



COMMENT

What can the randomness of missing values tell you about clinical practice in large data sets of children's vital signs?

Damian Roland^{1,2}, Neslihan Suzen³, Timothy J. Coats⁴, Jeremy Levesley³, Alexander N. Gorban³ and Evgeny M. Mirkes³

Pediatric Research (2021) 89:16–21; <https://doi.org/10.1038/s41390-020-0861-2>

INTRODUCTION

Currently, there are 5.5 million Emergency Department (ED) attendances in England by Children and Young People making up 26% of all age attendances.¹ The Nuffield trust has reported a 9% increase in emergency or unplanned admissions between June 2005 and June 2015² and the Royal College of Paediatrics and Child Health says “attendances of children are projected to increase by 50% by 2030 if current trends are maintained”.¹

Initiatives are needed to improve public education, expectation and acute care access as well as strategies to improve the health of children by addressing current huge societal inequalities. However, even with these interventions, attendances are still likely to grow and healthcare organisations are going to need to respond to this growth in demand. One facet of the response will be an increased use of information technology where there is already a national digital strategy. The desire is that by 2020, “technology and data in the form of digitally enabled care will be used by most citizens and will help to meet their demand for better and safer care”.³

In relation to Emergency Care there is a growing trend in the use of e-observations to record vital signs and other relevant assessments of patients. Vital signs are observations that are used to make clinical decisions and commonly include heart rate, breathing rate and so on, with other pertinent assessments being more subjective observations such as overall appearance. “e-Observations” are the records of data in an electronic format (i.e. the nurse will still need to manually assess the heart rate, but will record this result digitally).

Utilising e-Observations creates a paradigm shift in the ability of organisations to understand and respond to their workflows. Their use creates opportunities to improve patient safety through automated alert systems and also provides mechanisms for better quality improvement and assurance.⁴ However, given the volume of data collected, there will be new challenges in how these data are interpreted and utilised. Missing data are a ubiquitous problem in healthcare datasets,⁵ but particularly unique to child health care is the frequency of missing data as a result of being unable to obtain results (due to the clinical challenge of obtaining observations in infants) and the impact of the non-standard mechanisms of obtaining vital signs that are sometimes used. An example of the latter is the use of pulse oximetry devices (which measure the saturation of oxygen in the blood) to also take the heart rate (on which the pulse oximetry depends for its

calculation).⁶ A failure in either system may mean neither result is recorded.

There is already a body of evidence regarding missing values in large data sets. Moreover, there is special classification of different types of missing values.^{7,8} There are different ways to deal with missing data values, from deletion to imputation⁵ or specific modification of the database, for example, reweighting.⁵ A deep analysis of the patterns of missing data can reveal sources of omission and help to find ways for improving future data collection. These are general techniques, however, not specific to the practice of child health care. The increasing digitisation of healthcare will make large data sets more common and clinicians will be increasingly able to conduct their own analysis. This makes the need to understand missing data more important as the way in which data are missing and the way in which the missing data are handled in analysis, and modelling can cause bias and incorrect conclusions.

The paediatric observation priority score (POPS) is a methodology to assess the acuity of children presenting to urgent and emergency care environments. POPS generates a total score (0–16) as the combination of eight physiological, behavioural and known-risk parameters: oxygen saturations (Sats), level of alertness (AVPU), extent of breathing difficulty (Breathing), background history (Other), nurse gut feeling (Gut Feeling), heart rate (Pulse), respiratory rate (RR) and temperature (Temp).^{9,10} Each of the variables above has a range of responses, and each response is automatically converted to a number 0, 1 or 2. The total score is the sum of sub-scores to give a possible total of 16 points with the rules for transformation of raw POPS variables into score variables given in Table 1. POPS has been utilised in the ED at the Leicester Royal Infirmary since 2012 and has been utilised in both paper and digital forms.

