Temporal phenome analysis of a large electronic health record cohort enables identification of hospital-acquired complications

Jeremy L Warner,1,2 Amin Zollanvari,3,4 Quan Ding,3 Peijin Zhang,5 Graham M Snyder,6 Gil Alterovitz3,7,8

ABSTRACT

Objective To develop methods for visual analysis of temporal phenotype data available through electronic health records (EHR).

Materials and methods 24,580 adults from the multiparameter intelligent monitoring in intensive care V.6 (MIMIC II) EHR database of critically ill patients were analyzed, with significant temporal associations visualized as a map of associations between hospital length of stay (LOS) and ICD-9-CM codes. An expanded phenotype, using ICD-9-CM, microbiology, and computerized physician order entry data, was defined for hospital-acquired Clostridium difficile (HA-CDI). LOS, estimated costs, 30-day post-discharge mortality, and antecedent medication provider order entry were evaluated for HA-CDI cases compared to randomly selected controls.

Results Temporal phenome analysis revealed 191 significant codes (p value, adjusted for false discovery rate, ≤0.05). HA-CDI was identified in 414 cases, and was associated with longer median LOS, 20 versus 9 days, and adjusted HR 0.33 (95% CI 0.28 to 0.39). This prolongation carries an estimated annual incremental cost increase of US$1.2–2.0 billion in the USA alone.

Discussion Comprehensive EHR data have made large-scale phenome-based analysis feasible. Time-dependent pathological disease states have dynamic phenomic evolution, which may be captured through visual analytical approaches. Although MIMIC II is a single institutional retrospective database, our approach should be portable to other EHR data sources, including prospective ‘learning healthcare systems’. For example, interventions to prevent HA-CDI could be dynamically evaluated using the same techniques.

Conclusions The new visual analytical method described in this paper led directly to the identification of numerous hospital-acquired conditions, which could be further explored through an expanded phenotype definition.

BACKGROUND AND SIGNIFICANCE

Clinical phenomics, the measurement of the diversity of disease states across human subjects, is of increasing importance in the analysis of large clinical datasets.1 Electronic health records (EHR) now make it possible to evaluate disease prevalence, distribution, and correlation across a predefined comprehensive collection of phenotypes (a ‘phenome’) to a degree that was previously not achievable.2–4 Due to the complexity of EHR and related medical data, visual analytics has become of increasing prominence.5,6 Earlier work has demonstrated that a clinical phenome, for example, as defined by the prevalence of International Classification of Disease, Ninth Revision, Clinical Modification (ICD-9-CM) codes or aggregations of such codes, can be used to calculate a phenome-wide association of disease codes with single-nucleotide polymorphisms (SNP).7–9 However, most work in this area has focused on dichotomous variables, such as SNP which are amenable to display by conventional probability plotting (eg, Manhattan and Q-Q plots). In previous work, we demonstrated that complex phenomic information can be visualized as a function of a continuous laboratory value, such as the white blood cell count.10 As disease conditions are expected to change over time, capturing the continuous temporal evolution of a phenome could lead to new and valuable discoveries. For example, hospital-acquired complications are well known to prolong the length of hospitalizations, but may do so to varying degrees. Information about the temporal associations of specific hospital-acquired conditions, optimized for data visualization, may be evaluable through comprehensive EHR resources. Previous work on temporal phenotype extraction from EHR has focused primarily on individual patient trajectories or specific case scenarios (eg, LifeLines11, LifeLines2,12 VISITORS),13 we propose to examine a large cohort for undiscovered temporal associations.

Objective In this paper, we describe a new method for: visual analysis of phenomic associations as a function of time; and hypothesis generation based on the resultant temporal phenome map. These two steps comprise a substantial portion of the ‘learning healthcare system,’ as shown in figure 1. For a clinical use case, we demonstrate how the application of these visual analytical methods can lead to recognition of specific hospital-acquired complications, as a function of the length of inpatient hospitalization, in a diverse EHR database of critically ill patients.

