The quality of vital signs measurements in electronic medical records varies by hospital, specialty, and patient demographics

Niall Jackson¹, Jessica Woods², Peter Watkinson³, Andrew Brent¹, Tim EA Peto^{4,5}, A Sarah Walker^{4,5}, David W Eyre^{5,6}

¹Oxford University Hospitals NHS Foundation Trust, Oxford, UK

² N Family Club, London, UK

³ Nuffield Department of Clinical Neurosciences, University of Oxford, UK

⁴ Nuffield Department of Medicine, University of Oxford, Oxford, UK

⁵ NIHR Health Protection Research Unit in in Healthcare Associated Infections and Antimicrobial Resistance, University of Oxford, Oxford, UK

⁶ Big Data Institute, Nuffield Department of Population Health, University of Oxford, Oxford, UK

Corresponding author: David Eyre, Big Data Institute, University of Oxford, Old Road Campus, Oxford, OX3 7LF, UK. david.eyre@bdi.ox.ac.uk

Key words: Electronic healthcare records; Vital signs; Digit preference

Abstract

Objective: To assess the frequency of digit preference in recording of vital signs in electronic healthcare records (EHRs) and associated patient and hospital factors.

Study Design and Setting: We used EHR data from Oxford University Hospitals, UK, between 01-January-2016 and 30-June-2019 and multivariable logistic regression to investigate associations between temperature readings of 36.0°C or systolic and diastolic blood pressure (SBP/DBP) readings both ending in zero and patient age, sex, ethnicity, deprivation, comorbidities, calendar time, hour of day, days into admission, hospital, day of week and speciality.

Results: In 4,305,914 records from 143,352 patients, there was an excess of temperature readings of 36.0°C (15.1%, 649,976/4,305,914), compared to an expected 4.9% from the underlying distribution. 2.2% (95,215) BP readings had a SBP and DBP both ending in zero vs. 1% expected by chance. Digit preference was more common in older and male patients, as length of stay increased, following a previous normal set of vital signs and typically more common in medical vs. surgical specialities. Differences were seen between hospitals, however, digit preference reduced over calendar time.

Conclusion: Vital signs may not always be accurately documented. Allowances and adjustments may be needed in observational analyses using these factors as outcomes or exposures.

What's New?

Key findings

- Digit preference in the recording of vital signs in electronic healthcare records is common, affecting approximately 10% of temperature measurements and 1% of blood pressure recordings in a large UK teaching hospital group
- These findings were obtained in hospitals using a semi-automated data capture system that required manual re-entry of vital signs into a tablet computer prior to automated upload to electronic patient records
- Digit preference was associated with patient characteristics and was more common in older and male patients, as length of stay increased and following previously normal vital signs
- Digit preference varied between hospitals, but decreased over time
- Digit preference was generally more common in medical compared to surgical specialities

What this adds to what is known

- Most previous studies of data quality in electronic patient records have focused on the accuracy of coding
- This study focuses on the accuracy of numeric values in patient records, and also adds new data on patient and hospital factors associated with the accuracy of values in electronic patient records

Implications

- Clinicians and researchers need to be aware that vital signs may not always be accurately documented
- Appropriate allowances and adjustments for digit preference should be considered in observational analyses using these factors as outcomes or exposures.
- Further work is required to understand the mechanisms behind values preference on a systems, patient and clinician level

Introduction

Electronic healthcare records (EHR) have been widely adopted across different healthcare settings globally and have become an integral part of health infrastructure: saving time, improving communication and record keeping, and supporting learning.¹ As the scale and breadth of EHR data increases, so does its ability to fulfil secondary functions including quality improvement, product development, and research, contingent on appropriate regulation and transparency.² Example applications of EHR data include population-level epidemiological studies,^{3–5} machine learning-based diagnostic assistants for clinicians,⁶ screening for child maltreatment and family violence,⁷ and detecting and tracking infectious disease outbreaks.^{8,9}

However, conclusions rely on the reliability and accuracy of EHR data, which is not guaranteed.^{10,11} Indeed, the use of EHR data beyond its original purpose (clinical care and billing) raises specific challenges. Data collection in the clinical environment is imperfect¹² and often incomplete;¹³ it may lack comparability or reproducibility¹⁴ or even simply be wrong.^{15,16} Attempts to quantify or evaluate EHR data quality are limited, and even fewer have investigated causes for variability in quality.¹⁷ Most studies have focused on checking the accuracy of clinical and diagnostic codes rather than numerical observations. In one notable exception, evaluating the quality of vital sign data across multiple hospitals and EHR systems, there was a skew of completeness and correctness in favour of arriving patients and higher fidelity in wholly EHR based systems compared to a combination of paper and EHR.¹⁸

