0.491	0.065	.024	.101	-
4	0.467	0.497	0.052	.122	.294	-
5	0.493	0.489	0.027	.846	.908	-
6	0.473	0.511	0.067	.019	.091	-
7	0.490	0.494	0.063	.020	.108	-
8	0.496	0.505	0.032	.660	.870	-
9	0.478	0.514	0.077	4.50e-03	.044	T
10	0.491	0.480	0.030	.729	.881	-
11	0.489	0.496	0.029	.426	.728	-
12	0.499	0.498	0.024	.919	.919	-
13	0.492	0.494	0.042	.325	.590	-
14	0.497	0.493	0.035	.556	.809	-

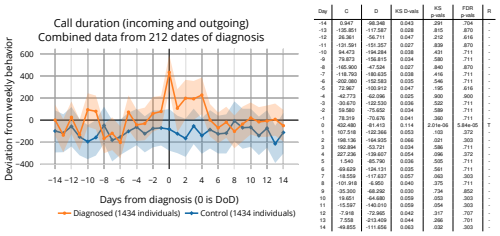
(c) Normalized fractional ranking.

Fig. S16. Feature graph and significance test results for ***call duration (outgoing)*** using raw, detrended, and ranked data.



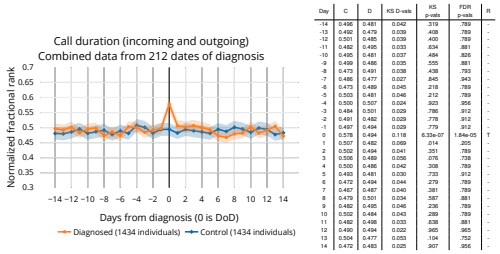
Day	C	D	KS D-vals	KS p-vals	FDR p-vals	R
-14	1950.829	2020.039	0.030	.741	.860	-
-13	1905.924	1912.624	0.018	.994	.994	-
-12	1916.872	1957.349	0.021	.970	.994	-
-11	1799.679	1877.626	0.056	.073	.350	-
-10	1917.143	1909.669	0.027	.847	.945	-
-9	1916.765	1923.075	0.054	.100	.350	-
-8	1859.155	1946.867	0.042	.299	.681	-
-7	1937.576	2011.048	0.038	.428	.821	-
-6	1878.102	1932.632	0.064	.036	.149	-
-5	1976.862	1946.342	0.018	.994	.994	-
-4	1967.084	1998.282	0.049	.156	.417	-
-3	1917.735	1955.673	0.044	.259	.458	-
-2	1981.848	1932.835	0.040	.375	.573	-
-1	2081.096	1995.051	0.034	.574	.793	-
0	2451.454	2120.435	0.099	5.64e-05	1.65e-03	T
1	2223.467	1957.754	0.053	.108	.350	-
2	2142.502	1922.516	0.049	.164	.417	-
3	2103.707	1982.100	0.033	.617	.813	-
4	2062.246	1934.410	0.046	.232	.458	-
5	1921.209	1960.534	0.045	.241	.458	-
6	1878.509	2025.220	0.073	8.04e-03	.117	-
7	1930.752	2054.402	0.068	.015	.144	-
8	1794.048	2106.134	0.054	.092	.350	-
9	1857.537	2023.546	0.044	.268	.458	-
10	1837.555	1927.016	0.030	.731	.860	-
11	1786.927	1975.198	0.065	.025	.149	-
12	1829.279	1991.683	0.049	.173	.417	-
13	1925.209	1928.151	0.031	.704	.860	-
14	1858.031	2003.609	0.044	.268	.458	-

(a) Raw feature value.



Day	C	D	KS D-vals	KS p-vals	FDR p-vals	R
-14	0.947	68.348	0.043	.291	.704	-
-13	-135.851	-117.567	0.039	.315	.870	-
-12	36.361	-56.711	0.047	.212	.616	-
-11	-131.591	-151.357	0.027	.839	.870	-
-10	94.473	-194.284	0.038	.431	.711	-
-9	79.873	-156.815	0.034	.580	.711	-
-8	-165.900	-47.534	0.027	.840	.870	-
-7	-118.760	-180.635	0.038	.416	.711	-
-6	-202.060	-152.583	0.035	.546	.711	-
-5	72.967	-100.912	0.047	.195	.616	-
-4	-42.773	-62.096	0.035	.500	.800	-
-3	-30.670	-122.530	0.036	.532	.711	-
-2	59.581	-75.652	0.034	.589	.711	-
-1	78.319	-70.676	0.041	.360	.711	-
0	402.480	81.413	0.114	2.01e-04	5.84e-03	T
1	167.518	-122.366	0.053	.103	.372	-
2	168.136	-164.935	0.066	.031	.303	-
3	192.894	-53.721	0.034	.586	.711	-
4	207.236	-139.627	0.054	.096	.372	-
5	1.540	-85.790	0.036	.505	.711	-
6	-69.626	-134.131	0.035	.561	.711	-
7	-18.559	-117.637	0.057	.063	.303	-
8	101.918	-6.993	0.040	.375	.711	-
9	35.300	-68.290	0.030	.734	.852	-
10	19.651	-64.880	0.059	.053	.303	-
11	-10.530	-140.519	0.059	.054	.303	-
12	-7.918	-72.965	0.042	.317	.707	-
13	7.558	-213.409	0.044	.266	.701	-
14	-49.855	-111.656	0.063	.032	.303	-

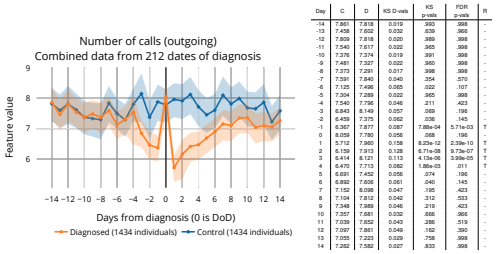
(b) Deviation from weekly average.



Day	C	D	KS D-vals	KS p-vals	FDR p-vals	R
-14	0.496	0.481	0.042	.319	.789	-
-13	0.492	0.479	0.039	.408	.789	-
-12	0.501	0.485	0.039	.400	.789	-
-11	0.482	0.495	0.033	.634	.881	-
-10	0.495	0.481	0.037	.484	.826	-
-9	0.499	0.486	0.035	.555	.881	-
-8	0.473	0.491	0.038	.438	.793	-
-7	0.486	0.477	0.027	.845	.943	-
-6	0.473	0.489	0.045	.218	.789	-
-5	0.503	0.481	0.046	.212	.789	-
-4	0.500	0.507	0.024	.923	.956	-
-3	0.484	0.501	0.029	.786	.912	-
-2	0.491	0.482	0.029	.779	.912	-
-1	0.497	0.494	0.029	.779	.912	-
0	0.578	0.494	0.118	8.33e-07	1.84e-05	T
1	0.507	0.482	0.069	.014	.205	-
2	0.502	0.494	0.041	.351	.789	-
3	0.506	0.489	0.056	.076	.738	-
4	0.500	0.486	0.042	.308	.789	-
5	0.493	0.481	0.030	.733	.912	-
6	0.472	0.494	0.044	.279	.789	-
7	0.487	0.487	0.040	.381	.789	-
8	0.479	0.501	0.034	.587	.881	-
9	0.482	0.495	0.046	.236	.789	-
10	0.502	0.484	0.043	.289	.789	-
11	0.482	0.498	0.033	.638	.881	-
12	0.496	0.481	0.042	.319	.789	-
13	0.504	0.477	0.053	.104	.752	-
14	0.472	0.483	0.025	.907	.956	-

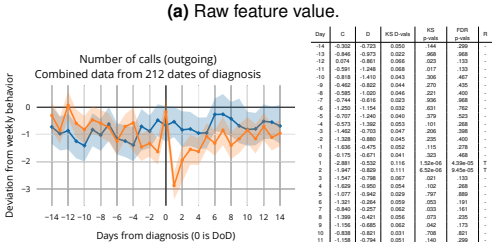
(c) Normalized fractional ranking.

Fig. S17. Feature graph and significance test results for ***call duration (both incoming and outgoing)*** using raw, detrended, and ranked data.



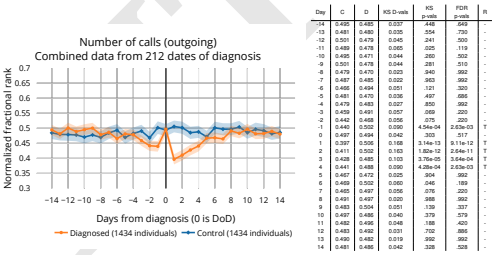
Day	C	D	KS D-vals	KS p-vals	FDR p-vals	R
-14	7.861	7.818	0.019	.993	.998	-
-13	7.458	7.602	0.032	.639	.966	-
-12	7.809	7.818	0.020	.989	.998	-
-11	7.540	7.617	0.022	.965	.998	-
-10	7.376	7.374	0.019	.991	.998	-
-9	7.481	7.327	0.022	.960	.998	-
-8	7.379	7.291	0.017	.998	.998	-
-7	7.591	7.845	0.040	.354	.570	-
-6	7.125	7.496	0.065	.022	.107	-
-5	7.304	7.289	0.022	.965	.998	-
-4	7.540	7.796	0.046	.211	.423	-
-3	6.843	8.149	0.057	.069	.196	-
-2	6.459	7.375	0.062	.036	.145	-
-1	6.367	7.877	0.087	7.88e-04	5.71e-03	T
0	8.059	7.780	0.056	.068	.196	-
1	5.712	7.980	0.158	8.23e-12	2.39e-10	T
2	6.159	7.913	0.128	6.71e-08	9.73e-07	T
3	6.414	8.121	0.113	4.13e-06	3.99e-05	T
4	6.470	7.713	0.082	1.86e-03	.011	T
5	6.691	7.452	0.056	.074	.196	-
6	6.892	7.606	0.061	.040	.145	-
7	7.152	8.098	0.047	.195	.423	-
8	7.104	7.812	0.042	.312	.533	-
9	7.348	7.989	0.046	.219	.423	-
10	7.357	7.681	0.032	.666	.966	-
11	7.039	7.652	0.043	.286	.519	-
12	7.097	7.861	0.049	.162	.390	-
13	7.055	7.223	0.029	.758	.998	-
14	7.262	7.582	0.027	.833	.998	-

(a) Raw feature value.