MATERIALS AND METHODS

Data source Multiparameter intelligent monitoring in intensive care V6 (MIMIC II), an EHR-based database of critically ill patients admitted to Beth Israel Deaconess Medical Center between 2001 and 2007, was the primary data source.14 MIMIC II contains detailed
Figure 1  Example schema of a ‘learning healthcare system’. This example demonstrates the ideal flow of a learning healthcare system environment, which begins with data analysis and visualization. Based on interpretation of these data, potential problems are recognized and hypotheses are generated. These lead to the development of interventions to mitigate or improve the identified problems, which are then implemented and evaluated in an iterative fashion. The first two steps of this process, represented by solid lines, are described within the paper.

Duration of healthcare exposure definition
We defined healthcare exposure duration as the interval from initial contact with the healthcare system to hospital discharge. In MIMIC II, hospital admissions and discharges are recorded as midnight-timed date stamps; intensive care unit (ICU) admission and discharge, POE orders, and laboratory tests are recorded with time stamps. We used the first time stamp for a laboratory test or POE obtained within ±48 h of the admission date stamp as a proxy for initial contact. For same-day hospital/ICU discharges, we used the ICU discharge time as the last contact. Otherwise, we used the last time stamp of a laboratory test or POE stop order occurring within 24 h of the hospital discharge date stamp as a proxy for last contact. If there were no stop orders or labs within 24 h of the discharge date stamp, the date stamp itself was used as the last contact time.

Phenome definition and temporal phenome mapping
The ICD-9-CM codes recorded for each hospitalization were used to define patient phenotypes. We aggregated individual hospitalizations into subgroups defined by healthcare exposure duration intervals. The intervals were defined such that approximately 200 patients would be included in each subgroup. All ICD-9-CM codes in each subgroup were tabulated, and 2×2 contingency tables were formulated comparing the number of instances of each ICD-9-CM code in the subgroup to those in the remainder of the population. Fisher’s exact test with two-sided hypothesis testing was used to determine significance for each contingency table. Adjusted p values of 0.05 or less, controlling for the false discovery rate using the method of Benjamini and Hochberg, were considered to report statistically significant associations of ICD-9-CM codes to subgroups. In order to study the robustness of our results to the number of patients in each subgroup, we explored how varying the healthcare exposure duration intervals affected the results.

A two-dimensional ‘temporal phenome map’ was rendered by plotting time as a function of ICD-9-CM codes, displaying line segments for significant adjusted p values, with segment length corresponding to the time interval containing the subgroup, and width proportionate to the negative logarithm of the adjusted p value. In the baseline analysis, an additional temporal phenome map with all adjusted p values less than 0.99 was rendered.

Visual analysis of the temporal phenome map
For the purposes of this pilot project, we limited the scope of extended evaluation to those ICD-9-CM codes in chapter one: Infectious and parasitic disease. These codes were systematically examined, and those that appeared to be consistent with hospital-acquired complications were selected for further study, with a final selection of one condition for further evaluation: hospital-acquired Clostridium difficile infection (HA-CDI).

Expanded case definitions
We then developed an expanded phenotypic definition for HA-CDI, using medication and microbiology information available in MIMIC II. Cases of HA-CDI were defined by one or more of the following occurring at least 48 h after initial contact: (1) a positive assay for C difficile toxin; (2) POE for oral or rectal vancomycin; (3) POE for oral or intravenous metronidazole and ICD-9-CM code 008.45: C difficile. For the second criterion, treatment of C difficile is the only common use for oral or rectal vancomycin so the ICD-9-CM code was not required; conversely, oral or intravenous metronidazole is used to treat other conditions, so the ICD-9-CM code was required for the third criterion. Non-HA-CDI was defined using the same criteria but with cutoffs before 48 h for case definition.

