



UNIVERSITY OF LEEDS

# Use of routinely collected healthcare data for analysis of indoor mobility and interaction

**Professor Ed Manley**

**Professor of Urban Analytics**  
School of Geography  
University of Leeds

**Turing Fellow**  
The Alan Turing Institute for Data  
Science and Artificial Intelligence

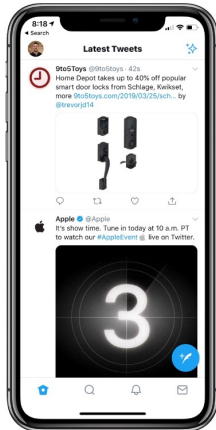
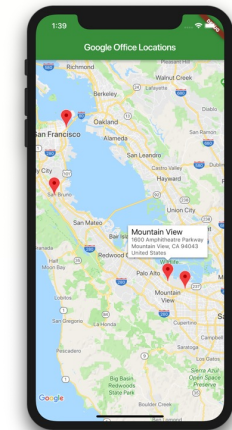
# Mobility Data

## Digital Footprints



UNIVERSITY OF LEEDS

- Available for many samples, large cohorts
- Often **passively** generated, as part of another activity
- Often associated **space** and **time** data
- Anonymised **individual** data and metadata can provide context (e.g. job role, derived 'home' areas)
- Produces observations in **aggregate** in situ at large scales
- Allows some **inference** of context affecting behaviour
- Usually lower **sampling bias** issues than other survey approaches
- **Privacy** preservation requires care
- Risks remain around how findings are used to penalise populations, so **proportionality** in use is key



# Mobility Data

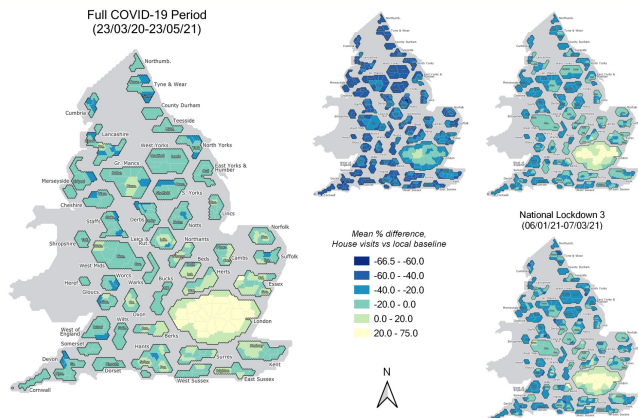
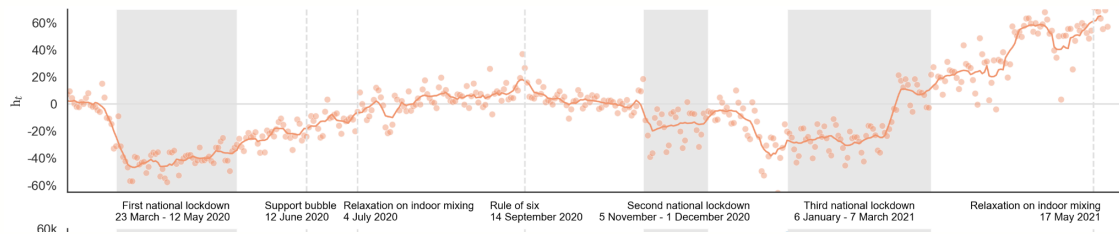
Digital Footprints



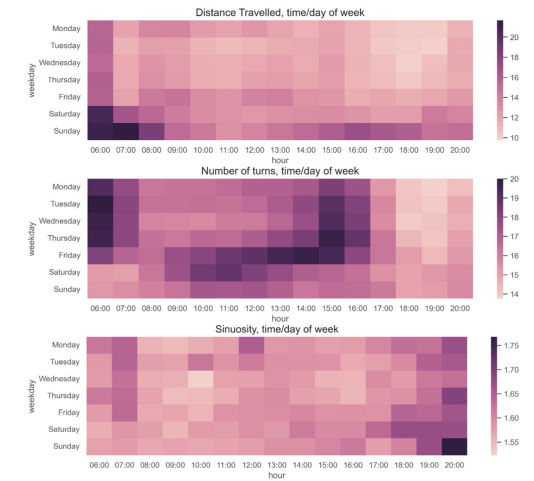
UNIVERSITY OF LEEDS

CDRC Leeds holds new, national scale mobility data from **Spectus** (app-derived) and **Wejo** (vehicle GPS)