In this study, we have analysed POPS data that were collected by a bespoke web-based application, completed by nurses on the initial assessment of children when they arrived at the Leicester Royal Infirmary children's ED. Only initial observations were recorded electronically and not all patients had an electronic POPS (ePOPS) recorded due to short-term problems with the computer system or the child being identified as so ill as to need immediate treatment.

The objective of this study was to analyse missing data where an ePOPS record was started, understand the distributions of missing data and identify the dependency of variables between one another.

¹SAPPHIRE Group, Health Sciences, University of Leicester, Leicester, UK; ²Paediatric Emergency Medicine Leicester Academic (PEMLA) Group, Children's Emergency Department, Leicester Royal Infirmary, Leicester, UK; ³School of Mathematics and Actuarial Science, University of Leicester, Leicester, UK and ⁴Emergency Medicine Academic Group, Emergency Department, Leicester Royal Infirmary, Leicester, UK
Correspondence: Damian Roland (dr98@le.ac.uk)

Received: 2 November 2019 Revised: 27 January 2020 Accepted: 26 February 2020
Published online: 15 April 2020

Table 1. POPS chart.

Age	Variable	Score					
		2	1	0	1	2	
Any	Sats	<90%	90–94%	>95%	90–94%	<90%	
Any	Breathing	Stridor	Audible grunt or wheeze	No distress	Mild or moderate recession	Severe recession	
Any	AVPU	Pain	Voice	Alert	Voice	Pain	
Any	Gut Feeling	High-level concern	Low-level concern	Well	Low-level concern	Child looks unwell	
Any	Other	Oncology patient	Significant PMH ^a		Significant PMH ^a	Congenital heart disease	
0–1	Pulse	<90	90–109	110–160	161–180	>180	
	RR	<25	25–29	30–40	41–50	>50	
	Temp	<35 °C	35–35.9 °C	36–37.5 °C	37.6–39 °C	>39 °C	
1–2	Pulse	<90	90–99	100–150	151–170	>170	
	RR	<20	20–24	25–35	36–50	>50	
	Temp	<35 °C	35–35.9 °C	36–37.9 °C	38–40 °C	>40 °C	
2–4	Pulse	<80	80–94	95–140	141–160	>160	
	RR	<20	20–24	25–30	31–40	>40	
	Temp	<35 °C	35–35.9 °C	36–37.9 °C	38–40 °C	>40 °C	
5–12	Pulse	<70	70–79	80–110	111–150	>150	
	RR	<15	15–19	20–25	26–40	>40	
	Temp	<35 °C	35–35.9 °C	36–37.9 °C	38–40 °C	>40 °C	
13–16	Pulse	<50	50–59	60–100	101–110	>110	
	RR	<12	12–14	15–20	21–25	>25	
	Temp	<35 °C	35–35.9 °C	36–37.9 °C	38–40 °C	>40 °C	

POPS (paediatric observation priority score) is copyrighted (creative commons attribution non-commercial share-alike 4.0), Dr. Damian Roland and Dr. Ffion Davies 2010. This is version 1.3, August 2016.

^aSignificant PMH includes: ex-premature, syndromic conditions, cardiac problems, asthma, diabetes, long-term steroids, all other chronic conditions.

METHODS

Data were collected from January of 2014 to September of 2016 and consisted of 56,691 records and 55 variables, which are in formats of date, number and text. Inappropriate data, for example, values such as “NA”, “NASA” and “pink” as responses in the oxygen saturation column, were transformed to missed values. After removing duplicated records and records with age >16 years, a database of 56,042 patients was used in the study.

The ePOPS score is calculated from the score variables included in POPS (heart rate, breathing rate, etc.).

Analysis of missing data requires an understanding of whether or not the value is missing at random, or whether there is some systemic reason for some of the data to be missing. The classification of randomness of missing values is defined as below:^{6,7}

- Data in specified variable are missed completely at random (MCAR) if probability of missing is independent of the missing value (of this variable) and of values of any other variables.
- Data in specified variable are missed at random (MAR) if probability of missing is independent of the missing value (of this variable), but can depend on other variables.
- Data in specified variable are missed not at random (MNAR) if probability of missing depends on the missing value of this variable.