Digit preference in vital sign measurement is well recognized in healthcare. Terminal digit preference for multiples of ten in blood pressure recordings has been shown to be extremely common,^{19,20} introduce systemic bias potentially effecting mortality²¹ and produce inaccurate epidemiological results.²² This phenomenon has been observed in other vital sign measurements such as respiratory rate, with attempts to rectify inaccuracy through continuous, automated monitoring.²³

In this study we investigate observations of vital signs gathered over 3.5 years from inpatients at a large UK teaching hospital group. We assess the frequency of recordings with preferences for a specific value (e.g., multiples of ten for blood pressure or temperature readings of 36.0°C) as a marker of sub-optimal data quality and establish associated patient and hospital factors.

Methods

Setting

We conducted a retrospective observational study at Oxford University Hospitals NHS Foundation Trust (OUH). OUH consists of four teaching hospitals with a total of 1000 beds: Hospital A (providing acute care, trauma and neurosurgery services); Hospital B (elective cancer surgery, transplant, haematology, oncology); Hospital C (district hospital, acute medical services) and Hospital D (elective orthopaedics). OUH acts as a tertiary referral centre for the surrounding region, generating approximately 1 million patient contacts a year, serving a population of around 655,000.

Data

Deidentified data were obtained from the Infections in Oxfordshire Research Database which has approvals from the National Research Ethics Service South Central – Oxford C Research Ethics Committee (19/SC/0403), the Health Research Authority and the national Confidentiality Advisory Group (19/CAG/0144). We used individual observations of vital signs conducted in OUH for adult patients (≥18y) between 01-January-2016 and 30-June-2019. The vital signs observed, with dates and times of collection, included respiratory rate (RR), heart rate (HR), tympanic temperature, systolic and diastolic blood pressure (SBP and DBP).

Observations were collected using a semi-automated vital sign observation system. HR, SBP and DBP were collected using an observation machine, RR manually timed, and temperature with a separate tympanic thermometer. All observations were then manually transcribed into a tablet computer attached to the same stand, which uploaded results into the EHR. Oxygen saturations, supplemental oxygen and alertness (alert, responsive to voice, pain or unresponsive, AVPU) are also all recorded. However, these measurements are not considered further here as the dynamic range of oxygen saturation measurements or number of categories for AVPU scoring were insufficient to distinguish digit preference from the underlying distribution of the data.

Additional data were obtained: hospital-level data (hospital and ward where the observation was made, the specialty managing the patient); and patient data (age, sex, ethnic group, index of multiple deprivation (IMD) score at home address, Charlson comorbidity score).

Statistical analysis

We defined preferences for a specific value or digit for each vital sign based on the observed distribution of values. Temperatures of 36.0°C, blood pressure (BP) with both SBP and DBP recorded as multiples of ten, RRs of 16 or 18, and HRs as multiples of ten were all more common than would be expected by chance.

Temperature had the clearest value preference signal; associated factors were investigated using multivariable logistic regression. Analyses were restricted to patients with complete data, and to complete vital sign sets (i.e. all of temperature, HR, RR, SBP, DBP recorded). We used natural cubic splines to account for non-linear relationships for continuous variables (allowing up to five default placed knots, selecting the final number of knots by minimising the Bayesian Information Criterion, BIC). To avoid undue influence of outlying values, continuous variables were truncated at the 1st and 99th percentiles. Pairwise interactions between model main effects were included where this improved model fit based on BIC. We used clustered robust standard errors to account for repeated measurements obtained from the same patient.

To investigate if associations were specific to temperature or applied to vital signs more widely, we also refitted the same model (i.e., with the same spline terms and interactions) with BP digit preference as the outcome. For RR the most common value preferences were also the most common normal values precluding separate investigation of value preference. Value preferences were less frequent for HR and not investigated further.

We also investigated if the presence of abnormal previous temperature or BP readings affected subsequent digit preference. Temperatures of $\leq 35.5^\circ\text{C}$ or $\geq 37.5^\circ\text{C}$ and SBP readings of >160 or <90 mmHg or DBP readings of >100 or <60 mmHg were arbitrarily considered abnormal. For each observation with a prior observation from the same patient within ≤ 36 hours, we selected the most recent prior observation for comparison. We then refitted the regression models above including if the prior temperature or BP reading had been abnormal as a covariate.