## Household mixing policy adherence (Spectus)



## Vehicle routing behaviour (Wejo)



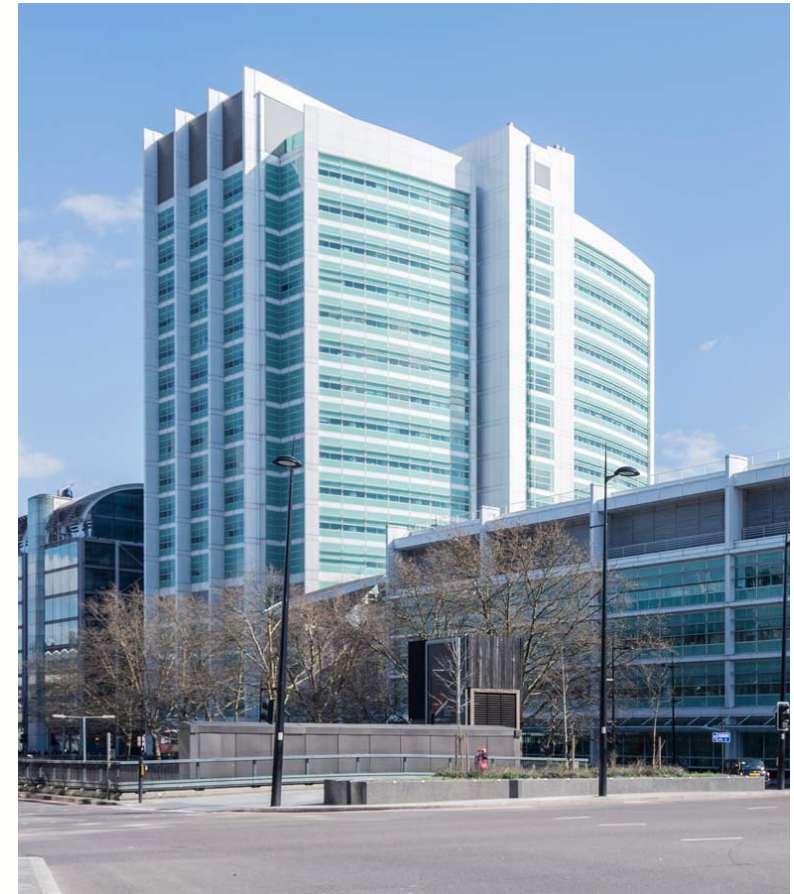
# Healthcare Worker Mobility

During COVID-19



UNIVERSITY OF LEEDS

- Healthcare workers have been at **significant risk** of catching COVID-19 and of inadvertently **transmitting** it to patients and colleagues - mobility and interaction, as well all know, are fundamental in transmission
- Our study focused on staff COVID-19 testing and tracking at a hospital, with a view to developing policies to reduce **hospital transmission** and **healthcare worker risk**
- **But how does one collect mobility and interaction data in a locked-down hospital?**
- We developed a **new set of proxy indicators of activity**, from **routinely collected** staff data to derive insights into staff mobility
- Questions: Who faces the most risk? Which environments and working practices affect risk? Can we estimate where and when transmission occurred?



# Trajectory Reconstruction

Data Sources



UNIVERSITY OF LEEDS

Electronic Staff Records – N = 4148 individuals

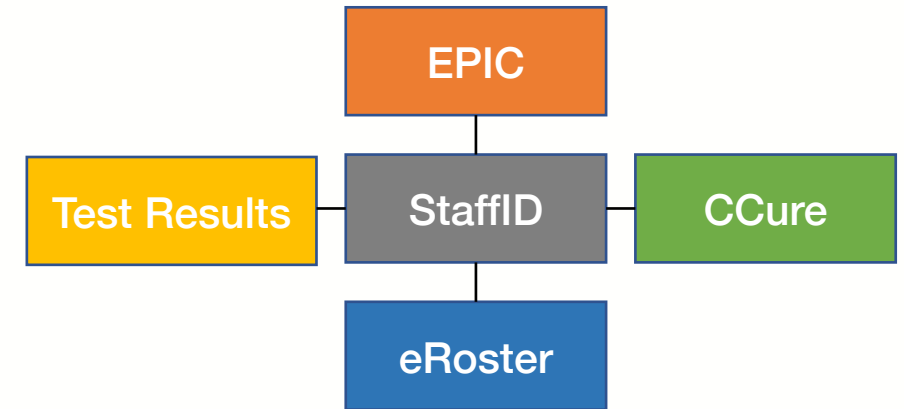
EPIC electronic healthcare records – Records of patient interactions with space (down to bay and bed location) and time

Card Access logs (CCure) – Door access with space (door access points) and time

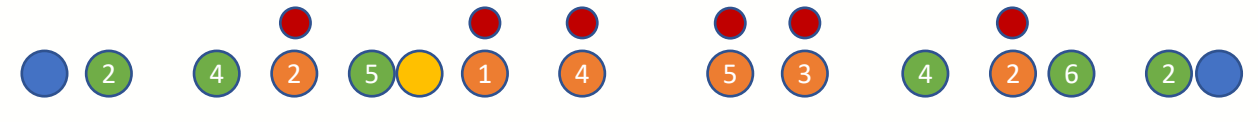
eRoster – including time of shift

COVID-19 tests – N= 25632 (negative = 24809, positive = 823; 3%)

Timeline – January 2020 to May 2021



Staff 'Footprint'