For example, each variable on the POPS score has 3 possible classifications (Table 1) giving a component score of 0, 1 and 2, (i.e. no derangement, moderate derangement and severe derangement). If the amount of data missing in each component is proportional to the amount of data in each, the data are MAR. Otherwise, we can conclude that data are MNAR.

Another relevant classification is that of randomness of observations is:⁶

- Data for a particular variable are observed at random (OAR) if the probability of missing is independent of the other variables.
- Data for a particular variable are observed not at random (ONAR) if the probability of missing is dependent on the other variables.

For example, if pulse rate is always taken from the saturation monitor, it is likely that pulse rate and oxygen saturation will be missing together. In this case data are ONAR.

Analysis of the patterns of missing data is a crucial part of any analysis of real data.^{6,11} For this analysis, it is necessary to transform score variables to a binary variable by the following way: a known value is encoded as 1 and missing value is encoded as 0. Such variables allow the analysis of correlation of the missing data: are the two variables missing together rarer or more frequent than independently distributed data.

Our first hypothesis is that missing values in all POPS variables are independent (OAR). If our hypothesis is wrong, then data are not MCAR and not OAR. To check this hypothesis, we evaluate the probability of missing the value of variable *i* as the proportion p_i of records with missing value in this variable among all records (this calculation is available in the Supplementary Appendix). The process proceeds by considering a pair of variables and testing the hypothesis that there is no correlation between the incidence of missing data in this pair. The Hamming distance between two sequences is linear function of the Pearson’s correlation coefficient (PCC) for binary variables, and we utilised the PCC as the ability to use confidence intervals. If PCC is significantly different from 0, we would conclude (if data value is missing in one of the pairs) that data value is not missing at random in the other of the pair (the properties of one of the pair can be inferred to some extent from the other).

We repeat the same analysis for all possible pairs. The statistical test of equality of proportions utilised for the different categories of missing data was the χ^2 test of contingency table data.

Ethical approval was obtained from NHS REC East Midlands 11/EM/0351.

RESULTS

A description of the data set of 56,042 patients is shown in Table 2. The number of readings and the fraction of missing values for each variable is shown in Table 3. The ePOPS score was recalculated according to the rules in Table 1.^{9,10} After all corrections, the data contain 18,221 records, which missed at least one missing value across all of the POPS variables.

To check the first hypothesis that missed data in all variables are missed independently (OAR), we use the fractions from Table 3 as estimates of p_i and check the probabilities of observing (i) complete (C) records (without any missing values, $m_i = 1, i = 1, \dots, 8$), (ii) completely missing (CM) records (all individual score variables are missing, $m_i = 0, i = 1, \dots, 8$), and all other (254) missing data patterns.

In each case, the p value was very small, although it should be noted that these values are not independent due to the marginal total being fixed. Results of the calculations for C and CM are given in Table 4. We have enough evidence to reject the hypothesis that data are missing independently since the p value from the χ^2 test for the full contingency table is $<10^{-300}$.

A random distribution of missing data across all cases would mean that most would be expected to have some missing data (Table 4); however, this is not the case as most of the POPS data set cases tend to have either all complete or all missing data. Thus, it is unlikely that data are missing (observed) at random.