Analyses were conducted using R, version 4.1.

Results

Between 01-January-2016 and 30-June-2019, a total of 5,799,224 sets of vital signs were recorded. For 482,257 (8.3%) records no deprivation score was documented, and these records were excluded. Complete data were available for all other hospital/patient variables. Within the 5,316,967 vital signs sets included, 758,955 (14.3%) did not include temperature, 473,433 (8.9%) were missing SBP and/or DBP, 535,534 (10.1%) missing RR, and 450,899 (8.5%) missing HR recordings. Restricting to complete sets of vital signs left 4,305,914 (81.0%) records in the final dataset from 143,352 patients. The median (IQR) patient age was 60 (42-75) years, 74,726 (52.1%) were female, and 105,751

(73.8%) were of white and 30,272 (21.1%) of unstated or unknown ethnicity. The most common specialties recording vital signs were general surgery (861,807, 20.0%), trauma and orthopaedics (640,259, 14.9%) and acute and emergency medicine (619,636, 14.4%).

Prevalence of value preferences in vital signs readings

Compared with the overall distribution of temperature values, there was an excess of temperature readings of 36.0°C (Figure 1), which accounted for 15.1% (649,976/4,305,914) of values. This compared to an expected 4.9% (Figure S1) from the underlying distribution, i.e., approximately 10% of all temperature recordings were likely inappropriately recorded as 36.0°C. Approximately 1% of all BP readings would be expected to have both a SBP and DBP ending in zero, however 2.2% (95,215) of readings showed this pattern. Digit preference was less common for HR readings, with 12.1% (522,101) ending in zero compared to an expected 10%. Although 51.4% (2,213,322) of RR readings were 16 or 18 reflecting a likely excess of these values, RR was not studied further as it was difficult to distinguish which associations reflected normal physiology and which digit preferences.

Value preference associations with patient demographics

Temperature was independently more likely to be recorded as 36.0°C with increasing age above 50 years and BP most likely to be recorded with SBP and DBP both ending in zero for those above 80 years (Figure 2A). Male patients were more likely than female patients to have readings with value preferences for both temperature (adjusted odds ratio, aOR=1.10 [95%CI 1.09-1.12]) and BP (1.04 [1.00-1.08]) (Table 1). Temperature value preference was slightly less common in patients from less deprived areas (aOR per 10 unit change in deprivation percentile=0.99 [0.99-0.99, higher percentiles are less deprived]), but with no evidence of a difference in BP value preference with variation in deprivation. There was no evidence for consistent differences in value preference by ethnicity, however temperatures of 36.0°C were more commonly recorded in patients of Asian ethnicity (aOR vs. white=1.07 [1.02-1.12]) and BPs ending in zero were more common in those of unstated or unknown ethnicity (aOR vs. white=1.12 [1.07-1.18]). Patients with higher Charlson scores were more likely to have recorded temperatures of 36.0°C (aOR per 5 unit increase=1.02 [1.02-1.02]), but less likely to have BP ending in zero (aOR per 5 unit increase=0.93 [0.92-0.94]).

Changes over calendar time, hour of day, and by hospital and time in admission

The frequency of temperature and BP measurements with value preferences decreased during the study (Figure 2B), e.g., for temperature from an estimated 17.7% at the start of 2016 to 12.5% at the start of 2019. Temperatures were most likely to be recorded as 36.0°C at around 6am, i.e., at the first routine set of observations performed per day in most patients, whereas BP was most likely to end in zero during the evening (Figure 2C).

Differences were also seen between hospitals in temperature value preferences; at the time of admission this was more common in hospitals A-C, and considerably more common in long stay patients at hospital C. There were small changes in temperature value preferences by day of the week (Table 2) and digit preference in BP became less common during the working week (aOR for Friday vs. Monday=0.87 [0.85-0.90]). As number of days into a hospital admission increased, temperatures were more likely to be recorded as 36.0°C in hospitals A-C with more marked increases at hospital C, but less likely in hospital B (Figure 2D). For BP, digit preference was relatively stable by day of admission and between hospitals, except at hospital B where it increased during hospital stay.

Variation by specialty

Temperature readings of 36.0°C and BP readings ending in zero were generally more common in medical vs. surgical specialties (Figure 3). Exceptions included nephrology where value preference was less common for both temperature and BP, haematology and oncology where temperatures of 36.0°C were less commonly recorded and cardiology where BP digit preference was less common. Additionally, temperature readings of 36.0°C were more common in specialist surgery (ear, nose and

throat surgery; plastic surgery; maxillofacial surgery; ophthalmology) and BP digit preference in haematology and oncology.