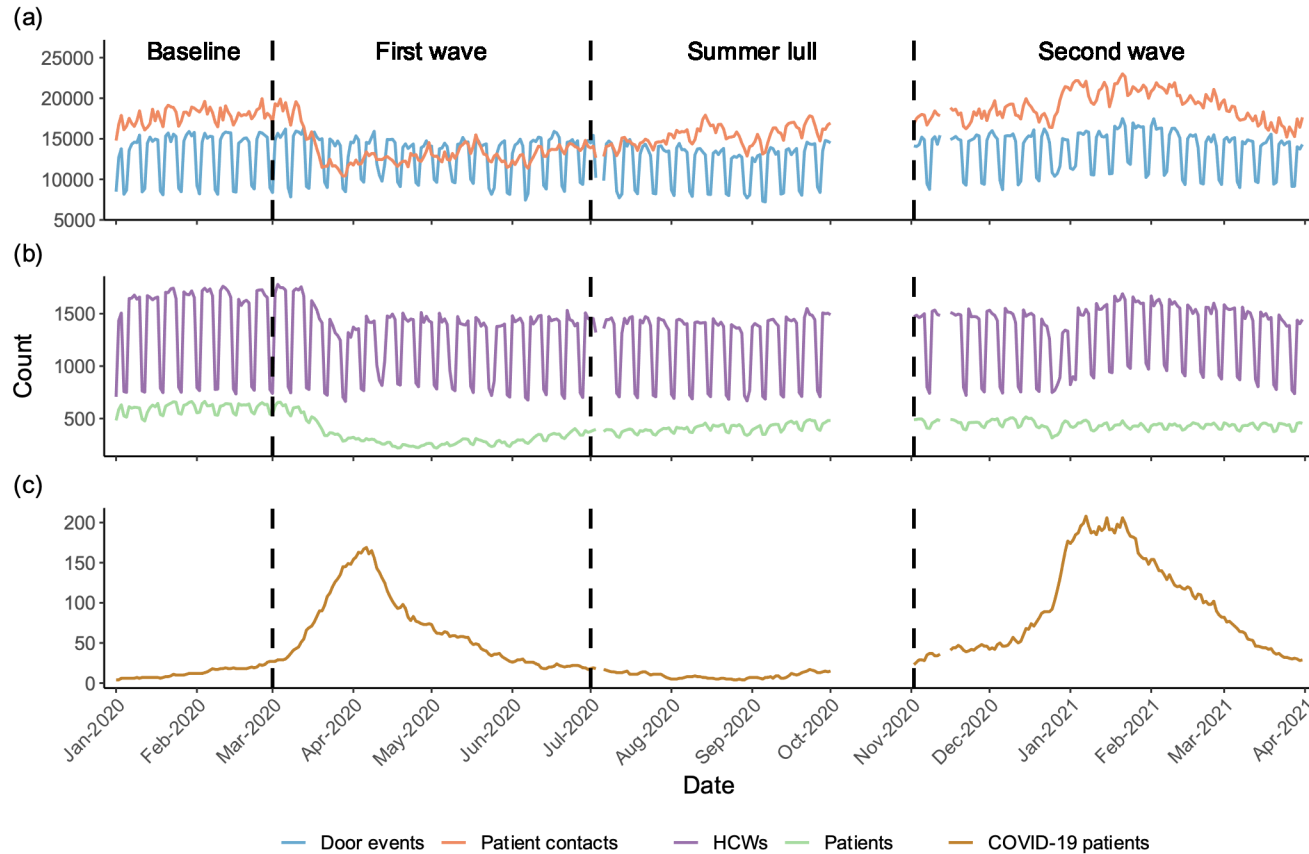
time

# Temporal Variation

Healthcare workers



UNIVERSITY OF LEEDS