The test of the hypothesis that the distribution of missed data for pairs of ePOPS variables are independent is based on the analysis of PCC, which are presented in Table 5. We can see that there is very high correlation between Breathing, AVPU, Gut Feeling and Other (at least 0.87 with 99% confidence interval [0.868, 0.872]) and between Sats, Pulse, RR and Temp (at least 0.90 with 99% confidence interval [0.898, 0.902]). We observe considerably less correlation for variables from different groups (at most 0.78 with 99% confidence interval [0.776, 0.784]). As we can see, the lower confidence limits of all confidence intervals of PCC inside each group are greater than the upper confidence limits of all confidence intervals of PCC between groups. This means that we can state with 99% confidence that the intra-correlations (inside each of the above described groups) are greater than the inter-correlations (between groups). This means that the grouping of Breathing, AVPU, Gut Feeling and Other behaves in a different way than Sats, Pulse, RR and Temp in relation to the extent of missing data.

Since Table 4 shows drastic differences between predicted (for independence hypothesis) and observed numbers of complete and CM records we repeated the test of the hypothesis of independence of the distribution of missing data for pairs of ePOPS variables with partially missed data only (records with at least one known and at least one missed ePOPS variable). PCCs for this test are presented in Table 6. We can see high correlations inside the group of Breathing, AVPU, Gut Feeling and Other (minimal PCC is 0.3 with 99% confidence interval [0.28, 0.32]) and inside the group of Sats, Pulse, RR, and Temp, exclude insignificant correlation of Temp with Sats (minimal PCC is 0.49 with 99% confidence interval [0.473, 0.507]). Maximal correlation coefficient between these two groups of variables is 0.23 with 99% confidence interval [0.209, 0.251]. This confirms the groupings found in Table 5.

The relationship between an initial assessment normal value (zero score) and subsequent missing measured values is shown in Table 7. If the initial assessment of breathing has a normal value,

Table 2. Sample description.

Variable	Value	No. of records	Fraction	
Total		56,042	100%	
POPS score	0	21,448	38.27%	
	1	6737	12.02%	
	2	3938	7.03%	
	3	2380	4.25%	
	4	1507	2.69%	
	5	892	1.59%	
	6	529	0.94%	
	7	246	0.44%	
	8	96	0.17%	
	9	34	0.06%	
	10	10	0.02%	
	11	3	0.01%	
	15	1	0.00%	
	Age	Missed	18,221	32.51%
		0	6,125	10.93%
1		7069	12.61%	
2		5514	9.84%	
3		4394	7.84%	
4		3626	6.47%	
5		3053	5.45%	
6		2,691	4.80%	
7		2481	4.43%	
8		2498	4.46%	
9		2617	4.67%	
10		2666	4.76%	
11		2755	4.92%	
12		2705	4.83%	
13		2618	4.67%	
14		2628	4.69%	
15		2578	4.60%	
16	24	0.04%		

Table 3. Missing data in the database for each field (totally 56,042 records).

Variable	Missing data	
	Number	Fraction
Age	0	0.00%
POPS variables		
Breathing score	22,444	40.05%
AVPU score	19,120	34.12%
Gut feeling score	19,944	35.59%
Other score	20,060	35.79%
Sats score	29,452	52.55%
Pulse score	28,417	50.71%
RR score	28,591	51.02%
Temp score	28,540	50.93%

each of the measured variables are missing more than 60% of patients. If either AVPU, Gut Feeling and Other is normal on initial assessment more than 80% of patients have missing values of Sats, Pulse, RR and Temp. If all of the initial assessment variables

Table 4. Expected and observed number of records with complete, partially missing and completely missing POPS variables.

	Expected number of records	Observed number of records	Observed fraction of record	p Value
Complete	67	25,114	45%	<10 ⁻³⁰⁰
Partially missing	55,461	12,707	23%	<10 ⁻³⁰⁰
Completely missing	514	18,221	32%	<10 ⁻³⁰⁰

Table 5. Pearson's correlation coefficient for each pair of POPS variables.