Matching controls to cases
For the set of HA-CDI cases, the outlying 1st percentile and 99th percentile of healthcare exposure duration were excluded before selection of a matching control group. Candidate controls were randomly selected from the remaining MIMIC II cohort, excluding cases of non-HA-CDI. Candidates were excluded if their healthcare exposure duration was less than the 1st percentile or greater than the 99th percentile of the case hospitalizations. If candidates had at least one laboratory value measurement during the first 48 h of hospitalization, they were included as a control. This criterion was set to exclude any test patients present in MIMIC II who appear identical to real patients but do not have laboratory information recorded. Candidate evaluation continued until a 1:1 match was achieved. Demographics (age, gender, ethnicity, and Elixhauser comorbidity scores) were recorded for all cases and controls; Elixhauser comorbidity is pre-calculated for the MIMIC II cohort. In order to explore patterns of antecedent medication use, which can be associated with propensity to HA-CDI, medication POE data were used to develop three aggregate groupings: (1) antibacterial agents not known to be associated with C difficile (low-risk antibacterial agents); (2) antibacterial agents known to be associated with C difficile (high-risk antibacterial agents); and (3) proton pump inhibitors and H2 receptor antagonists (H2-blocker).

Statistical and general methods
Multiple hospitalizations of the same patient were treated as independent events, and the adjusted p values for each subgroup...
were freely available at http://www.dryang.org/temporalphenome/ for generating and zooming into the temporal phenome maps, http://cran.r-project.org/). The R scripts, which include routines were conducted with R (The R Project for Statistical Computing postgresql.org). Queries to the database and the visual analysis (The PostgreSQL Global Development Group, http://www. system.19 prospective analysis of the hospital cost report information

infection due to C difficile (ICD-9-CM 008.45, temporal range 28.7–294.5 days, OR range 3.15–5.16); Pseudomonas aeruginosa infection (ICD-9-CM 041.7, temporal range 42.4–49.6 days and 61.5–294.5 days, OR 10.6); and aspergillosis (ICD-9-CM 117.3, temporal range 61.5–294.5 days, OR 35.4).

In a sensitivity analysis, the calculations were repeated using subgroup sizes of 25, 50, 100, and 400 patients; these results are shown in figure 3. The smaller subgroups (25, 50 patients) had very few significant findings—40 (0.7%) and 73 (1.3%) of the possible ICD-9-CM codes, respectively. Conversely, the larger subgroup (400 patients) had more findings than the baseline case (285, 5%), but the graphic was distorted by the large time intervals.

HA-CDI associated with longer hospital stays, increased mortality, and greater antecedent medication POE

In the MIMIC II database, we identified 414 cases (1.5% of hospitalizations) that met our expanded phenotypic definition of HA-CDI. For HA-CDI cases, first identification was as follows: 56% by a positive C difficile toxin, 41% by POE for metronidazole plus C difficile ICD-9 code, and 3% by POE for oral or rectal vancomycin.

For the matching analysis, 362 HA-CDI cases (87%) met inclusion criteria, after exclusion for outlying length of stay (LOS) and/or lack of POE information. Randomly selected candidate controls met the eligibility criteria 63.5% of the time. Compared to cases, controls were well matched for age, gender, ethnicity, and Elixhauser comorbidity index (p value non-significant for all comparisons, table 1). Length of hospitalization was significantly longer for cases, median 20 days versus 9 days (HA-CDI vs controls—table 2 and figure 4); this 11-day prolongation in length of hospitalization would be expected to add US$13 500–27 500 (2004 dollars) to the cost of care for a patient with HA-CDI, dependent on how much time was spent in the ICU versus the standard hospital floor. Cases were also more likely to die within 30 days of hospital discharge: 22% vs 12% (HA-CDI vs controls), although this comparison was no longer significant after adjustments for demographic factors—table 2 and figure 4.

The aggregated medication categories, as described above, are detailed in supplementary Table S2 (available online only). HA-CDI cases were almost twice as likely to have antecedent POE for high-risk antibacterial agents up to 24 h before HA-CDI case definition, compared to control POE during the first 48 h of hospitalization (66% vs 34%, table 2). They were