Door events + Patient contacts

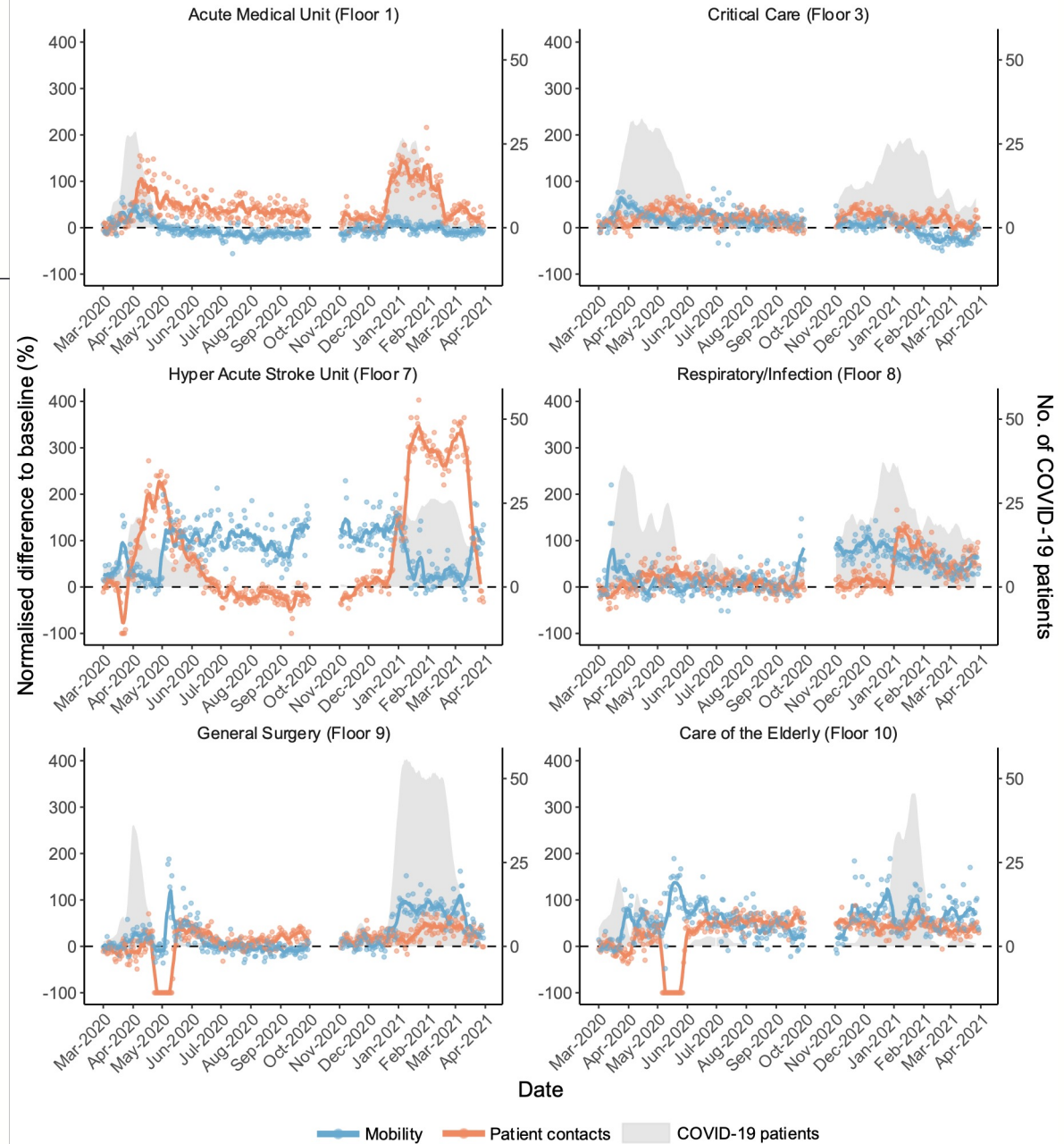
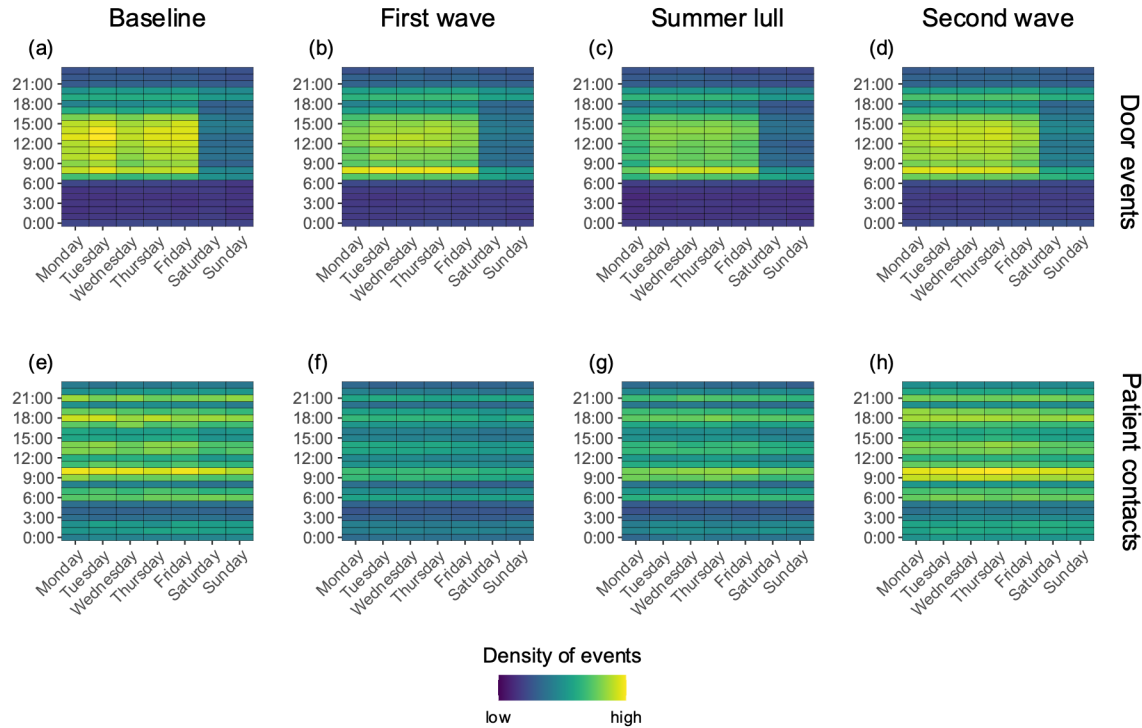
Patient counts + HCW counts