	Breathing	AVPU	Gut Feeling	Other	Sats	Pulse	RR	Temp
Breathing	–	0.87	0.90	0.89	0.76	0.78	0.78	0.76
AVPU	0.87	–	0.96	0.94	0.67	0.69	0.69	0.67
Gut Feeling	0.90	0.96	–	0.97	0.69	0.71	0.71	0.70
Other	0.89	0.94	0.97	–	0.69	0.72	0.71	0.70
Sats	0.76	0.67	0.69	0.69	–	0.96	0.95	0.90
Pulse	0.78	0.69	0.71	0.72	0.96	–	0.98	0.93
RR	0.78	0.69	0.71	0.71	0.95	0.98	–	0.93
Temp	0.76	0.67	0.70	0.70	0.90	0.93	0.93	–

Bold shows pairs with correlation coefficient ≥0.8.

Table 6. Correlations for each pair of score variables (partially missed data only).

	Breathing	AVPU	Gut Feeling	Other	Sats	Pulse	RR	Temp
Breathing	–	0.30	0.47	0.42	0.10	0.22	0.21	0.11
AVPU	0.30	–	0.61	0.41	–0.13	–0.05	–0.05	–0.23
Gut Feeling	0.47	0.61	–	0.74	–0.03	0.05	0.05	–0.08
Other	0.42	0.41	0.74	–	–0.04	0.04	0.04	–0.08
Sats	0.10	–0.13	–0.03	–0.04	–	0.72	0.57	0.15
Pulse	0.22	–0.05	0.05	0.04	0.72	–	0.87	0.53
RR	0.21	–0.05	0.05	0.04	0.57	0.87	–	0.49
Temp	0.11	–0.23	–0.08	–0.08	0.15	0.53	0.49	–

Bold shows pairs with correlation coefficient ≥0.3.

are normal (Breathing, AVPU, Gut Feeling and Other), then 56% (5212) of patients had no further observations recorded (missed values for all of Sats, Pulse, RR, and Temp).

Table 8 shows that fraction of abnormal values in Sats, Pulse, RR and Temp variables is relatively variable for records with normal values of Breathing, AVPU, Gut Feeling and Other.

DISCUSSION

Missing data are a common problem in digital healthcare databases due to the volume of the data and this problem is likely to increase as the use of electronic patient records becomes more widespread. The data used in this study are typical for healthcare as they include significant amount of missing data, from electronic records, for both scores and score domains, with nearly 32% of records do not contain any POPS variables (Table 4). Data gaps are similarly high for any of the eight POPS variables. Nearly 22% of patients were not recorded in at least one of these variables. Missing data are not a problem unique to the POPS system,¹² and in situations where children are being assessed, it is likely, regardless of the skill of the practitioner, it will not always be possible to obtain an accurate reading in a short time scale. There is an important distinction between “a zero value”, “recording

attempted but cannot be made at this time” and “recording not attempted”. This distinction is not often made in healthcare data. The use of specific data notations to highlight why an observation has not been recorded is important. Like many NHS applications, the web-based application recording the POPS data did not have the ability to record why data were not entered. This has subsequently been improved upon in our hospital's current electronic observation system.

The lack of recording of the reasons for missing data in this data set means that we had to explore aggregate data and examine whether there were any patterns to the missing data. The missing data were not missing at random, that is, there was a dependence of the distribution of missing data on the individual components of POPS. There were groups where all of the data were missing, but within the patients where data were partially missing, there seemed to be two groups of variables with heart rate, breathing rate, temperature and oxygen saturations forming one group and AVPU, Work of Breathing, Gut Feeling and Other forming a second grouping. Within each group having missing values was highly correlated, that is, missing values in heart rate were likely to correlate with missing values in breathing rate.

The two groups might correspond to the usual clinical practice of staff taking observations. When a child first presents to initial

Table 7. Fraction of missing values in Sats, Pulse, RR and Temp, which corresponds to normal (0) value of Breathing, AVPU, Gut Feeling and Other among all missed values.