Effect of previous abnormal measurements

A total of 4,037,232 sets of observations had a previous measurement from the same patient within ≤ 36 hours. Temperature readings of 36.0°C were less frequent following an abnormal prior temperature measurement, 20,272/351,619 (5.8%), than following a normal prior measurement, 587,856/3,685,613 (16.0%). Given it may take time for temperature to normalise it would not be expected that those with a previously abnormal temperature would have the same proportion of true temperatures of 36.0°C as the overall hospital population (estimated above as 4.9%). However even allowing for this, preference for recording a temperature of 36.0°C was more common with a normal prior measurement. In an analysis, adjusted for the same factors as in Table 1, a previous abnormal temperature reduced the odds of a recording of 36.0°C (aOR=0.35, [95%CI 0.34-0.35]). Similarly, 24,394/1,291,295 (1.9%) of BP observations had SBP and DBP both ending in zero after an abnormal SBP or DBP, compared to 66,253/2,745,937 (2.4%) without a prior abnormal reading (aOR=0.87 [0.85-0.89]).

Discussion

In this analysis of records from a large UK teaching hospital group, we show preference for specific values or digits in vital sign records in EHRs. Our findings have implications for patient management, quality improvement initiatives and for research conducted using EHRs.

A key question is whether value preferences indicate simply convenience rounding on transcribing values or whether they are also a marker for incompletely observed observations. Potentially favouring the latter, we observed differences in the relative frequency of value preferences, with temperature recordings of 36.0°C three times more common than expected based on the remainder of the temperature distribution. It is also possible that given 36.0°C was the lower bound of the normal range in the hospitals' early warning score, that implausibly low readings, e.g., when thermometers were mis-calibrated were recorded as 36.0°C . In contrast, BP readings ending in zero were twice as common, but HR readings ending in zero only around 20% above expected. This may reflect the additional burden of recording a patient's temperature, requiring a second piece of equipment that may occasionally be missing or broken. Conversely it is possible that a greater proportion of all temperature values are rounded to a single value.

We found differences between hospital specialties, even after adjustment for other factors. Generally surgical specialties recorded vital signs with greater precision. However, the prevalence of value preferences also potentially reflects the culture within a speciality, where greater importance may be placed on measured values, e.g. in nephrology, or on specific vital signs, e.g. temperature in neutropenic and other immunosuppressed patients in haematology and oncology or BP in cardiology. We also found marked differences in temperature measurement between the four hospitals in the organisation, even following adjustment for the specialties present. This may reflect systemic factors, e.g., staffing levels and the importance placed on vital signs may vary by setting. In higher acuity settings, reliance on vital signs for treatment escalation could increase vital sign fidelity compared to less acute settings focused on rehabilitation. Although patients are admitted to acute medicine as an emergency, the increased digit preference seen in this specialty may reflect that for many longer staying patients rehabilitation and provision of social care are the dominant issues for much of each admission. In keeping with this, for both acute hospitals, A and C, we found that temperatures were more likely to be recorded as 36.0°C as length of stay increased. We also found that normal prior measurements were more likely to be followed by digital preference in subsequent observations, with previous abnormal measurements being associated with greater accuracy in subsequent observations.

Older patients and male patients were more likely to have vital sign recordings with value preferences. Further work is required to better understand the reasons for this. The reasons for variation in temperature recording by age may reflect differences in the acuity of patients and associated culture around vital sign measurement, the relative importance placed on curative treatment vs. patient comfort, and physical barriers to temperature measurement including patient agitation. There were no systematic differences by ethnicity across both vital sign modalities.

Changes over time suggest institution-wide improvement is possible, with increased precision of both temperature and BP seen during the study. The study builds on previous studies of vital sign recording quality,²⁴ and highlights that institutions may wish to monitor vital sign recording to identify areas of the hospital or patient groups where specific interventions to improve quality may be required.

Multiple variables representing the timing of measurements were investigated. Temperature appeared to be most likely to be recorded accurately in the afternoon or evening. Conversely BP measurements appeared most likely to be rounded in the evening. Generally, vital sign precision was greatest around the time of hospital admission, likely reflecting that patients are most unwell when first presenting to hospital and so vital signs are performed and recorded carefully and in full.