Table 1  Baseline demographics of the MIMIC II V.6 dataset

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Total adult hospitalizations (n=28 061)</th>
<th>HA-CDI cases (n=362)</th>
<th>HA-CDI controls (n=362)</th>
<th>p Value, HA-CDI cases compared to controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, median (IQR), year</td>
<td>65 (51–77)</td>
<td>68 (22–99)</td>
<td>65 (21–95)</td>
<td>0.079</td>
</tr>
<tr>
<td>Men, no. (%)</td>
<td>15 781 (56)</td>
<td>199 (55)</td>
<td>201 (56)</td>
<td>0.940</td>
</tr>
<tr>
<td>Race/ethnicity, no. (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>586 (2)</td>
<td>5 (1)</td>
<td>13 (4)</td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>2362 (8)</td>
<td>27 (7)</td>
<td>27 (7)</td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>825 (3)</td>
<td>7 (2)</td>
<td>10 (3)</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>19 704 (70)</td>
<td>282 (78)</td>
<td>257 (71)</td>
<td></td>
</tr>
<tr>
<td>Other/unknown</td>
<td>4594 (16)</td>
<td>41 (11)</td>
<td>55 (15)</td>
<td></td>
</tr>
<tr>
<td>Elixhauser comorbidity index, median (IQR)</td>
<td>2 (1–4)</td>
<td>2 (1–3)</td>
<td>2 (1–3)</td>
<td>0.480</td>
</tr>
</tbody>
</table>

1Total adult hospitalizations include all patients 16 years of age and older at the time of hospital admission. HA-CDI cases and randomly selected controls are well-matched by age, gender, ethnicity, and comorbidity. HA-CDI, hospital-acquired C difficile; MIMIC II, multiparameter intelligent monitoring in intensive care V.6.

also more likely to have POE for low-risk antibacterial agents and for H2-blockers or proton pump inhibitors.

**DISCUSSION**

We have demonstrated a new method for the quantification and visualization of temporal patterns of risk of disease, as defined by ICD-9-CM codes available from EHR data. Given the high dimensionality of the EHR data under consideration, the human visual system is well suited to high-level interpretation of the data. Although tabular views may provide the same information, the pattern recognition abilities of the visual system are not engaged under this circumstance. We were immediately able to recognize temporal patterns, as well as regions of interest, such as ICD-9-CM chapter one, which were selected for more

---

**Figure 2** Temporal phenome maps. Temporal phenome-wide association of International Classification of Disease, Ninth Revision, Clinical Modification (ICD-9-CM) codes: (A) Only associations with adjusted p values of 0.05 or less are shown. (B) All associations with adjusted p value less than 0.99 are shown. For both visualizations, the width of the rendered line segments is proportionate to the negative logarithm of the adjusted p value. Each chapter of the ICD-9-CM coding schema is shown in a separate color, with vertical bars separating the chapters, and V and E codes on the right. Times are shown in days, using deciles, ranging from the 0th to 100th percentile. Median duration of hospitalization and IQR for the entire multiparameter intelligent monitoring in intensive care V.6 (MIMIC II) adult population are shown as horizontal lines.

**Figure 3** Temporal phenome map sensitivity analysis. Calculations were repeated using a subgroup size of (A) 25, (B) 50, (C) 100, and (D) 400 patients. Only associations with adjusted p values of 0.05 or less are shown. ICD-9-CM, International Classification of Disease, Ninth Revision, Clinical Modification.
HA-CDI cohort of increased LOS and a trend towards increased mortality in the eligible pool will affect control characteristics (critically ill adult and neonatal patients, in the case of MIMIC II). Date of death information is not available through MIMIC II data; however, cause of death information is available through MIMIC II; therefore, the trend towards increased mortality observed in the HA-CDI group may have been directly from HA-CDI or from unrelated causes.

Preventable hospital-acquired complications with modifiable risk factors, such as HA-CDI, are a major cause of morbidity and mortality, and add significantly to hospitalization costs. \(^{20–22}\) This increased cost is supported by our study, which suggests average incremental costs of US$13,500–27,500 for an individual diagnosed with HA-CDI, as defined by our phenotyping algorithm. With extrapolation to the total USA population of critically ill patients, the estimated average annual incremental cost is US$1.2–2.0 billion for cases of HA-CDI, or approximately 1.4–2.3% of USA critical care expenditures. \(^{19}\)