Normal value of	Missed values of				
	Sats	Pulse	RR	Temp	All
Breathing	65%	61%	62%	64%	61%
AVPU	94%	94%	94%	96%	96%
Gut Feeling	84%	83%	83%	84%	85%
Other	84%	83%	83%	85%	85%
All	59%	50%	51%	53%	56%

Table 8. Fraction of normal values in Sats, Pulse, RR, and Temp for records with normal value of Breathing, AVPU, Gut Feeling and Other for complete subsample of data set.

	Sats	Pulse	RR	Temp
Breathing	99.02%	70.54%	78.29%	77.72%
AVPU	97.02%	67.80%	71.86%	76.81%
Gut Feeling	98.78%	71.89%	78.45%	78.76%
Other	97.20%	67.85%	72.36%	76.79%

assessment a number of the variables are immediately obvious (such as level of consciousness, pattern of breathing and the clinician has an immediate gut feeling about the child). These more subjective components of POPS are usually determined just by looking at the patient during initial assessment, and therefore group together. The other group of variables are recorded slightly later and need to be measured, normally using a pulse oximeter to simultaneously measure pulse rate and oxygen saturation, a separate device (thermometer) to measure temperature and the clinician to stop and time 60 s to measure respiratory rate. This set of measurements is often performed as a single episode. Each of these two sets of activity is either done or not done, leading to grouping of the components of the POPS score. Analysis of patterns of missing data by mathematicians with no knowledge of clinical workflows was able to describe groups, which corresponded with the complexity of clinical activity.

Our data demonstrated that if the initial assessment variables were normal (0), then there was a higher than expected chance that the measured variables would be missing. This would fit with the clinical practice model. The assessment variables are much quicker to ascertain as they are slightly more subject and based on observation alone. In minor presentations (such as a sprained ankle) staff might feel that there is no need to record their initial assessment subjective impression as it so obvious that the child will be normal (so all POPS score components will be missing). Even in presentations that could be more serious, if the nursing staff felt that the child was very well, subjectively, they may be less inclined to undertake a full set of vital sign measurements (measurement variables). While it might be considered best practice to always undertake a full set of observations in real life, this is unlikely to occur for children with minor injuries or very minor illness where only a rapid assessment approach may be undertaken.

Understanding a pattern of clinician behaviour through analysis of missing data has the potential to improve both clinical practice and data acquisition. Clinical practice could be improved by using an analysis of whether or not the clinician's assumption (further observations are not needed) was correct or not. Data acquisition could be improved if the analysis of missing data shows that the

electronic record system is not well designed to fit with clinical practice.

From a missing data analysis, it is not possible to know whether or not the clinician correctly decided to skip a recording of data. Either (1) clinicians are correctly deciding to skip unnecessary steps in a data collection, system which is not sufficiently individualised, or (2) clinicians are for some reason giving poor care. To differentiate these, future research could look at outcomes (length of ED stay, final diagnosis, admission, etc.) in patients who have missing data to judge whether or not it was appropriate to miss out this step in documentation.

It is easy to design an electronic record that mandates data entry, but inflexible systems that do not fit with clinical workflows or are poorly designed may have adverse unintended consequences.¹³ The missing data analysis presented here points to a more flexible approach—if the measured variables are not recorded by the clinician, the system could automatically double check if this is the sort of patient type or presentation (defined by clinical rules or an AI-based algorithm) in which the measured observations (heart rate, breathing rate, temperature and oxygen saturations) are not needed. If data collection was usually needed in this type of patient, a prompt could be given to the clinician. This “precision medicine” approach would individualise clinical data collection and provide the clinician with an “intelligent assistant”.

The way that healthcare professionals work means that in clinical data science missing data are likely to be non-random. The extremes of the very sick (observations not being able to be recorded due to pressure of time or changing physiology) or the very well (observations deliberately missed as they were not felt to be necessary) increase the non-random nature of missing data. This must be accounted for in clinical data systems with appropriate labelling of blank fields, otherwise the available data may not be representative of the cohort being studied and analytics (whether statistical, artificial intelligence or machine learning) may lead to bias in the analysis and incorrect conclusions.