COVID-19 patients

# Temporal Variation

Floor changes

Activity normalised against a pre-COVID baseline



# Spatial Variation

Granularity



UNIVERSITY OF LEEDS

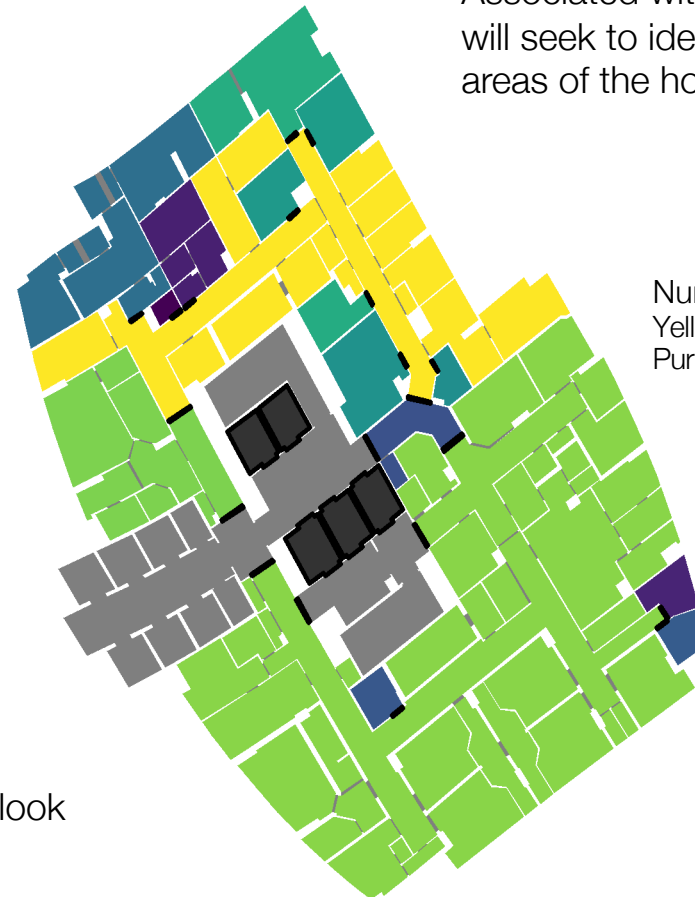
Tower Floor 5  
17 Zones



Gold star = Staff only  
Blue star = Public access  
Remainder = Clinical

Card access data allows us to look at movement into floor zones

Associated with test data, we will seek to identify 'riskier' areas of the hospital