While the association of HA-CDI with increased LOS and increased costs is not novel, the finding increases confidence in our approach. Less well-studied conditions, such as multidrug-resistant Gram negative bacteria (particularly carbapenem-resistant Enterobacteriaceae), may also be identifiable through the described method. A brief review of other ICD-9-CM chapters revealed additional possible hospital-acquired conditions, for example: iatrogenic pulmonary embolism and infarction (ICD-9-CM 415.11); iatrogenic pneumothorax (ICD-9-CM 512.1); decubitus ulcer (ICD-9-CM 707.0); and infection due to other vascular device, implant, and graft (ICD-9-CM 996.62). Identifying hospital-acquired complication phenotypes through large-scale analysis of EHR data, as we have demonstrated in this pilot study, could be a precursor to targeted proactive measures, such as clinical decision support, which could yield significant cost savings while also reducing morbidity and mortality rates. In addition, this tool can be implemented to identify characteristics of conditions at the local level, and then change factors at the local level. For example, if \(C\) difficile infection is not associated with preceding antibiotic use in the reviewer’s hospital, then antimicrobial stewardship might not be the first intervention made to improve infection rates. Such evidence-driven clinical decision support, based on global as well as local factors, is essential to successful learning healthcare systems.

There are several notable limitations to this study. Primarily, tests of association as used in the temporal phenotype analysis cannot distinguish between cause and effect. Currently, therefore, it is necessary to apply external medical domain knowledge to determine whether an identified code is or is not a potential hospital-acquired complication. Second, we chose to use ICD-9-CM codes due to their ready availability, despite the caveat that they are known to be recorded with variable accuracy. \(^{23–25}\) Furthermore, ICD-9-CM codes are generated at the end of a (potentially lengthy) hospitalization and are thus

---

**Table 2** Exposures and outcomes of HA-CDI cases as compared to matched controls

<table>
<thead>
<tr>
<th>Exposure or outcome</th>
<th>HA-CDI cases (n=362)</th>
<th>HA-CDI controls (n=362)</th>
<th>p Value</th>
<th>HR or OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length of stay, median (IQR), days</td>
<td>20 (13–33)</td>
<td>9 (6–14)</td>
<td>&lt;0.001</td>
<td>0.34 (0.30 to 0.41)*</td>
</tr>
<tr>
<td>30-Day post-discharge mortality, no. (%)</td>
<td>80 (22)</td>
<td>42 (12)</td>
<td>0.05</td>
<td>1.48 (1.00 to 2.17)*</td>
</tr>
<tr>
<td>High-risk antibacterial exposure, no. (%)§‡</td>
<td>240 (66)</td>
<td>124 (34)</td>
<td>&lt;0.001</td>
<td>3.77 (2.74 to 5.20¶)</td>
</tr>
<tr>
<td>Low-risk antibacterial exposure, no. (%)§‡</td>
<td>249 (69)</td>
<td>168 (46)</td>
<td>&lt;0.001</td>
<td>2.54 (1.86 to 3.49¶)</td>
</tr>
<tr>
<td>PPI or H2-blocker exposure, no. (%)§‡</td>
<td>297 (82)</td>
<td>265 (73)</td>
<td>0.006</td>
<td>1.67 (1.16 to 2.43¶)</td>
</tr>
</tbody>
</table>

*Unadjusted HR.
†Adjusted HR.
‡Medication POE data collected up to 24 h before diagnosis for cases, and for the first 48 h of healthcare exposure, for controls.
§Aggregate medication categories are defined in supplementary Table S2 (available online only).
¶OR.

---

**Figure 4** Outcomes for hospital-acquired \(C\) difficile (HA-CDI) cases compared to randomly selected controls. (A) Hospitalization duration is significantly longer for HA-CDI cases; (B) 30-day post-discharge mortality is slightly worse for HA-CDI cases compared to controls. HR, 95% CI, and p values are shown within figure panels.
subject to the recency effect, suggesting that important diagnostic events occurring early in a prolonged admission may not be captured. It is likely that the accuracy of the phenotype maps would be improved through expanded phenotype definitions based on structured clinical information (eg, encoded problem lists/summaries) as well as narrative clinical elements.26–29 While definitions such as those provided by the phenotype knowledge base (http://www.phenkb.org/) hold promise as more specific and sensitive phenotype definitions, they are not yet extensive enough to be practicable across the entire human phenome.24,30 Replication of this study in less well-structured environments than MIMIC II would be difficult; fortunately, the use of robust clinical data warehouses as well as secondary solutions such as Informatics for Integrating Biology and the Bedside31 is becoming more widespread.