Day	C	D	KS D-vals	KS p-vals	FDR p-vals	R
-14	-0.332	-0.723	0.059	.144	.299	-
-13	-0.946	-0.973	0.022	.968	.968	-
-12	0.074	-0.861	0.066	.023	.133	-
-11	-0.591	-1.246	0.068	.017	.133	-
-10	-0.818	-1.410	0.043	.306	.467	-
-9	-0.462	-0.822	0.044	.270	.435	-
-8	-0.585	-1.020	0.046	.221	.400	-
-7	-0.744	-0.616	0.023	.936	.968	-
-6	-1.250	-1.154	0.032	.631	.762	-
-5	-0.707	-1.240	0.040	.379	.523	-
-4	-0.573	-1.392	0.053	.101	.268	-
-3	-1.462	-0.703	0.047	.206	.398	-
-2	-1.328	-0.880	0.045	.235	.400	-
-1	-1.636	-0.475	0.052	.115	.278	-
0	-0.175	-0.671	0.041	.323	.468	-
1	-2.881	-0.532	0.116	1.52e-06	4.39e-05	T
2	-1.947	-0.829	0.111	6.52e-06	9.45e-05	T
3	-1.547	-0.798	0.067	.021	.133	-
4	-1.629	-0.950	0.054	.102	.268	-
5	-1.077	-0.942	0.029	.797	.889	-
6	-1.321	-0.264	0.059	.053	.191	-
7	-0.840	-0.257	0.062	.033	.161	-
8	-1.399	-0.421	0.056	.073	.235	-
9	-1.156	-0.685	0.062	.042	.173	-
10	-0.839	-0.821	0.031	.708	.821	-
11	-1.158	-0.794	0.051	.140	.299	-
12	-0.704	-0.515	0.023	.941	.968	-
13	-1.109	-0.547	0.037	.486	.641	-
14	-0.955	-0.687	0.035	.542	.683	-

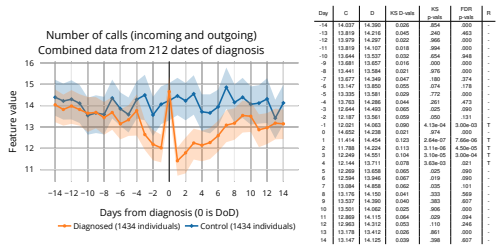
(b) Deviation from weekly average.



Day	C	D	KS D-vals	KS p-vals	FDR p-vals	R
-14	0.495	0.485	0.037	.448	.649	-
-13	0.481	0.480	0.035	.554	.730	-
-12	0.501	0.479	0.045	.241	.500	-
-11	0.489	0.478	0.065	.025	.119	-
-10	0.495	0.471	0.044	.260	.502	-
-9	0.501	0.478	0.044	.281	.510	-
-8	0.479	0.470	0.023	.940	.992	-
-7	0.487	0.485	0.022	.963	.992	-
-6	0.486	0.494	0.051	.121	.320	-
-5	0.481	0.470	0.036	.497	.686	-
-4	0.479	0.483	0.027	.850	.992	-
-3	0.459	0.491	0.057	.069	.220	-
-2	0.442	0.468	0.056	.075	.220	-
-1	0.440	0.502	0.090	4.54e-04	2.63e-03	T
0	0.497	0.494	0.042	.303	.517	-
1	0.397	0.506	0.168	3.14e-13	9.11e-12	T
2	0.411	0.502	0.163	1.82e-12	2.64e-11	T
3	0.438	0.485	0.103	3.76e-05	3.64e-04	T
4	0.441	0.488	0.090	4.28e-04	2.63e-03	T
5	0.467	0.472	0.025	.904	.992	-
6	0.469	0.502	0.065	.046	.189	-
7	0.465	0.497	0.056	.076	.220	-
8	0.491	0.497	0.020	.988	.992	-
9	0.483	0.504	0.051	.139	.337	-
10	0.497	0.486	0.040	.379	.579	-
11	0.482	0.496	0.048	.188	.420	-
12	0.483	0.492	0.031	.702	.886	-
13	0.490	0.482	0.019	.992	.992	-
14	0.481	0.486	0.042	.328	.528	-

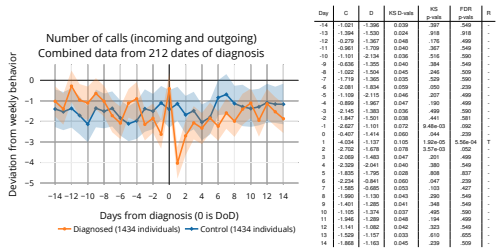
(c) Normalized fractional ranking.

Fig. S18. Feature graph and significance test results for ***calls count (outgoing)*** using raw, detrended, and ranked data.



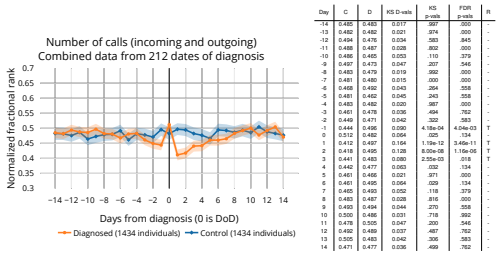
Day	C	D	KS D-vals	KS p-vals	FDR p-vals	R
-14	14.037	14.390	0.026	.854	.000	-
-13	13.819	14.218	0.040	.240	.463	-
-12	13.979	14.297	0.022	.966	.000	-
-11	13.919	14.107	0.018	.994	.000	-
-10	13.644	13.537	0.032	.654	.948	-
-9	13.681	13.657	0.016	.000	.000	-
-8	13.441	13.584	0.021	.976	.000	-
-7	13.677	14.349	0.047	.180	.374	-
-6	13.147	13.850	0.055	.074	.178	-
-5	13.335	13.581	0.029	.772	.000	-
-4	13.783	14.288	0.044	.261	.473	-
-3	12.644	14.492	0.065	.025	.090	-
-2	12.187	13.561	0.059	.050	.131	-
-1	12.021	14.063	0.090	4.13e-04	3.00e-03	T
0	14.652	14.238	0.021	.974	.000	-
1	11.414	14.454	0.123	2.64e-07	7.68e-06	T
2	11.788	14.224	0.113	3.11e-06	4.50e-05	T
3	12.240	14.551	0.104	3.19e-05	9.26e-04	T
4	12.144	13.711	0.078	3.63e-03	.021	T
5	12.269	13.858	0.065	.025	.090	-
6	12.594	13.946	0.067	.019	.090	-
7	13.084	14.858	0.062	.035	.101	-
8	13.176	14.150	0.041	.333	.569	-
9	13.537	14.390	0.040	.383	.607	-
10	13.501	14.062	0.025	.906	.000	-
11	12.869	14.115	0.064	.029	.094	-
12	12.962	14.012	0.053	.110	.284	-
13	13.178	13.412	0.026	.861	.000	-
14	13.147	14.125	0.039	.398	.607	-

(a) Raw feature value.



Day	C	D	KS D-vals	KS p-vals	FDR p-vals	R
-14	-1.021	-1.396	0.039	.397	.549	-
-13	-1.394	-1.530	0.024	.918	.978	-
-12	-0.279	-1.367	0.048	.176	.499	-
-11	-0.961	-1.709	0.040	.367	.549	-
-10	-1.101	-2.134	0.036	.516	.590	-
-9	-0.636	-1.355	0.040	.384	.549	-
-8	-1.022	-1.504	0.040	.246	.509	-
-7	-1.719	-1.365	0.035	.529	.590	-
-6	-2.081	-1.834	0.059	.050	.239	-
-5	-1.109	-2.115	0.046	.207	.499	-
-4	-0.899	-1.967	0.047	.180	.499	-
-3	-2.145	-1.383	0.036	.499	.590	-
-2	-1.847	-1.501	0.038	.441	.581	-
-1	-2.627	-1.101	0.072	9.46e-03	.092	-
0	-0.407	-1.414	0.060	.044	.239	-
1	-4.024	-1.137	0.105	1.92e-05	5.58e-04	T
2	-2.702	-1.678	0.078	3.57e-03	.052	-
3	-2.669	-1.483	0.047	.201	.499	-
4	-2.329	-2.041	0.040	.380	.549	-
5	-1.825	-1.795	0.028	.808	.837	-
6	-2.234	-1.841	0.060	.047	.239	-
7	-1.585	-1.685	0.053	.103	.427	-
8	-1.990	-1.139	0.043	.290	.549	-
9	-1.401	-1.285	0.041	.348	.549	-
10	-1.105	-1.374	0.037	.495	.590	-
11	-1.946	-1.289	0.048	.194	.499	-
12	-1.141	-1.082	0.042	.323	.549	-
13	-1.529	-1.157	0.033	.610	.655	-
14	-1.868	-1.163	0.045	.239	.509	-

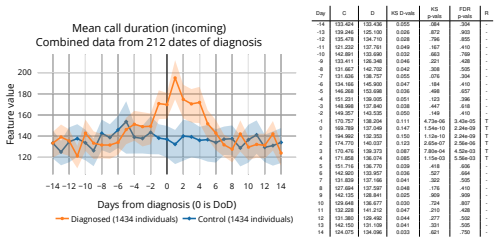
(b) Deviation from weekly average.