Digit preference is a well described phenomenon.^{19,20} However, particularly for temperature measurement, the question that arises from our findings is; if a vital sign is more difficult to measure for some reason then why does current culture potentially favour documenting an inaccurate reading instead of leaving it missing. There may be explicit or implied pressure to always record a complete set of vital signs (although ~20% of observations in our study were excluded because of missing one or more vital signs), or it may be that recording an observation as unavailable may be more onerous and require entering a justification. There may also be disincentives to recording abnormal values if this requires escalation of care and additional action, but this would be expected to potentially impact all values rather than just those showing value preference.

Limitations of our study include that it is based on a single organisation and data entry system for recording vital signs. Further studies are required to confirm if our findings are replicated more widely. We did not investigate more granular variation in vital sign recording by hospital ward or individual staff member, the latter as the identity of the healthcare worker recording the vital signs was not available in our data extract. We also did not investigate the downstream consequences of vital sign value preferences, and this could be looked at in future work, e.g. considering associations with length of stay or mortality, although care would be required to avoid reverse causation where delays in discharge or a more palliative focus change value preferences.

Our study provides evidence that vital sign measurement displays value preference to a such a degree that it could affect conclusions based on unadjusted vital sign data, in both clinical and research settings. We show that hospital, speciality, admission stage and patient age all have important impacts on the accuracy of vital signs. Changes over time in our hospital suggest improvements in accuracy are possible. Ultimately fully connected systems that do not require re-entry of vital signs into the EHR are likely to address many of the issues identified. Work with institutions and individuals is required to fully elucidate and understand the mechanisms behind values preference on a systems, patient and clinician level. In the meantime, clinicians and researchers need to be aware that vital signs may not always be accurately documented, and to make appropriate allowances and adjustments for this in observational analyses using these factors as outcomes or exposures.

Acknowledgements

This work uses data provided by patients and collected by the UK's National Health Service as part of their care and support. We thank all the people of Oxfordshire who contribute to the Infections in Oxfordshire Research Database. Research Database Team: L Butcher, H Boseley, C Crichton, DW Crook, DW Eyre, O Freeman, J Gearing (community), R Harrington, K Jeffery, M Landray, A Pal, TEA Peto, TP Quan, J Robinson (community), J Sellors, B Shine, AS Walker, D Waller. Patient and Public Panel: G Blower, C Mancey, P McLoughlin, B Nichols.

Funding

This work was supported by the National Institute for Health Research Health Protection Research Unit (NIHR HPRU) in Healthcare Associated Infections and Antimicrobial Resistance at Oxford University in partnership with the UK Health Security Agency (NIHR200915), and the NIHR Biomedical Research Centre, Oxford. DWE is a Big Data Institute Robertson Fellow. ASW is an NIHR Senior Investigator. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, the Department of Health or the UK Health Security Agency. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Transparency declarations

DWE declares lecture fees from Gilead outside the submitted work. No other author has a conflict of interest to declare.

Data availability

The data analysed are not publicly available as they contain personal data but are available from the Infections in Oxfordshire Research Database (<https://oxfordbrc.nihr.ac.uk/research-themes-overview/antimicrobial-resistance-and-modernising-microbiology/infections-in-oxfordshire-research-database-iord/>), subject to an application and research proposal meeting on the ethical and governance requirements of the Database.