Number of Entries  
Yellow = Highest  
Purple = Lowest

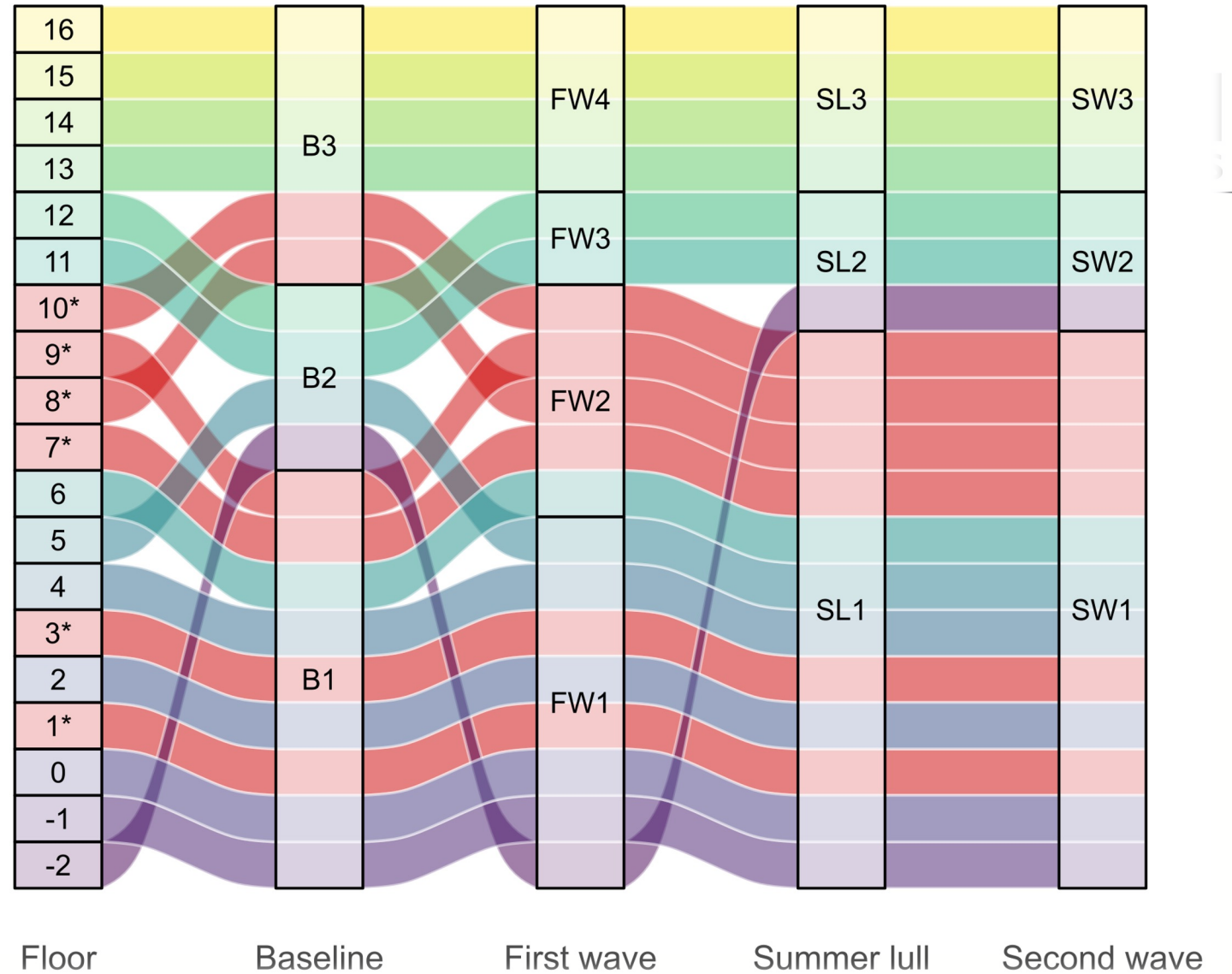


# Spatial Variation

## Floor interactions

Figure shows clusters of interactions between floors, through staff movements

Indications of floor reorganisation and staff cohorting as COVID progressed



# Risk Factors

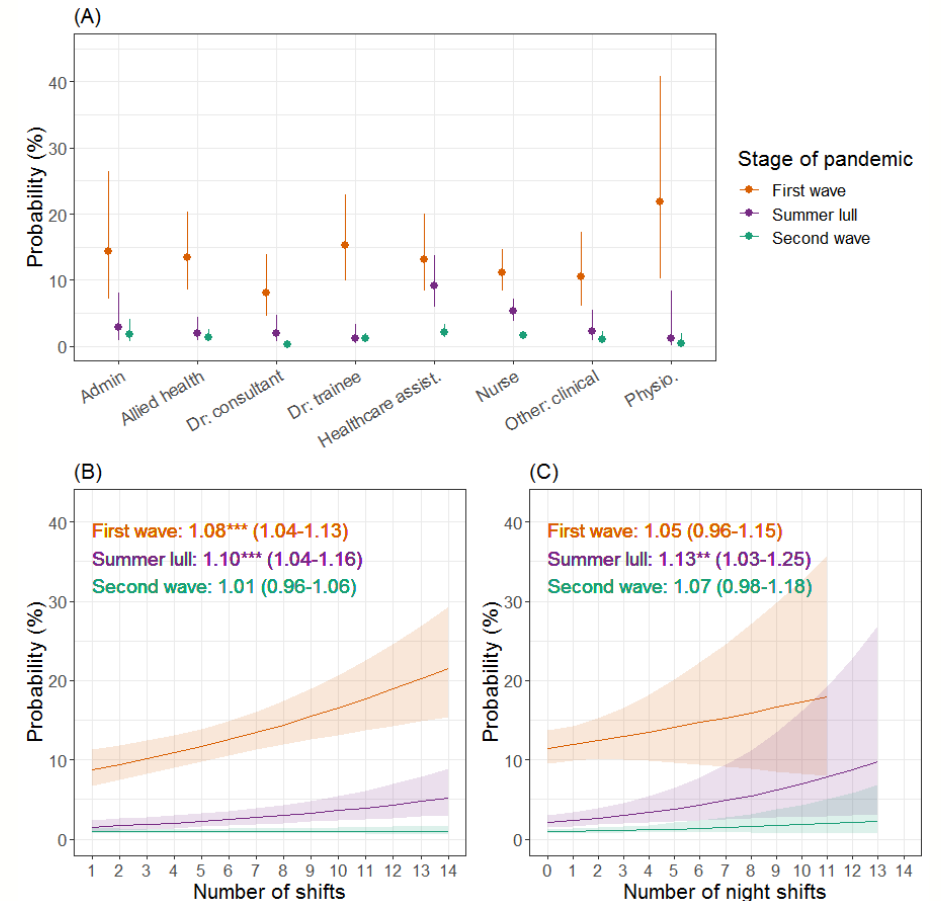
## COVID risk



UNIVERSITY OF LEEDS

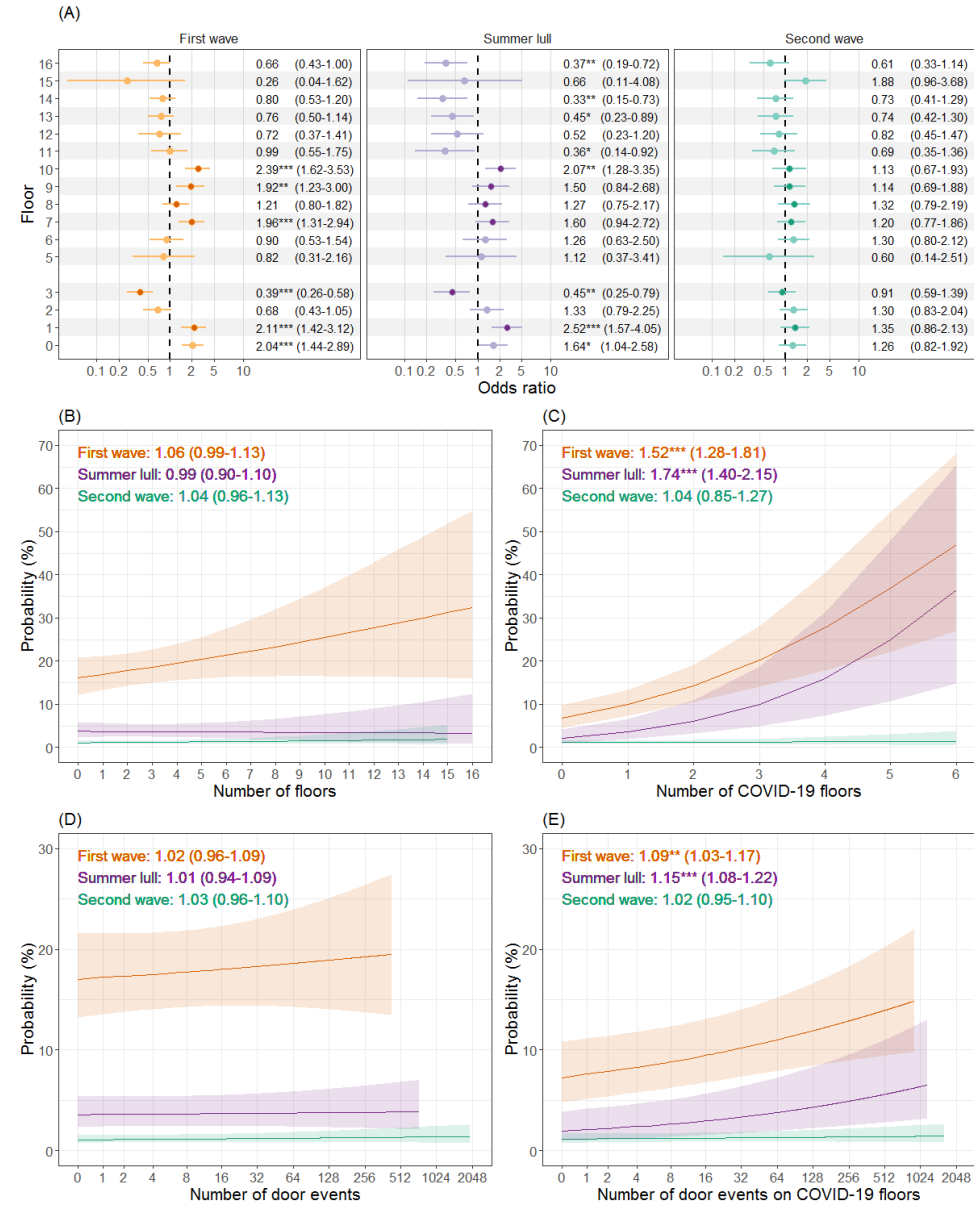
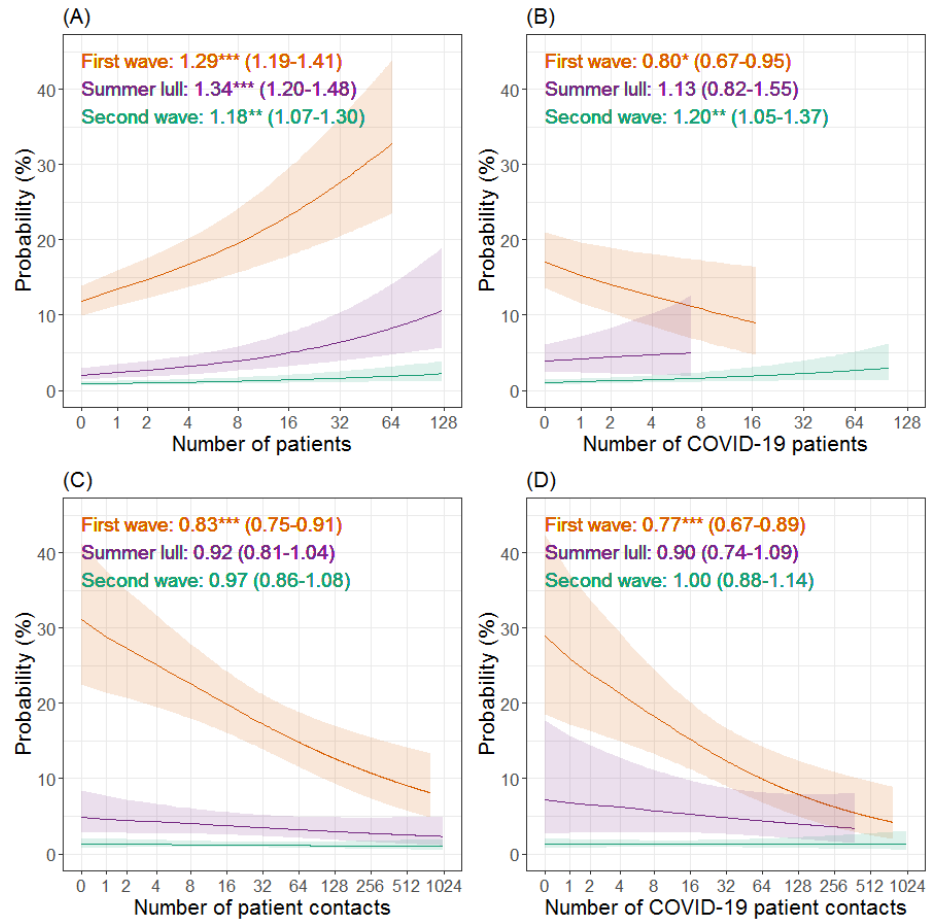
Table S6. Predictors for the risk of healthcare workers testing positive for COVID-19 during the pandemic. The odds ratio, 95% confidence intervals and statistical significance is reported from models using metrics derived from either 14 days, 7 days or 2 days of data prior to a COVID-19 test being taken. Specific post-hoc comparisons of interest are reported. Results for the models using 7 and 2 days of data are presented in bold if the result is different to that in the models using 14 days of data. For the model using 2 days of data, tests for the physiotherapist staff group were excluded as the records were too few for the model to be stable, and the model for activity on floor 5 was not included as again the model was not stable.