Day	C	D	KS D-vals	KS p-vals	FDR p-vals	R
-14	0.485	0.483	0.017	.997	.000	-
-13	0.482	0.482	0.021	.974	.000	-
-12	0.494	0.476	0.034	.583	.845	-
-11	0.488	0.487	0.028	.802	.000	-
-10	0.486	0.485	0.053	.110	.279	-
-9	0.497	0.473	0.047	.207	.546	-
-8	0.483	0.479	0.019	.992	.000	-
-7	0.481	0.480	0.015	.000	.000	-
-6	0.468	0.492	0.043	.264	.558	-
-5	0.481	0.482	0.045	.243	.558	-
-4	0.483	0.482	0.020	.987	.000	-
-3	0.463	0.478	0.036	.494	.762	-
-2	0.449	0.471	0.042	.322	.583	-
-1	0.444	0.496	0.090	4.18e-04	4.04e-03	T
0	0.512	0.482	0.064	.029	.134	-
1	0.412	0.497	0.164	1.19e-12	3.46e-11	T
2	0.418	0.495	0.128	8.00e-08	1.16e-06	T
3	0.441	0.483	0.080	2.55e-03	.018	T
4	0.442	0.477	0.063	.032	.134	-
5	0.461	0.466	0.021	.971	.000	-
6	0.461	0.495	0.064	.029	.134	-
7	0.465	0.493	0.052	.118	.379	-
8	0.483	0.487	0.028	.816	.000	-
9	0.493	0.494	0.044	.270	.558	-
10	0.500	0.486	0.031	.718	.992	-
11	0.478	0.505	0.047	.200	.546	-
12	0.492	0.489	0.037	.487	.762	-
13	0.505	0.483	0.042	.306	.583	-
14	0.471	0.477	0.036	.499	.762	-

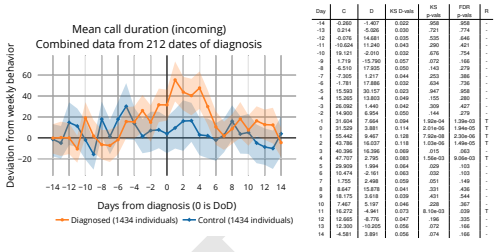
(c) Normalized fractional ranking.

Fig. S19. Feature graph and significance test results for ***calls count (both incoming and outgoing)*** using raw, detrended, and ranked data.



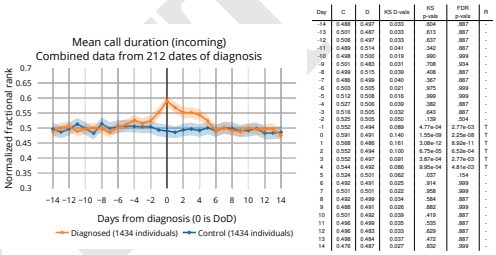
Day	C	D	KS D-vals	KS p-vals	FDR p-vals	R
-14	133.424	132.436	0.055	.084	.304	-
-13	139.246	135.100	0.026	.872	.993	-
-12	135.478	134.710	0.028	.796	.855	-
-11	131.232	137.761	0.049	.167	.410	-
-10	142.891	133.690	0.032	.663	.769	-
-9	133.411	136.248	0.046	.231	.438	-
-8	131.667	142.702	0.042	.308	.595	-
-7	131.636	138.757	0.055	.076	.304	-
-6	134.166	145.930	0.047	.184	.410	-
-5	146.268	153.698	0.036	.498	.657	-
-4	151.231	139.095	0.051	.123	.396	-
-3	148.998	137.840	0.038	.447	.618	-
-2	149.357	143.535	0.050	.149	.410	-
-1	170.757	138.224	0.111	4.73e-06	3.40e-05	T
0	169.769	137.049	0.147	1.04e-10	2.24e-09	T
1	194.992	132.333	0.150	1.12e-10	2.24e-09	T
2	174.779	146.037	0.123	2.65e-07	2.56e-06	T
3	170.476	139.373	0.087	7.80e-04	4.52e-03	T
4	171.858	136.074	0.085	1.15e-03	5.56e-03	T
5	151.716	136.770	0.039	.418	.696	-
6	142.980	133.957	0.036	.527	.664	-
7	131.839	137.166	0.041	.322	.595	-
8	127.694	137.597	0.048	.176	.410	-
9	142.135	128.841	0.025	.909	.909	-
10	129.648	136.677	0.030	.704	.807	-
11	132.228	141.212	0.047	.210	.438	-
12	131.380	129.492	0.044	.277	.532	-
13	142.152	131.159	0.041	.331	.595	-
14	124.075	134.096	0.033	.621	.730	-

(a) Raw feature value.



Day	C	D	KS D-vals	KS p-vals	FDR p-vals	R
-14	-0.260	-1.407	0.022	.958	.958	-
-13	0.214	-5.026	0.030	.721	.774	-
-12	-0.076	14.681	0.035	.535	.646	-
-11	-10.624	11.240	0.043	.290	.421	-
-10	19.121	-2.010	0.032	.676	.754	-
-9	1.719	-15.790	0.057	.072	.166	-
-8	-6.516	17.935	0.050	.143	.379	-
-7	-7.305	1.217	0.044	.253	.386	-
-6	-1.781	17.886	0.032	.634	.736	-
-5	15.593	30.157	0.023	.947	.958	-
-4	15.265	13.803	0.049	.155	.280	-
-3	26.092	1.440	0.042	.309	.427	-
-2	14.900	6.954	0.050	.144	.279	-
-1	31.604	7.664	0.094	1.60e-04	1.28e-03	T
0	31.539	3.881	0.114	2.01e-05	1.94e-05	T
1	55.442	9.467	0.128	7.60e-06	2.30e-06	T
2	43.786	16.037	0.116	1.03e-06	1.49e-05	T
3	40.396	16.398	0.069	.015	.063	-
4	47.707	2.795	0.083	1.56e-03	9.06e-03	T
5	29.909	1.994	0.064	.029	.103	-
6	16.474	-2.161	0.063	.032	.103	-
7	1.755	2.498	0.059	.051	.149	-
8	8.647	15.878	0.041	.331	.436	-
9	18.175	3.618	0.039	.431	.544	-
10	7.467	5.197	0.046	.228	.367	-
11	16.272	-4.941	0.073	8.19e-03	.039	T
12	12.665	-8.776	0.047	.196	.335	-
13	12.390	-10.205	0.056	.072	.166	-
14	-4.581	3.891	0.056	.074	.166	-

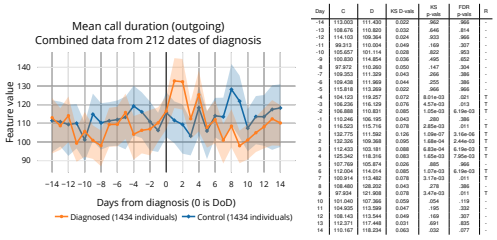
(b) Deviation from weekly average.



Day	C	D	KS D-vals	KS p-vals	FDR p-vals	R
-14	0.488	0.497	0.033	.604	.887	-
-13	0.503	0.487	0.033	.613	.887	-
-12	0.506	0.497	0.033	.637	.887	-
-11	0.489	0.514	0.041	.342	.887	-
-10	0.498	0.500	0.019	.992	.999	-
-9	0.501	0.483	0.031	.708	.934	-
-8	0.499	0.515	0.039	.408	.887	-
-7	0.498	0.499	0.049	.167	.887	-
-6	0.503	0.505	0.021	.975	.999	-
-5	0.512	0.508	0.016	.999	.999	-
-4	0.527	0.506	0.039	.382	.887	-
-3	0.516	0.505	0.032	.643	.887	-
-2	0.525	0.505	0.050	.139	.504	-
-1	0.552	0.494	0.089	4.77e-04	2.77e-03	T
0	0.591	0.491	0.149	1.55e-09	2.25e-08	T
1	0.568	0.486	0.161	3.08e-12	8.92e-11	T
2	0.552	0.494	0.100	6.75e-05	6.52e-04	T
3	0.552	0.497	0.091	3.87e-04	2.77e-03	T
4	0.544	0.492	0.086	9.95e-04	4.81e-03	T
5	0.524	0.501	0.062	.037	.154	-
6	0.492	0.491	0.025	.914	.999	-
7	0.501	0.501	0.022	.958	.999	-
8	0.492	0.499	0.034	.584	.887	-
9	0.488	0.491	0.026	.882	.999	-
10	0.501	0.492	0.039	.419	.887	-
11	0.496	0.499	0.035	.535	.887	-
12	0.496	0.483	0.033	.629	.887	-
13	0.498	0.484	0.037	.472	.887	-
14	0.476	0.487	0.027	.832	.999	-

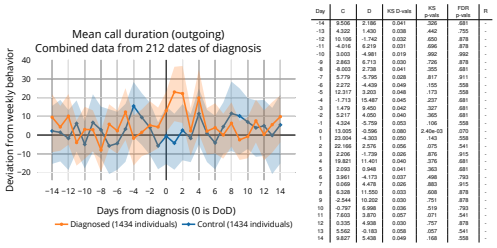
(c) Normalized fractional ranking.