Table

	Temperature, descriptive		BP, descriptive		Temperature, multivariable			BP, multivariable		
Characteristic	Not 36.0°C, N = 3,655,938 ¹	36.0°C, N = 649,976 ¹	Does not end in 0, N = 4,210,699 ¹	Ends in 0, N = 95,215 ¹	OR ²	95% CI ²	p-value	OR ²	95% CI ²	p-value
Age	70 (55, 81)	74 (59, 84)	71 (55, 82)	67 (54, 80)	See Figure 2A					
Sex										
F	1,838,516 (85%)	312,952 (15%)	2,106,417 (98%)	45,051 (2.1%)	—	—		—	—	
M	1,817,422 (84%)	337,024 (16%)	2,104,282 (98%)	50,164 (2.3%)	1.10	1.09, 1.12	<0.001	1.04	1.00, 1.08	0.050
Index of multiple deprivation percentile (lower = more deprived), OR per 10 unit change	73 (52, 88)	73 (52, 88)	73 (52, 88)	74 (53, 89)	0.99	0.99, 0.99	<0.001	1.00	0.99, 1.01	0.63
Ethnic group										
White	2,869,404 (85%)	520,658 (15%)	3,320,215 (98%)	69,847 (2.1%)	—	—		—	—	
Asian	79,755 (85%)	14,018 (15%)	91,392 (97%)	2,381 (2.5%)	1.07	1.02, 1.12	0.006	1.12	0.96, 1.29	0.14
Black	37,973 (86%)	6,100 (14%)	42,840 (97%)	1,233 (2.8%)	1.00	0.94, 1.08	0.91	1.15	0.94, 1.41	0.16
Mixed	20,336 (86%)	3,217 (14%)	22,920 (97%)	633 (2.7%)	1.00	0.91, 1.10	0.97	1.18	0.85, 1.64	0.32
Other	25,768 (86%)	4,041 (14%)	29,260 (98%)	549 (1.8%)	0.97	0.90, 1.04	0.37	0.92	0.81, 1.05	0.21
Not stated or unknown	622,702 (86%)	101,942 (14%)	704,072 (97%)	20,572 (2.8%)	0.99	0.97, 1.00	0.14	1.12	1.07, 1.18	<0.001
Charlson score, OR per 5 unit change	4 (0, 13)	4 (0, 14)	4 (0, 13)	4 (0, 12)	1.02	1.02, 1.02	<0.001	0.93	0.92, 0.94	<0.001
Years since study start	1.79 (0.95, 2.65)	1.64 (0.85, 2.45)	1.77 (0.94, 2.62)	1.54 (0.76, 2.48)	See Figure 2B					
Hour of day	13 (7, 18)	11 (6, 17)	13 (7, 18)	13 (7, 19)	See Figure 2C					
Day of admission	4 (1, 10)	4 (1, 12)	4 (1, 11)	5 (2, 14)	See Figure 2D					
Hospital										
A	2,062,683 (84%)	380,605 (16%)	2,404,339 (98%)	38,949 (1.6%)						
B	822,680 (87%)	123,483 (13%)	903,336 (95%)	42,827 (4.5%)						
C	397,330 (80%)	99,995 (20%)	489,641 (98%)	7,684 (1.5%)						
D	373,245 (89%)	45,893 (11%)	413,383 (99%)	5,755 (1.4%)						
Day of the week										
Monday	464,451 (85%)	84,128 (15%)	535,487 (98%)	13,092 (2.4%)	—	—		—	—	
Tuesday	519,132 (85%)	92,847 (15%)	598,002 (98%)	13,977 (2.3%)	1.01	1.00, 1.02	0.021	0.98	0.95, 1.00	0.094
Wednesday	543,116 (85%)	96,521 (15%)	625,927 (98%)	13,710 (2.1%)	1.01	1.00, 1.03	0.022	0.92	0.89, 0.94	<0.001
Thursday	544,826 (85%)	96,041 (15%)	627,316 (98%)	13,551 (2.1%)	1.01	1.00, 1.02	0.18	0.90	0.88, 0.93	<0.001