We chose to aggregate by fixed patient numbers due to the marked non-linearity of time epochs. In addition, it was unclear a priori what time intervals could or should be considered clinically significant. Future enhancements to this tool will include the ability to toggle between patient and temporal groupings. Our sensitivity analysis demonstrates that the choice of subgroup size can have a strong influence on the results. Clearly, a fixed subgroup size of 25 was too underpowered to generate meaningful results. In addition to the choice of interval size for fixed subgroups, several other questions and complications arise: (1) Should subgroup size be non-linear, for example, dependent on the degree to which a time value is ‘out of range’? It is clear that patients with extreme out of range values are markedly different both from the general population and perhaps from each other, which is why they are conventionally excluded from analyses, including our case-control analysis. (2) Should subgroups have ‘soft’ boundaries such that patients are potentially included in more than one subgroup? (3) Computational complexity increases substantially with smaller subgroups. Future work will also focus on developing new methods to define optimally sized subgroups.

Our time definitions may lead to several systemic errors, which are likely to be widely present in EHR databases. For example, if a patient was transferred from another institution, the exposure at the previous institution would not be captured by MIMIC II. Conversely, the day of discharge should be fairly accurate but the hour of discharge is subject to unmeasured variability, leading to systemic underestimation of hospitalization time. Case definition was timed using the surrogate of a treatment decision, recorded as POE. It is assumed that the time lapse between a treatment decision and the entering of POE for that treatment is minimal, although this may not always be the case. Ultimately, we chose to use POE information because it is readily available in structured format, whereas the diagnostic information that prompts treatment decisions is typically in narrative format and difficult to access.32 As such, patients who were diagnosed with HA-CDI by means other than a positive toxin but not treated (for whatever reason) would not be captured by our phenotype definition. Whether this presumably small unmeasured cohort may have influenced our findings is unknown; future work will focus on whether such cohorts can be identified by other means.

CONCLUSION
There is a general need for ‘learning healthcare systems’, which will enable effective translation of EHR-driven discoveries into clinical interventions at the institutional, regional, and national levels.13,34 This work presents a new methodology for visual analytics and testable hypothesis generation from EHR data, which reveals patterns of context-specific complications with clinical implications. This analysis and interpretation begins to create a prototypical learning healthcare system, as shown in figure 1. In order to ‘complete the loop’ these findings would have to be implemented as an intervention, the effects of which could then be measured and re-evaluated. This process could continue until an identified problem has been resolved, and could also be used to provide objective evidence of quality improvement efforts.

Author affiliations
1Department of Medicine, Division of Hematology and Oncology, Vanderbilt University, Nashville, Tennessee, USA
2Department of Biomedical Informatics, Vanderbilt University, Nashville, Tennessee, USA
3Center for Biomedical Informatics, Harvard Medical School, Boston, Massachusetts, USA
4Department of Statistics, Texas A&M University, College Station, Texas, USA
5Program for Research in Mathematics, Engineering and Science for High School Students, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA
6Department of Health Care Quality, Division of Infection Control and Hospital Epidemiology, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, Massachusetts, USA
7Children’s Hospital Informatics Program at Harvard-MIT Division of Health Science, Boston, Massachusetts, USA
8Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

Acknowledgements The authors would like to extend special thanks to Peter Szolovits, Leo Celi, Federico Cismondi, Daniel Scott, Tom Lasko, and Josh Denny, for their support and advice in regards to MIMIC II and general considerations. The authors would also like to thank Peter Yang for hosting files.

Contributors JLW had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. JLW and GA conceived the study design. JLW and PZ performed the experiments. JLW, AZ, PZ, and GMS analyzed the data. JLW, QD, AZ, PZ, and GMS contributed to the manuscript writing; all authors approved the final manuscript.

Funding This work was supported in part by grants SR21DA025168-02 (GA), 1R01HG004836-01 (GA), and 4R01LM009826-03 (GA). The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests None.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement The R code developed for the study is freely available through a website described in the paper.

REFERENCES