Fig. S20. Feature graph and significance test results for ***mean call duration (incoming)*** using raw, detrended, and ranked data.



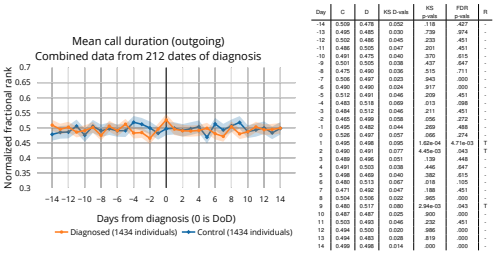
Day	C	D	KS D-vals	KS p-vals	FDR p-vals	R
-14	113.003	111.430	0.022	.962	.966	-
-13	108.676	110.820	0.032	.646	.814	-
-12	114.103	109.364	0.024	.933	.966	-
-11	99.313	110.004	0.049	.169	.307	-
-10	105.657	101.114	0.026	.822	.953	-
-9	105.833	114.854	0.036	.495	.652	-
-8	97.972	110.260	0.050	.147	.304	-
-7	109.353	111.329	0.043	.266	.396	-
-6	109.438	111.969	0.044	.255	.386	-
-5	115.818	113.269	0.022	.966	.966	-
-4	104.103	119.257	0.072	8.01e-03	.021	T
-3	106.236	116.129	0.076	4.57e-03	.013	T
-2	106.888	110.831	0.065	1.05e-03	6.19e-03	T
-1	110.246	106.195	0.043	.280	.396	-
0	116.523	115.716	0.078	2.85e-03	.011	T
1	132.775	111.592	0.126	1.09e-07	3.16e-06	T
2	132.336	109.368	0.095	1.68e-04	2.44e-03	T
3	112.433	103.181	0.088	6.83e-04	6.19e-03	T
4	125.342	118.316	0.083	1.65e-03	7.95e-03	T
5	107.769	105.874	0.026	.885	.966	-
6	112.004	114.014	0.065	1.07e-03	6.19e-03	T
7	100.914	113.482	0.078	3.17e-03	.011	T
8	109.480	109.202	0.043	.278	.396	-
9	97.934	121.906	0.078	3.47e-03	.011	T
10	101.040	107.366	0.059	.054	.119	-
11	104.935	113.599	0.047	.195	.332	-
12	108.143	113.544	0.049	.169	.307	-
13	112.371	117.448	0.031	.691	.835	-
14	110.167	118.234	0.063	.032	.077	-

(a) Raw feature value.



Day	C	D	KS D-vals	KS p-vals	FDR p-vals	R
-14	8.506	2.186	0.041	.326	.681	-
-13	4.322	1.430	0.038	.442	.755	-
-12	10.169	-1.742	0.032	.650	.878	-
-11	-4.016	6.219	0.031	.696	.878	-
-10	3.003	-4.981	0.019	.992	.992	-
-9	2.863	6.713	0.030	.726	.878	-
-8	-6.003	2.738	0.041	.355	.681	-
-7	5.779	5.795	0.028	.817	.911	-
-6	2.272	-4.439	0.049	.155	.558	-
-5	12.317	3.203	0.048	.173	.558	-
-4	-1.713	15.487	0.045	.237	.681	-
-3	1.479	9.450	0.042	.327	.681	-
-2	5.217	4.050	0.040	.365	.681	-
-1	4.324	-5.759	0.053	.106	.558	-
0	13.005	-0.596	0.080	2.49e-03	.070	-
1	23.004	-4.303	0.050	.143	.558	-
2	22.166	2.576	0.056	.075	.541	-
3	2.206	-1.739	0.026	.876	.915	-
4	19.821	11.401	0.040	.370	.681	-
5	2.093	0.948	0.041	.363	.681	-
6	3.961	4.173	0.037	.498	.793	-
7	0.069	4.478	0.026	.883	.915	-
8	6.328	11.550	0.033	.608	.878	-
9	-2.544	10.202	0.030	.751	.878	-
10	-0.797	8.998	0.036	.519	.793	-
11	7.603	3.870	0.057	.071	.541	-
12	0.335	4.938	0.030	.757	.878	-
13	5.562	-0.183	0.058	.057	.541	-
14	9.827	5.438	0.049	.169	.558	-

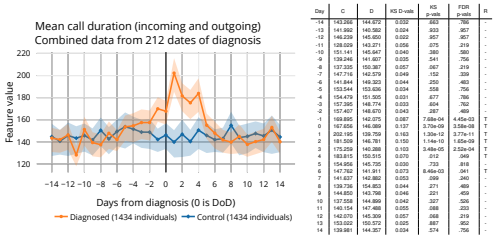
(b) Deviation from weekly average.



Day	C	D	KS D-vals	KS p-vals	FDR p-vals	R
-14	0.509	0.478	0.052	.118	.427	-
-13	0.495	0.485	0.030	.739	.974	-
-12	0.502	0.486	0.045	.233	.451	-
-11	0.486	0.505	0.047	.201	.451	-
-10	0.491	0.475	0.040	.370	.615	-
-9	0.501	0.505	0.038	.437	.647	-
-8	0.475	0.490	0.036	.515	.711	-
-7	0.506	0.497	0.023	.943	.000	-
-6	0.490	0.490	0.024	.917	.000	-
-5	0.512	0.491	0.046	.209	.451	-
-4	0.483	0.518	0.059	.013	.098	-
-3	0.484	0.512	0.046	.211	.451	-
-2	0.465	0.499	0.058	.056	.272	-
-1	0.495	0.482	0.044	.269	.488	-
0	0.526	0.497	0.057	.066	.274	-
1	0.495	0.498	0.095	1.62e-04	4.71e-03	T
2	0.490	0.491	0.077	4.45e-03	.043	T
3	0.488	0.496	0.051	.139	.448	-
4	0.491	0.503	0.036	.446	.647	-
5	0.498	0.499	0.040	.382	.615	-
6	0.480	0.513	0.067	.018	.105	-
7	0.471	0.492	0.047	.188	.451	-
8	0.504	0.506	0.022	.965	.000	-
9	0.480	0.517	0.080	2.94e-03	.043	T
10	0.487	0.487	0.025	.900	.000	-
11	0.503	0.493	0.046	.232	.451	-
12	0.494	0.500	0.020	.986	.000	-
13	0.494	0.483	0.028	.819	.000	-
14	0.499	0.498	0.014	.000	.000	-

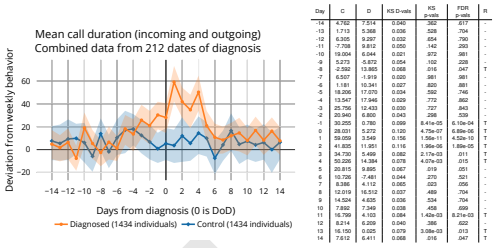
(c) Normalized fractional ranking.

Fig. S21. Feature graph and significance test results for ***mean call duration (outgoing)*** using raw, detrended, and ranked data.



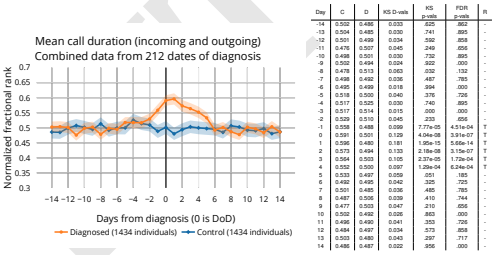
Day	C	D	KS D-vals	KS p-vals	FDR p-vals	R
-14	143.266	144.672	0.032	.663	.786	-
-13	141.992	140.382	0.024	.933	.957	-
-12	146.239	145.650	0.022	.957	.957	-
-11	128.509	143.271	0.056	.075	.219	-
-10	151.141	145.647	0.040	.380	.580	-
-9	139.246	141.607	0.035	.541	.756	-
-8	137.305	150.387	0.057	.067	.219	-
-7	147.716	142.579	0.049	.152	.339	-
-6	141.844	149.323	0.044	.250	.483	-
-5	153.544	153.636	0.034	.558	.756	-
-4	134.439	151.506	0.031	.677	.786	-
-3	157.395	148.774	0.033	.604	.762	-
-2	157.407	148.675	0.043	.287	.469	-
-1	169.895	142.075	0.087	7.68e-04	4.45e-03	T
0	167.636	146.089	0.137	3.70e-09	3.58e-08	T
1	202.195	139.759	0.163	1.30e-12	3.77e-11	T
2	181.508	146.781	0.150	1.14e-10	1.65e-09	T
3	175.259	140.288	0.103	3.48e-05	2.52e-04	T
4	183.815	150.515	0.070	.012	.049	T
5	154.956	145.735	0.030	.733	.818	-
6	147.762	141.911	0.073	8.46e-03	.041	T
7	141.637	142.882	0.053	.099	.240	-
8	139.736	154.853	0.044	.271	.469	-
9	144.859	143.798	0.046	.221	.459	-
10	137.558	144.699	0.042	.327	.526	-
11	140.134	147.498	0.055	.088	.233	-
12	142.070	145.309	0.057	.068	.219	-
13	153.022	150.572	0.025	.887	.952	-
14	139.981	144.357	0.034	.574	.756	-

(a) Raw feature value.