Friday	553,595 (85%)	98,125 (15%)	638,411 (98%)	13,309 (2.0%)	1.02	1.01, 1.03	0.001	0.87	0.85, 0.90	<0.001
Saturday	545,098 (85%)	96,060 (15%)	627,183 (98%)	13,975 (2.2%)	1.01	1.00, 1.02	0.23	0.94	0.91, 0.97	<0.001
Sunday	485,720 (85%)	86,254 (15%)	558,373 (98%)	13,601 (2.4%)	0.99	0.98, 1.00	0.21	1.01	0.98, 1.04	0.46
Specialty group										
Acute and emergency medicine	508,050 (82%)	111,586 (18%)	608,483 (98%)	11,153 (1.8%)	—	—		—	—	
General surgery	745,512 (87%)	116,295 (13%)	850,044 (99%)	11,763 (1.4%)	0.83	0.82, 0.85	<0.001	0.65	0.62, 0.69	<0.001
Cardiothoracic surgery	91,147 (85%)	15,799 (15%)	105,359 (99%)	1,587 (1.5%)	0.90	0.86, 0.93	<0.001	0.81	0.76, 0.86	<0.001
ENT, plastic surgery, maxillofacial surgery, ophthalmology	139,133 (82%)	31,556 (18%)	168,002 (98%)	2,687 (1.6%)	1.28	1.25, 1.32	<0.001	0.76	0.71, 0.81	<0.001
Neurosurgery	166,657 (85%)	28,451 (15%)	192,541 (99%)	2,567 (1.3%)	0.95	0.92, 0.98	0.003	0.69	0.65, 0.73	<0.001
Trauma & orthopaedics	560,531 (88%)	79,728 (12%)	631,204 (99%)	9,055 (1.4%)	0.79	0.77, 0.82	<0.001	0.79	0.76, 0.82	<0.001
Obstetrics and gynaecology	110,459 (90%)	12,569 (10%)	121,549 (99%)	1,479 (1.2%)	0.71	0.68, 0.73	<0.001	0.63	0.59, 0.68	<0.001
Cardiology	166,860 (82%)	36,390 (18%)	200,358 (99%)	2,892 (1.4%)	1.00	0.98, 1.02	0.95	0.80	0.77, 0.84	<0.001
Gastroenterology	105,353 (83%)	22,055 (17%)	125,433 (98%)	1,975 (1.6%)	0.93	0.90, 0.96	<0.001	0.89	0.83, 0.95	<0.001
Geratology	352,487 (82%)	77,061 (18%)	421,072 (98%)	8,476 (2.0%)	0.96	0.94, 0.97	<0.001	1.09	1.05, 1.13	<0.001
Infectious diseases	178,380 (83%)	37,148 (17%)	211,708 (98%)	3,820 (1.8%)	0.93	0.91, 0.95	<0.001	0.95	0.89, 1.01	0.10
Nephrology	77,836 (87%)	11,646 (13%)	88,091 (98%)	1,391 (1.6%)	0.76	0.73, 0.80	<0.001	0.74	0.67, 0.81	<0.001
Haematology and oncology	246,620 (89%)	29,390 (11%)	243,882 (88%)	32,128 (12%)	0.65	0.62, 0.67	<0.001	5.56	5.13, 6.03	<0.001
Respiratory medicine	53,766 (82%)	11,945 (18%)	64,535 (98%)	1,176 (1.8%)	1.13	1.08, 1.19	<0.001	0.98	0.91, 1.05	0.53
Other	153,147 (84%)	28,357 (16%)	178,438 (98%)	3,066 (1.7%)	0.92	0.90, 0.95	<0.001	0.92	0.87, 0.97	0.003

Table 1. Relationship between temperature readings of 36.0°C and blood pressure readings ending in zero and associated factors. Blood pressure readings where both the systolic and diastolic blood pressure end in zero are shown. Descriptive data for patient factors (age, sex, ethnicity, Index of multiple deprivation, Charlson score), are summarised per observation, rather than per patient. Therefore, for example as older patients had longer hospital stays and more observations the median age is higher than the median age on a per patient basis. ENT, ear, nose and throat surgery; BP, blood pressure. ¹Median (IQR); n (%). ²OR = Odds Ratio, CI = Confidence Interval.

Figures

Figure 1. Observed distribution of temperature, systolic blood pressure, respiratory rate and heart rate recordings (n=4,305,914). Readings showing potential value preferences are shown in orange/red. SBP, systolic blood pressure; DBP, diastolic blood pressure. Values below the 1st percentile or above the 99th percentile are omitted for visualisation purposes.

Figure 2. Non-linear associations between factors and temperature readings of 36.0°C and blood pressure readings ending in zero. Model predictions from a multivariable model are shown with all other factors set to the reference category or median value as shown in Table 1. The shaded ribbon indicates the 95% confidence interval. BP, blood pressure.

Figure 3. Relationship between specialty and temperature readings of 36.0°C and blood pressure readings ending in zero. ENT, ear, nose and throat; plastics, plastic surgery; maxfax, maxillofacial surgery; BP, blood pressure.

Supplementary Figure

Figure S1. Observed and fitted distribution of temperature measurements. The red line indicates a gamma distribution fitted to the data. Based on the gamma distribution (with parameters: shape=4416, rate=121) 4.9% of values are expected to lie between 35.95°C and 36.05°C.