CONCLUSION

Describing missing data is an important part of data analysis and can be linked to healthcare professional practice patterns. It is important that lead clinicians with responsibility for big data sets understand the concepts of classification of randomness of missing data and as these data sets become more available to clinicians, and analytics easier to perform, there must be increased awareness of the challenges, and dangers, presented by missing data.

ACKNOWLEDGEMENTS

This study received a small grant (£10000) from the Leicester University REF Impact Fund.

AUTHOR CONTRIBUTIONS

Substantial contributions to conception and design, acquisition of data, or analysis and interpretation of data were made by D.R., N.S., A.N.G. and E.M.M. Drafting the article or revising it critically for important intellectual content were made by D.R., J.L., T.J.C., A.N.G. and E.M.M. Final approval of the version to be published was done by all authors.

ADDITIONAL INFORMATION

The online version of this article (<https://doi.org/10.1038/s41390-020-0861-2>) contains supplementary material, which is available to authorised users.

Competing interests: The authors declare no competing interests.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES

1. Viner, R., Ward, J., Cheung, R., Wolfe, I. & Hargreaves, D. *Child Health in 2030 in England: Comparisons with Other Wealth Countries*. RCPCH 2030 Report https://www.rcpch.ac.uk/sites/default/files/2018-10/child_health_in_2030_in_england_report_2018-10.pdf (2018).
2. Kossarova, L., Hargreaves, D., Cheung, R. & Keeble, E. *Admissions of Inequality: Emergency Hospital Use for Children and Young People*. Briefing, Nuffield Trust <https://www.nuffieldtrust.org.uk/research/admissions-of-inequality-emergency-hospital-use-for-children-and-young-people> (2017).
3. Health and Social Care Information Centre. *Information and Digital Technology for Better Care: Our Strategy for 2015–2020*. NHS Digital <https://digital.nhs.uk/about-nhs-digital/corporate-information-and-documents/our-strategy>. Accessed 23 Nov 2018.
4. Sefton, G. et al. Accuracy and efficiency of recording pediatric early warning scores using an electronic physiological surveillance system compared with traditional paper-based documentation. *Comput. Inf. Nurs.* **35**, 228–236 (2017).
5. Mirkes, E. M., Coats, T. J., Levesley, J. & Gorban, A. N. Handling missing data in large healthcare dataset: a case study of unknown trauma outcomes. *Comput. Biol. Med.* **75**, 203–216 (2016).
6. Singh, JKS et al. Accuracy of pulse oximetry in assessing heart rate of infants in the neonatal intensive care unit. *J. Paediatr. Child Health* **44**, 273–275 (2008).
7. Little, R. J. & Rubin, D. B. *Statistical Analysis with Missing Data*, Vol. 333 (Wiley, 2014).
8. Cismondi, F. et al. Missing data in medical databases: impute, delete or classify? *Artif. Intell. Med.* **58**, 63–72 (2013).
9. Roland, D., Arshad, F., Coats, T. & Davies, F. Baseline characteristics of the paediatric observation priority score in emergency departments outside its centre of derivation. *Biomed. Res. Int.* <https://doi.org/10.1155/2017/9060852> (2017).
10. Roland, D. et al. The paediatric observation priority score: a system to aid detection of serious illness and assist in safe discharge. *Open J. Emerg. Med.* **4**, 38–44 (2016).
11. Heitjan, D. F. Annotation: what can be done about missing data? Approaches to imputation. *Am. J. Public Health* **87**, 548–550 (1997).
12. Gravel, J., Opatrny, L. & Gouin, S. High rate of missing vital signs data at triage in a paediatric emergency department. *Paediatr. Child Health* **11**, 211–215 (2006).
13. Sittig, D. F., Wright, A., Ash, J. & Singh, H. New unintended adverse consequences of electronic health records. *Yearb. Med. Inform.* **25**, 7–12 (2016).