Day	C	D	KS D-vals	KS p-vals	FDR p-vals	R
-14	4.792	7.514	0.040	.382	.617	-
-13	1.713	5.368	0.036	.528	.704	-
-12	6.305	9.297	0.032	.654	.790	-
-11	-7.708	9.812	0.050	.142	.293	-
-10	19.004	6.044	0.021	.972	.981	-
-9	5.272	-5.872	0.054	.102	.228	-
-8	-2.592	13.865	0.068	.016	.047	T
-7	6.507	-1.919	0.020	.981	.981	-
-6	1.181	10.341	0.027	.820	.881	-
-5	18.206	17.070	0.034	.592	.746	-
-4	13.547	17.946	0.029	.772	.862	-
-3	25.756	12.433	0.030	.727	.843	-
-2	20.940	6.800	0.043	.298	.539	-
-1	30.255	0.780	0.099	8.41e-05	6.10e-04	T
0	28.031	5.272	0.120	4.75e-07	6.89e-06	T
1	59.059	3.540	0.156	1.56e-11	4.52e-10	T
2	41.835	11.951	0.116	1.96e-06	1.89e-05	T
3	34.730	5.499	0.082	2.17e-03	.011	T
4	50.226	14.384	0.078	4.07e-03	.015	T
5	20.815	9.895	0.067	.019	.051	T
6	10.726	-7.481	0.044	.270	.521	-
7	8.386	4.112	0.065	.023	.056	-
8	12.019	16.512	0.037	.499	.704	-
9	14.524	4.635	0.036	.534	.704	-
10	7.892	7.349	0.036	.458	.699	-
11	16.756	4.103	0.084	1.42e-03	8.21e-03	T
12	8.214	6.209	0.040	.386	.622	-
13	16.150	0.025	0.079	3.08e-03	.013	T
14	7.612	6.411	0.068	.016	.047	T

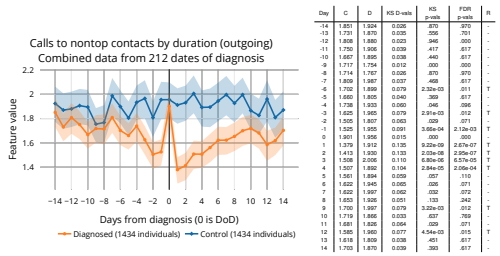
(b) Deviation from weekly average.



Day	C	D	KS D-vals	KS p-vals	FDR p-vals	R
-14	0.532	0.486	0.033	.625	.862	-
-13	0.504	0.485	0.030	.741	.895	-
-12	0.501	0.499	0.034	.592	.858	-
-11	0.476	0.507	0.045	.249	.656	-
-10	0.498	0.501	0.030	.732	.895	-
-9	0.502	0.494	0.024	.922	.000	-
-8	0.478	0.513	0.063	.032	.132	-
-7	0.498	0.492	0.036	.487	.785	-
-6	0.495	0.499	0.018	.994	.000	-
-5	0.518	0.500	0.040	.376	.726	-
-4	0.517	0.525	0.030	.740	.895	-
-3	0.517	0.514	0.015	.000	.000	-
-2	0.528	0.510	0.045	.233	.656	-
-1	0.558	0.488	0.099	7.77e-05	4.51e-04	T
0	0.591	0.501	0.128	4.04e-08	3.91e-07	T
1	0.596	0.480	0.181	1.95e-15	5.68e-14	T
2	0.573	0.494	0.133	2.18e-08	3.15e-07	T
3	0.564	0.503	0.105	2.37e-05	1.72e-04	T
4	0.552	0.500	0.097	1.29e-04	8.24e-04	T
5	0.533	0.497	0.059	.051	.185	-
6	0.492	0.495	0.042	.325	.725	-
7	0.501	0.485	0.036	.495	.785	-
8	0.487	0.506	0.039	.410	.744	-
9	0.472	0.503	0.047	.210	.656	-
10	0.502	0.492	0.026	.863	.000	-
11	0.496	0.490	0.041	.353	.726	-
12	0.494	0.497	0.034	.573	.858	-
13	0.503	0.480	0.043	.297	.717	-
14	0.486	0.487	0.022	.956	.000	-

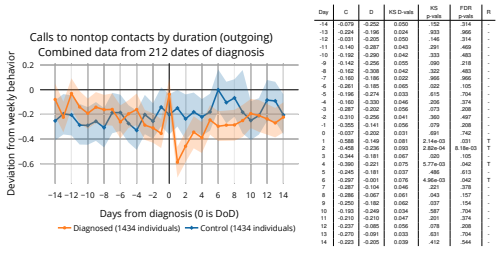
(c) Normalized fractional ranking.

Fig. S22. Feature graph and significance test results for ***mean call duration (both incoming and outgoing)*** using raw, detrended, and ranked data.



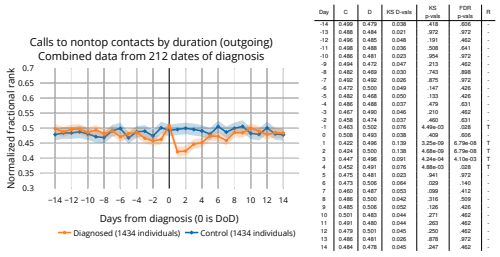
Day	C	D	KS D-vals	KS p-vals	FDR p-vals	R
-14	1.851	1.924	0.026	.870	.970	-
-13	1.731	1.870	0.035	.556	.701	-
-12	1.808	1.880	0.023	.946	.000	-
-11	1.750	1.906	0.039	.417	.617	-
-10	1.667	1.895	0.038	.440	.617	-
-9	1.717	1.754	0.012	.000	.000	-
-8	1.714	1.767	0.026	.870	.970	-
-7	1.809	1.987	0.037	.468	.617	-
-6	1.702	1.899	0.079	2.32e-03	.011	T
-5	1.660	1.805	0.040	.369	.617	-
-4	1.736	1.933	0.060	.046	.096	-
-3	1.625	1.965	0.079	2.91e-03	.012	T
-2	1.505	1.807	0.063	.029	.071	T
-1	1.525	1.955	0.091	3.66e-04	2.12e-03	T
0	1.901	1.956	0.015	.000	.000	-
1	1.379	1.912	0.135	9.22e-09	2.67e-07	T
2	1.413	1.930	0.133	2.03e-08	2.95e-07	T
3	1.538	2.006	0.110	6.83e-06	6.57e-05	T
4	1.507	1.892	0.104	2.84e-05	2.06e-04	T
5	1.561	1.894	0.059	.057	.110	-
6	1.622	1.945	0.065	.026	.071	-
7	1.622	1.997	0.062	.032	.072	-
8	1.653	1.926	0.051	.133	.268	-
9	1.700	1.997	0.079	3.22e-03	.012	T
10	1.719	1.866	0.033	.637	.769	-
11	1.681	1.826	0.064	.029	.071	-
12	1.585	1.960	0.077	4.54e-03	.015	T
13	1.618	1.859	0.038	.451	.617	-
14	1.703	1.870	0.039	.393	.617	-

(a) Raw feature value.



Day	C	D	KS D-vals	KS p-vals	FDR p-vals	R
-14	-0.079	-0.252	0.050	.152	.314	-
-13	-0.224	-0.196	0.024	.933	.966	-
-12	-0.031	-0.205	0.050	.146	.314	-
-11	-0.140	-0.287	0.040	.291	.469	-
-10	-0.192	-0.290	0.042	.333	.483	-
-9	-0.142	-0.256	0.032	.590	.778	-
-8	-0.162	-0.308	0.042	.322	.483	-
-7	-0.135	-0.186	0.022	.966	.966	-
-6	-0.261	-0.185	0.065	.022	.105	-
-5	-0.196	-0.274	0.033	.615	.704	-
-4	-0.160	-0.330	0.046	.206	.374	-
-3	-0.287	-0.202	0.056	.073	.208	-
-2	-0.310	-0.256	0.041	.360	.497	-
-1	-0.355	-0.141	0.056	.079	.208	-
0	-0.037	-0.202	0.031	.691	.742	-
1	-0.588	-0.149	0.081	2.14e-03	.031	T
2	-0.458	-0.226	0.093	2.82e-04	8.18e-03	T
3	-0.344	-0.181	0.067	.020	.105	-
4	-0.388	-0.221	0.075	5.77e-03	.042	T
5	-0.245	-0.181	0.037	.486	.613	-
6	-0.297	-0.001	0.076	4.96e-03	.042	T
7	-0.287	-0.104	0.046	.221	.378	-
8	-0.286	-0.067	0.061	.043	.157	-
9	-0.250	-0.182	0.042	.037	.154	-
10	-0.193	-0.249	0.034	.587	.704	-
11	-0.210	-0.210	0.047	.201	.374	-
12	-0.237	-0.085	0.056	.078	.208	-
13	-0.235	-0.091	0.033	.631	.704	-
14	-0.225	-0.205	0.039	.412	.544	-

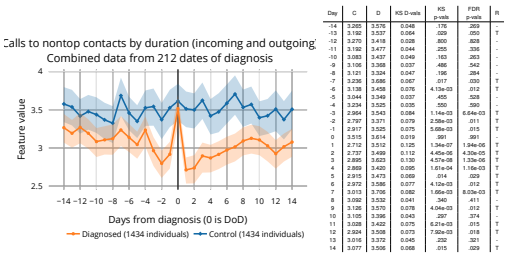
(b) Deviation from weekly average.