References

- 1 Evans RS. Electronic Health Records: Then, Now, and in the Future. *Yearb Medical Informatics* 2016; 25: S48–61.
- 2 Safran C, Bloomrosen M, Hammond WE, *et al.* Toward a National Framework for the Secondary Use of Health Data: An American Medical Informatics Association White Paper. *J Am Med Inform Assn* 2007; 14: 1–9.
- 3 Casey JA, Schwartz BS, Stewart WF, Adler NE. Using Electronic Health Records for Population Health Research: A Review of Methods and Applications. *Annu Rev Publ Health* 2015; 37: 1–21.
- 4 Lin H, Tang X, Shen P, *et al.* Using big data to improve cardiovascular care and outcomes in China: a protocol for the CHinese Electronic health Records Research in Yinzhou (CHERRY) Study. *Bmj Open* 2018; 8: e019698.
- 5 Morley KI, Wallace J, Denaxas SC, *et al.* Defining Disease Phenotypes Using National Linked Electronic Health Records: A Case Study of Atrial Fibrillation. *Plos One* 2014; 9: e110900.
- 6 Yang Z, Huang Y, Jiang Y, Sun Y, Zhang Y-J, Luo P. Clinical Assistant Diagnosis for Electronic Medical Record Based on Convolutional Neural Network. *Sci Rep-uk* 2018; 8: 6329.
- 7 Syed S, Ashwick R, Schlosser M, Gonzalez-Izquierdo A, Li L, Gilbert R. Predictive value of indicators for identifying child maltreatment and intimate partner violence in coded electronic health records: a systematic review and meta-analysis. *Arch Dis Child* 2020; 106: 44–53.
- 8 Cooper GF, Villamarin R, Tsui F-C (Rich), Millett N, Espino JU, Wagner MM. A method for detecting and characterizing outbreaks of infectious disease from clinical reports. *J Biomed Inform* 2015; 53: 15–26.
- 9 Greene SK, Huang J, Abrams AM, *et al.* Gastrointestinal Disease Outbreak Detection Using Multiple Data Streams from Electronic Medical Records. *Foodborne Pathog Dis* 2012; 9: 431–41.
- 10 Ray WA. Improving Automated Database Studies. *Epidemiology* 2011; 22: 302–4.
- 11 Terris DD, Litaker DG, Koroukian SM. Health state information derived from secondary databases is affected by multiple sources of bias. *J Clin Epidemiol* 2007; 60: 734–41.
- 12 Verheij RA, Curcin V, Delaney BC, McGilchrist MM. Possible Sources of Bias in Primary Care Electronic Health Record Data Use and Reuse. *J Med Internet Res* 2018; 20: e185.
- 13 Köpcke F, Trinczek B, Majeed RW, *et al.* Evaluation of data completeness in the electronic health record for the purpose of patient recruitment into clinical trials: a retrospective analysis of element presence. *Bmc Med Inform Decis* 2013; 13: 37.
- 14 Verweij LM, Tra J, Engel J, Verheij RA, Bruijne MC de, Wagner C. Data quality issues impede comparability of hospital treatment delay performance indicators. *Neth Heart J* 2015; 23: 420–7.
- 15 Brennan L, Watson M, Klaber R, Charles T. The importance of knowing context of hospital episode statistics when reconfiguring the NHS. *Bmj Br Medical J* 2012; 344: e2432.

- 16 Fawcett N, Young B, Peto L, *et al.* 'Caveat emptor': the cautionary tale of endocarditis and the potential pitfalls of clinical coding data—an electronic health records study. *Bmc Med* 2019; 17: 169.
- 17 Hogan WR, Wagner MM. Accuracy of Data in Computer-based Patient Records. *J Am Med Inform Assn* 1997; 4: 342–55.
- 18 Skyttberg N, Chen R, Blomqvist H, Koch S. Exploring Vital Sign Data Quality in Electronic Health Records with Focus on Emergency Care Warning Scores. *Appl Clin Inform* 2017; 08: 880–92.
- 19 HESSEL PA. Terminal Digit Preference in Blood Pressure Measurements: Effects on Epidemiological Associations. *Int J Epidemiol* 1986; 15: 122–5.
- 20 Nietert PJ, Wessell AM, Feifer C, Ornstein SM. Effect of Terminal Digit Preference on Blood Pressure Measurement and Treatment in Primary Care. *Am J Hypertens* 2006; 19: 147–52.
- 21 Wingfield D, Freeman GK, Bulpitt CJ, (GPHSG) GPHSG. Selective recording in blood pressure readings may increase subsequent mortality. *Qjm Int J Medicine* 2002; 95: 571–7.
- 22 Hense HW, Kuulasmaa K, Zaborskis A, Kupsc W, Tuomilehto J. Quality assessment of blood pressure measurements in epidemiological surveys. The impact of last digit preference and the proportions of identical duplicate measurements. WHO Monica Project [corrected]. *Revue D'épidémiologie Et De Santé Publique* 1990; 38: 463–8.
- 23 Granholm A, Pedersen NE, Lippert A, Petersen LF, Rasmussen LS. Respiratory rates measured by a standardised clinical approach, ward staff, and a wireless device. *Acta Anaesth Scand* 2016; 60: 1444–52.
- 24 Clifton DA, Clifton L, Sandu D-M, *et al.* 'Errors' and omissions in paper-based early warning scores: the association with changes in vital signs—a database analysis. *Bmj Open* 2015; 5: e007376.