Predictors	Contrast / Reference	14 Days (n = 28,909)	7 Days (n = 26,243)	2 Days (n = 18,688)
Age	-	0.99 (0.98-1.00)	0.99 (0.98-1.00)	0.99 (0.98-1.00)
Ethnicity	BAME / White	<b>1.75 (1.40-2.20)***</b>	<b>1.79 (1.42-2.26)***</b>	<b>1.66 (1.28-2.15)***</b>
Period	First wave / Summer lull	2.86 (2.25-3.65)***	2.94 (2.29-3.77)***	3.31 (2.48-4.42)***
	First wave / Second wave	9.35 (7.41-11.81)***	9.68 (7.59-12.34)***	10.92 (8.24-14.46)***
	Summer lull / Second wave	3.27 (2.53-4.23)***	3.29 (2.52-4.30)***	3.30 (2.41-4.52)***
Admin	First wave / Summer lull	5.83 (1.24-27.51)*	5.04 (1.04-24.51)*	<b>5.44 (0.66-45.01)</b>
	First wave / Second wave	9.51 (2.58-35.11)***	8.16 (2.14-31.16)***	5.13 (1.15-22.78)*
	Summer lull / Second wave	1.63 (0.33-8)	1.62 (0.33-7.98)	0.94 (0.13-7.03)
Allied health professional	First wave / Summer lull	7.84 (2.66-23.1)***	9.16 (2.92-28.76)***	7.85 (2.26-27.26)***
	First wave / Second wave	11.78 (4.84-28.64)***	11.29 (4.61-27.66)***	12.5 (4.56-34.27)***
	Summer lull / Second wave	1.5 (0.45-5)	1.23 (0.35-4.3)	1.59 (0.4-6.28)
Doctor: consultant	First wave / Summer lull	4.47 (1.3-15.38)*	4.5 (1.29-15.64)*	8.78 (1.8-42.94)**
	First wave / Second wave	32.37 (6.91-151.71)***	31.12 (6.59-147.07)***	42.22 (6.68-266.73)***
	Summer lull / Second wave	7.23 (1.28-41.03)*	6.92 (1.22-39.28)*	<b>4.81 (0.52-44.06)</b>



Mixed effects logistic regression on positive test, under a DAGs framework

Assessing risk based on personal attributes and working patterns, for two weeks prior to test, for the entire hospital



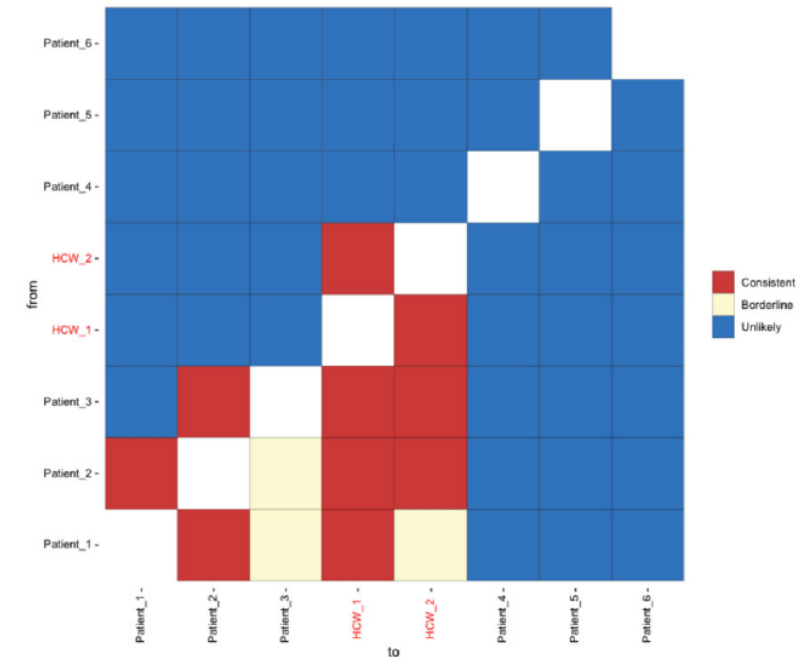
# Healthcare Worker Trajectories

During COVID-19



UNIVERSITY OF LEEDS

- The project emphasises the potential of **mobility trajectories** to highlight variations in behaviour within the hospital context
- Indications of **risk factors** affecting healthcare workers associated with working patterns, which is valuable for hospital policymaking to reduce further infection
- Current work relates to tracking **interactions** and **transmissions** between staff and patients using **mobility** and **RNA sequencing data**
- Digital footprints data can be ‘created’ from **repurposing** routine data!



# Acknowledgements

MRC-funded SAFER

Jared Wilson Aggarwal

Nick Gotts

Moira Spyer

Cat Houlihan

Eleni Nastouli

+ UCLH Staff



UNIVERSITY OF LEEDS



Thank You

Questions?

Ed Manley

[e.j.manley@leeds.ac.uk](mailto:e.j.manley@leeds.ac.uk)

[UrbanMovements.co.uk](http://UrbanMovements.co.uk)