Day	C	D	KS D-vals	KS p-vals	FDR p-vals	R
-14	0.499	0.479	0.038	.418	.606	-
-13	0.488	0.484	0.031	.772	.772	-
-12	0.496	0.485	0.048	.191	.462	-
-11	0.498	0.488	0.036	.508	.641	-
-10	0.486	0.481	0.023	.954	.972	-
-9	0.494	0.472	0.047	.213	.462	-
-8	0.482	0.469	0.030	.743	.868	-
-7	0.492	0.492	0.026	.875	.972	-
-6	0.472	0.500	0.049	.147	.426	-
-5	0.482	0.468	0.050	.133	.426	-
-4	0.488	0.488	0.037	.479	.631	-
-3	0.467	0.490	0.046	.210	.462	-
-2	0.458	0.474	0.037	.460	.631	-
-1	0.463	0.502	0.076	4.49e-03	.028	T
0	0.508	0.493	0.038	.409	.606	-
1	0.422	0.496	0.139	3.25e-09	6.79e-08	T
2	0.424	0.500	0.136	4.68e-09	6.79e-08	T
3	0.441	0.496	0.091	4.24e-04	4.13e-03	T
4	0.452	0.491	0.076	4.88e-03	.028	T
5	0.475	0.481	0.023	.941	.972	-
6	0.473	0.506	0.064	.029	.140	-
7	0.460	0.487	0.053	.099	.412	-
8	0.486	0.500	0.042	.316	.509	-
9	0.485	0.506	0.052	.126	.426	-
10	0.501	0.483	0.044	.271	.462	-
11	0.491	0.480	0.044	.263	.462	-
12	0.479	0.501	0.045	.250	.462	-
13	0.486	0.481	0.026	.878	.972	-
14	0.484	0.478	0.045	.247	.462	-

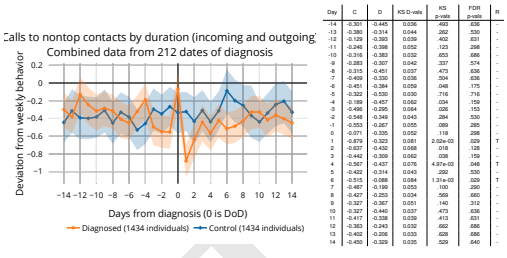
(c) Normalized fractional ranking.

Fig. S23. Feature graph and significance test results for ***nontop contacts by duration (outgoing)*** using raw, detrended, and ranked data.



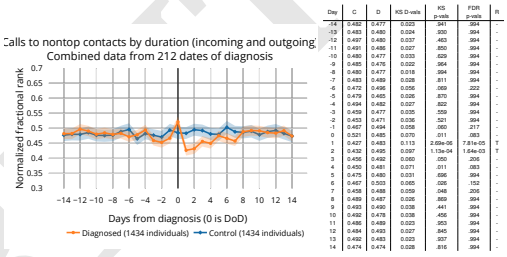
Day	C	D	KS D-vals	KS p-vals	FDR p-vals	R
-14	3.265	3.576	0.048	.176	.269	-
-13	3.192	3.537	0.064	.029	.050	T
-12	3.270	3.418	0.029	.800	.828	-
-11	3.192	3.477	0.044	.255	.336	-
-10	3.083	3.437	0.049	.163	.263	-
-9	3.166	3.368	0.037	.486	.542	-
-8	3.121	3.324	0.047	.196	.284	-
-7	3.236	3.686	0.067	.017	.030	T
-6	3.138	3.458	0.076	4.13e-03	.012	T
-5	3.044	3.349	0.037	.455	.528	-
-4	3.234	3.525	0.035	.550	.590	-
-3	2.964	3.545	0.084	1.14e-03	6.64e-03	T
-2	2.797	3.371	0.079	2.58e-03	.011	T
-1	2.917	3.525	0.075	5.68e-03	.015	T
0	3.515	3.614	0.019	.991	.991	-
1	2.712	3.512	0.125	1.34e-07	1.94e-06	T
2	2.737	3.499	0.112	4.45e-06	4.30e-05	T
3	2.885	3.623	0.130	4.57e-08	1.33e-06	T
4	2.869	3.420	0.095	1.61e-04	1.16e-03	T
5	2.915	3.473	0.069	.014	.029	T
6	2.972	3.586	0.077	4.12e-03	.012	T
7	3.013	3.706	0.082	1.66e-03	8.03e-03	T
8	3.092	3.532	0.041	.340	.411	-
9	3.126	3.570	0.078	4.04e-03	.012	T
10	3.105	3.396	0.043	.287	.374	-
11	3.028	3.422	0.075	6.21e-03	.015	T
12	2.924	3.508	0.073	7.92e-03	.018	T
13	3.016	3.372	0.045	.232	.321	-
14	3.077	3.506	0.068	.015	.029	T

(a) Raw feature value.



Day	C	D	KS D-vals	KS p-vals	FDR p-vals	R
-14	-0.301	-0.440	0.036	.493	.636	-
-13	-0.380	-0.314	0.044	.262	.530	-
-12	-0.129	-0.393	0.039	.402	.631	-
-11	-0.246	-0.398	0.052	.123	.298	-
-10	-0.316	-0.383	0.032	.653	.686	-
-9	-0.283	-0.307	0.042	.327	.574	-
-8	-0.315	-0.451	0.037	.473	.636	-
-7	-0.409	-0.330	0.036	.504	.636	-
-6	-0.451	-0.384	0.059	.048	.175	-
-5	-0.322	-0.530	0.030	.716	.716	-
-4	-0.189	-0.457	0.062	.034	.159	-
-3	-0.496	-0.295	0.064	.026	.153	-
-2	-0.548	-0.349	0.043	.284	.530	-
-1	-0.553	-0.267	0.055	.089	.285	-
0	-0.071	-0.335	0.052	.118	.298	-
1	-0.879	-0.323	0.081	2.02e-03	.029	T
2	-0.637	-0.432	0.068	.018	.128	-
3	-0.442	-0.309	0.062	.038	.159	-
4	-0.562	-0.427	0.076	4.97e-03	.048	T
5	-0.422	-0.314	0.043	.292	.530	-
6	-0.515	-0.088	0.084	1.31e-03	.029	T
7	-0.497	-0.199	0.053	.100	.290	-
8	-0.427	-0.253	0.034	.569	.660	-
9	-0.327	-0.367	0.051	.140	.312	-
10	-0.327	-0.440	0.037	.473	.636	-
11	-0.417	-0.338	0.039	.413	.631	-
12	-0.393	-0.243	0.032	.662	.686	-
13	-0.402	-0.256	0.033	.628	.686	-
14	-0.459	-0.329	0.035	.529	.640	-

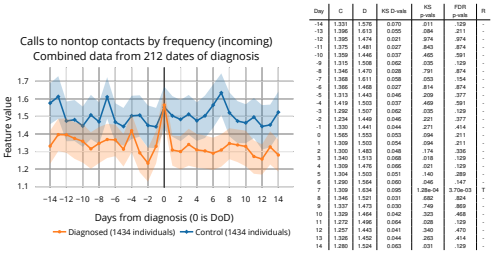
(b) Deviation from weekly average.



Day	C	D	KS D-vals	KS p-vals	FDR p-vals	R
-14	0.482	0.477	0.023	.941	.994	-
-13	0.483	0.480	0.024	.930	.994	-
-12	0.497	0.480	0.037	.463	.994	-
-11	0.491	0.488	0.027	.850	.994	-
-10	0.480	0.477	0.033	.629	.994	-
-9	0.485	0.476	0.022	.964	.994	-
-8	0.488	0.477	0.018	.984	.994	-
-7	0.483	0.489	0.028	.811	.994	-
-6	0.472	0.498	0.056	.069	.222	-
-5	0.479	0.485	0.026	.870	.994	-
-4	0.494	0.482	0.027	.822	.994	-
-3	0.459	0.477	0.035	.559	.994	-
-2	0.453	0.471	0.036	.521	.994	-
-1	0.497	0.494	0.058	.060	.217	-
0	0.521	0.485	0.070	.011	.083	-
1	0.427	0.483	0.113	2.69e-06	7.81e-05	T
2	0.432	0.495	0.097	1.13e-04	1.64e-03	T
3	0.456	0.492	0.069	.020	.206	-
4	0.450	0.481	0.071	.011	.083	-
5	0.475	0.480	0.031	.696	.994	-
6	0.467	0.503	0.065	.026	.152	-
7	0.458	0.488	0.059	.048	.206	-
8	0.486	0.487	0.026	.869	.994	-
9	0.493	0.490	0.038	.441	.994	-
10	0.492	0.478	0.038	.456	.994	-
11	0.486	0.489	0.023	.953	.994	-
12	0.484	0.493	0.027	.845	.994	-
13	0.482	0.483	0.023	.937	.994	-
14	0.474	0.474	0.029	.816	.994	-

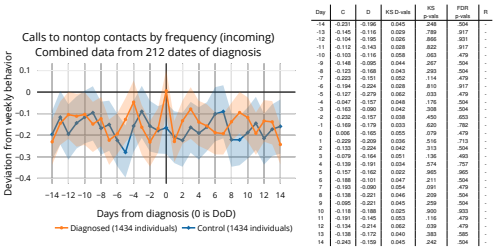
(c) Normalized fractional ranking.

Fig. S24. Feature graph and significance test results for ***nontop contacts by duration (both incoming and outgoing)*** using raw, detrended, and ranked data.



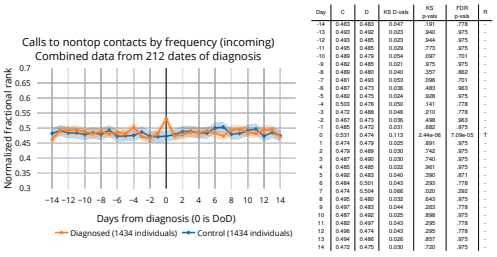
Day	C	D	KS D-vals	KS p-vals	FDR p-vals	R
-14	1.331	1.576	0.070	.011	.129	-
-13	1.396	1.613	0.055	.094	.211	-
-12	1.395	1.474	0.021	.974	.974	-
-11	1.375	1.481	0.027	.843	.874	-
-10	1.359	1.446	0.037	.465	.591	-
-9	1.315	1.508	0.062	.035	.129	-
-8	1.346	1.470	0.038	.791	.874	-
-7	1.368	1.611	0.058	.053	.154	-
-6	1.366	1.498	0.027	.814	.874	-
-5	1.313	1.443	0.046	.209	.377	-
-4	1.419	1.503	0.037	.469	.591	-
-3	1.292	1.507	0.062	.035	.129	-
-2	1.234	1.449	0.046	.221	.377	-
-1	1.330	1.441	0.044	.271	.414	-
0	1.565	1.553	0.053	.094	.211	-
1	1.309	1.503	0.054	.094	.211	-
2	1.300	1.483	0.048	.174	.336	-
3	1.340	1.513	0.068	.018	.129	-
4	1.309	1.476	0.066	.021	.129	-
5	1.304	1.503	0.051	.140	.289	-
6	1.290	1.564	0.060	.046	.147	-
7	1.309	1.634	0.095	1.28e-04	3.70e-03	T
8	1.346	1.521	0.051	.682	.824	-
9	1.337	1.473	0.030	.749	.869	-
10	1.329	1.464	0.042	.323	.468	-
11	1.272	1.496	0.064	.028	.129	-
12	1.257	1.443	0.041	.340	.470	-
13	1.326	1.452	0.044	.263	.414	-
14	1.280	1.524	0.063	.031	.129	-

(a) Raw feature value.



Day	C	D	KS D-vals	KS p-vals	FDR p-vals	R
-14	-0.231	-0.196	0.045	.346	.504	-
-13	-0.145	-0.116	0.029	.789	.917	-
-12	-0.104	-0.195	0.026	.899	.931	-
-11	-0.112	-0.143	0.028	.822	.917	-
-10	-0.103	-0.116	0.058	.063	.479	-
-9	-0.148	-0.095	0.044	.267	.504	-
-8	-0.123	-0.168	0.043	.293	.504	-
-7	-0.223	-0.151	0.052	.114	.479	-
-6	-0.194	-0.224	0.028	.810	.917	-
-5	-0.127	-0.279	0.062	.033	.479	-
-4	-0.047	-0.157	0.048	.176	.504	-
-3	-0.163	-0.090	0.042	.308	.504	-
-2	-0.232	-0.157	0.038	.455	.601	-
-1	-0.169	-0.179	0.033	.620	.782	-
0	0.006	-0.165	0.055	.079	.479	-
1	-0.229	-0.209	0.036	.516	.713	-
2	-0.133	-0.224	0.042	.313	.504	-
3	-0.079	-0.164	0.051	.136	.493	-
4	-0.139	-0.191	0.034	.574	.757	-
5	-0.157	-0.162	0.022	.965	.965	-
6	-0.188	-0.101	0.047	.211	.504	-
7	-0.193	-0.099	0.054	.091	.479	-
8	-0.138	-0.221	0.046	.209	.504	-
9	-0.095	-0.221	0.045	.259	.504	-
10	-0.118	-0.188	0.025	.900	.933	-
11	-0.191	-0.145	0.053	.116	.479	-
12	-0.134	-0.214	0.062	.039	.479	-
13	-0.138	-0.172	0.040	.383	.585	-
14	-0.243	-0.159	0.045	.242	.504	-

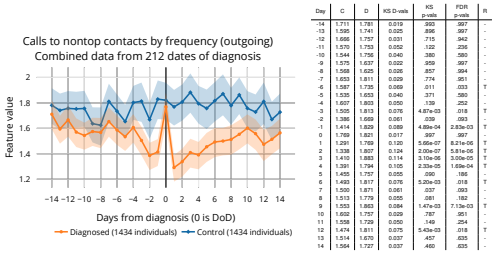
(b) Deviation from weekly average.



Day	C	D	KS D-vals	KS p-vals	FDR p-vals	R
-14	0.483	0.483	0.047	.191	.778	-
-13	0.493	0.492	0.023	.940	.975	-
-12	0.493	0.485	0.023	.944	.975	-
-11	0.495	0.485	0.029	.773	.975	-
-10	0.489	0.479	0.054	.097	.701	-
-9	0.482	0.485	0.021	.975	.975	-
-8	0.488	0.480	0.040	.357	.852	-
-7	0.481	0.493	0.053	.096	.701	-
-6	0.487	0.473	0.036	.483	.963	-
-5	0.482	0.475	0.024	.928	.975	-
-4	0.503	0.476	0.050	.141	.778	-
-3	0.472	0.488	0.046	.210	.778	-
-2	0.467	0.473	0.036	.498	.963	-
-1	0.485	0.472	0.031	.682	.975	-
0	0.531	0.474	0.113	2.44e-06	7.09e-05	T
1	0.474	0.479	0.025	.891	.975	-
2	0.479	0.489	0.030	.742	.975	-
3	0.487	0.490	0.030	.740	.975	-
4	0.485	0.485	0.022	.961	.975	-
5	0.492	0.483	0.040	.390	.871	-
6	0.484	0.501	0.043	.293	.778	-
7	0.474	0.504	0.066	.020	.292	-
8	0.484	0.480	0.032	.643	.975	-
9	0.497	0.483	0.044	.263	.778	-
10	0.487	0.492	0.025	.898	.975	-
11	0.482	0.497	0.043	.295	.778	-
12	0.496	0.474	0.043	.295	.778	-
13	0.494	0.486	0.026	.857	.975	-
14	0.472	0.475	0.030	.720	.975	-

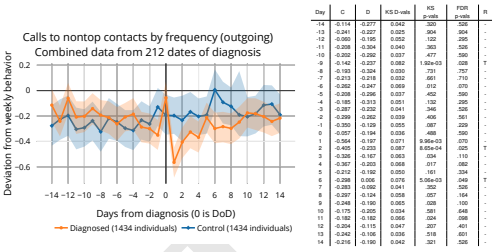
(c) Normalized fractional ranking.

Fig. S25. Feature graph and significance test results for ***nontop contacts by frequency (incoming)*** using raw, detrended, and ranked data.



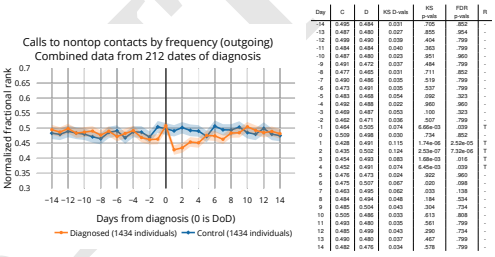
Day	C	D	KS D-vals	KS p-vals	FDR p-vals	R
-14	1.711	1.781	0.019	.993	.997	-
-13	1.598	1.741	0.025	.896	.997	-
-12	1.666	1.757	0.031	.715	.942	-
-11	1.570	1.753	0.052	.122	.256	-
-10	1.544	1.756	0.040	.360	.560	-
-9	1.575	1.637	0.022	.959	.997	-
-8	1.568	1.625	0.026	.857	.994	-
-7	1.653	1.811	0.029	.774	.951	-
-6	1.587	1.735	0.068	.011	.033	T
-5	1.535	1.653	0.040	.371	.580	-
-4	1.607	1.803	0.050	.139	.282	-
-3	1.555	1.813	0.076	4.87e-03	.018	T
-2	1.386	1.669	0.061	.039	.093	T
-1	1.414	1.829	0.089	4.89e-04	2.83e-03	T
0	1.769	1.821	0.017	.997	.997	-
1	1.291	1.769	0.120	5.66e-07	8.21e-06	T
2	1.338	1.807	0.124	2.00e-07	5.81e-06	T
3	1.410	1.883	0.114	3.10e-06	3.00e-05	T
4	1.391	1.794	0.105	2.33e-05	1.69e-04	T
5	1.455	1.757	0.055	.090	.186	-
6	1.493	1.817	0.076	5.20e-03	.018	T
7	1.500	1.871	0.061	.037	.093	-
8	1.513	1.778	0.055	.091	.182	-
9	1.553	1.863	0.084	1.47e-03	7.13e-03	T
10	1.602	1.757	0.029	.787	.951	-
11	1.556	1.729	0.050	.149	.254	-
12	1.474	1.811	0.075	5.43e-03	.018	T
13	1.514	1.670	0.037	.457	.635	-
14	1.564	1.727	0.037	.460	.635	-

(a) Raw feature value.



Day	C	D	KS D-vals	KS p-vals	FDR p-vals	R
-14	-0.114	-0.277	0.042	.320	.526	-
-13	-0.241	-0.227	0.025	.904	.904	-
-12	-0.060	-0.195	0.052	.122	.295	-
-11	-0.208	-0.304	0.040	.363	.536	-
-10	-0.202	-0.292	0.037	.477	.590	-
-9	-0.142	-0.237	0.062	1.92e-03	.028	T
-8	-0.193	-0.324	0.030	.731	.757	-
-7	-0.213	-0.218	0.032	.661	.710	-
-6	-0.262	-0.247	0.069	.012	.070	-
-5	-0.208	-0.296	0.037	.452	.590	-
-4	-0.185	-0.313	0.051	.132	.295	-
-3	-0.287	-0.232	0.041	.346	.526	-
-2	-0.299	-0.262	0.039	.406	.561	-
-1	-0.350	-0.129	0.055	.087	.229	-
0	-0.057	-0.194	0.036	.488	.590	-
1	-0.564	-0.197	0.071	9.96e-03	.070	-
2	-0.405	-0.233	0.087	8.65e-04	.025	T
3	-0.326	-0.167	0.063	.034	.110	-
4	-0.367	-0.203	0.068	.017	.082	-
5	-0.212	-0.192	0.050	.161	.334	-
6	-0.298	0.006	0.076	5.06e-03	.049	T
7	-0.283	-0.092	0.041	.352	.526	-
8	-0.297	-0.124	0.058	.057	.164	-
9	-0.246	-0.190	0.065	.028	.100	-
10	-0.175	-0.205	0.034	.581	.648	-
11	-0.192	-0.182	0.066	.024	.098	-
12	-0.204	-0.115	0.047	.207	.401	-
13	-0.242	-0.196	0.036	.518	.601	-
14	-0.216	-0.190	0.042	.321	.526	-

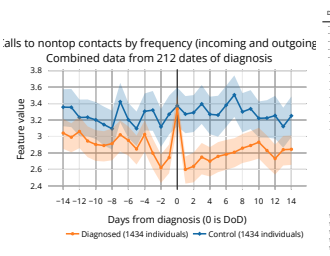
(b) Deviation from weekly average.



Day	C	D	KS D-vals	KS p-vals	FDR p-vals	R
-14	0.495	0.484	0.031	.705	.852	-
-13	0.487	0.480	0.027	.855	.954	-
-12	0.499	0.490	0.039	.404	.799	-
-11	0.494	0.484	0.040	.363	.799	-
-10	0.487	0.480	0.023	.951	.960	-
-9	0.491	0.472	0.037	.484	.799	-
-8	0.477	0.485	0.031	.711	.852	-
-7	0.490	0.486	0.035	.519	.799	-
-6	0.473	0.491	0.035	.537	.799	-
-5	0.483	0.468	0.054	.092	.323	-
-4	0.482	0.488	0.022	.960	.960	-
-3	0.469	0.487	0.053	.100	.323	-
-2	0.462	0.471	0.036	.507	.799	-
-1	0.464	0.505	0.074	6.66e-03	.039	T
0	0.509	0.498	0.030	.734	.852	-
1	0.428	0.491	0.115	1.74e-06	2.52e-05	T
2	0.435	0.502	0.124	2.53e-07	7.32e-06	T
3	0.454	0.493	0.083	1.68e-03	.016	T
4	0.452	0.491	0.074	6.45e-03	.039	T
5	0.476	0.473	0.024	.922	.960	-
6	0.475	0.507	0.067	.020	.098	-
7	0.463	0.495	0.062	.033	.138	-
8	0.484	0.494	0.048	.184	.334	-
9	0.485	0.504	0.043	.304	.734	-
10	0.506	0.486	0.033	.613	.808	-
11	0.493	0.480	0.035	.561	.799	-
12	0.485	0.499	0.043	.290	.734	-
13	0.467	0.480	0.037	.467	.799	-
14	0.482	0.476	0.034	.578	.799	-

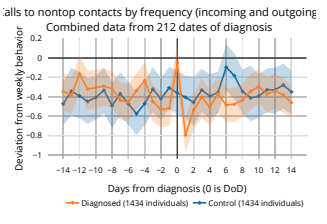
(c) Normalized fractional ranking.

Fig. S26. Feature graph and significance test results for ***nontop contacts by frequency (outgoing)*** using raw, detrended, and ranked data.



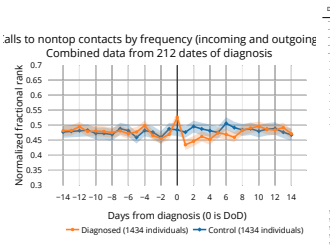
Day	C	D	KS D-vals	KS p-vals	FDR p-vals	R
-14	3.042	3.357	0.047	.190	.239	-
-13	2.991	3.354	0.055	.085	.144	-
-12	3.061	3.231	0.029	.768	.796	-
-11	2.945	3.234	0.061	.041	.074	-
-10	2.902	3.202	0.043	.281	.317	-
-9	2.890	3.144	0.042	.328	.352	-
-8	2.914	3.094	0.047	.188	.239	-
-7	3.021	3.422	0.064	.024	.046	T
-6	2.953	3.203	0.051	.122	.197	-
-5	2.848	3.096	0.050	.133	.203	-
-4	3.025	3.306	0.043	.284	.317	-
-3	2.797	3.320	0.080	2.55e-03	.012	T
-2	2.620	3.118	0.078	3.36e-03	.014	T
-1	2.744	3.270	0.071	.011	.028	T
0	3.334	3.374	0.018	.995	.995	-
1	2.600	3.272	0.110	5.54e-06	8.03e-05	T
2	2.638	3.290	0.098	8.08e-05	5.85e-04	T
3	2.750	3.395	0.117	1.41e-06	4.09e-05	T
4	2.700	3.270	0.108	1.12e-05	1.08e-04	T
5	2.758	3.260	0.077	4.52e-03	.015	T
6	2.783	3.381	0.076	5.05e-03	.015	T
7	2.809	3.506	0.087	6.89e-04	4.00e-03	T
8	2.859	3.300	0.048	.170	.235	-
9	2.891	3.336	0.076	4.92e-03	.015	T
10	2.931	3.221	0.047	.210	.254	-
11	2.830	3.225	0.069	.015	.034	T
12	2.731	3.254	0.067	.019	.040	T
13	2.840	3.122	0.049	.161	.233	-
14	2.845	3.251	0.070	.011	.028	T

(a) Raw feature value.



Day	C	D	KS D-vals	KS p-vals	FDR p-vals	R
-14	-0.345	-0.473	0.037	.463	.671	-
-13	-0.386	-0.343	0.019	.992	.992	-
-12	-0.165	-0.390	0.041	.348	.593	-
-11	-0.320	-0.448	0.035	.559	.737	-
-10	-0.305	-0.408	0.030	.740	.796	-
-9	-0.290	-0.333	0.035	.537	.737	-
-8	-0.316	-0.491	0.032	.649	.795	-
-7	-0.436	-0.368	0.045	.229	.532	-
-6	-0.456	-0.471	0.045	.239	.532	-
-5	-0.335	-0.575	0.042	.313	.593	-
-4	-0.232	-0.470	0.057	.070	.403	-
-3	-0.450	-0.322	0.050	.152	.442	-
-2	-0.531	-0.420	0.039	.415	.633	-
-1	-0.519	-0.308	0.040	.372	.600	-
0	-0.051	-0.360	0.051	.132	.426	-
1	-0.793	-0.405	0.067	.018	.262	-
2	-0.538	-0.457	0.060	.048	.348	-
3	-0.404	-0.331	0.055	.088	.426	-
4	-0.557	-0.394	0.060	.047	.348	-
5	-0.369	-0.354	0.031	.700	.795	-
6	-0.486	-0.095	0.071	.011	.262	-
7	-0.476	-0.182	0.051	.130	.426	-
8	-0.435	-0.345	0.031	.702	.795	-
9	-0.342	-0.412	0.042	.316	.593	-
10	-0.293	-0.393	0.047	.198	.523	-
11	-0.373	-0.326	0.052	.123	.426	-
12	-0.338	-0.329	0.031	.696	.795	-
13	-0.380	-0.278	0.030	.722	.795	-
14	-0.455	-0.349	0.042	.329	.593	-

(b) Deviation from weekly average.



Day	C	D	KS D-vals	KS p-vals	FDR p-vals	R
-14	0.482	0.478	0.032	.648	.997	-
-13	0.483	0.480	0.024	.927	.997	-
-12	0.495	0.482	0.030	.739	.997	-
-11	0.480	0.484	0.018	.997	.997	-
-10	0.481	0.473	0.027	.842	.997	-
-9	0.479	0.473	0.018	.996	.997	-
-8	0.476	0.469	0.025	.883	.997	-
-7	0.481	0.489	0.029	.753	.997	-
-6	0.472	0.482	0.039	.387	.935	-
-5	0.478	0.460	0.040	.349	.935	-
-4	0.498	0.483	0.032	.655	.997	-
-3	0.466	0.477	0.034	.571	.997	-
-2	0.456	0.460	0.031	.694	.997	-
-1	0.470	0.488	0.043	.279	.894	-
0	0.526	0.485	0.071	9.31e-03	.090	-
1	0.435	0.477	0.084	1.11e-03	.016	T
2	0.447	0.496	0.088	6.45e-04	.016	T
3	0.463	0.488	0.048	.186	.673	-
4	0.453	0.482	0.054	.095	.459	-
5	0.475	0.476	0.023	.943	.997	-
6	0.470	0.506	0.061	.041	.241	-
7	0.460	0.492	0.069	.012	.090	-
8	0.483	0.485	0.035	.543	.997	-
9	0.491	0.489	0.037	.484	.997	-
10	0.498	0.480	0.050	.148	.614	-
11	0.487	0.487	0.021	.975	.997	-
12	0.494	0.489	0.024	.935	.997	-
13	0.493	0.477	0.036	.495	.997	-
14	0.472	0.468	0.041	.357	.935	-

(c) Normalized fractional ranking.

Fig. S27. Feature graph and significance test results for ***nontop contacts by frequency (both incoming and outgoing)*** using raw, detrended, and ranked data.