**Appendix 1 - Manual data validation sampling methodology**

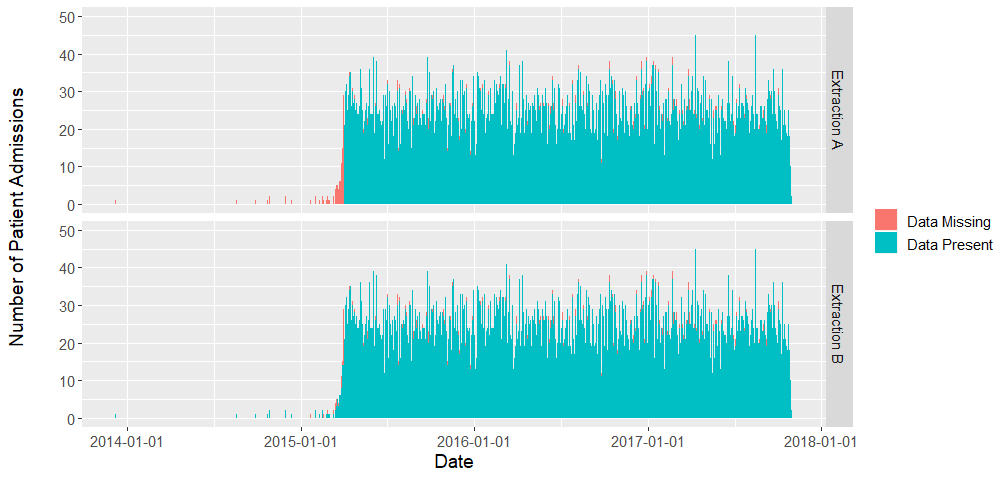
Data extraction, computational data quality checks, and manual data validation all occurred over two rounds: 1) patients discharged between April 1, 2010 and March 31, 2015 and 2) patients discharged between April 1, 2015 and October 31, 2017.

For the first round of manual data validation, we randomly selected 100 admissions to abstract data about radiology testing and laboratory testing at each hospital, and 200 admissions to abstract the admitting and discharging physician name.

For rare events (i.e. death, red blood cell transfusion, and ICU transfer) we used the extracted data to oversample rare events and to select a common sample of cases in which none of the events occurred to enable more efficient data abstraction. The data abstractor was blinded to the extracted clinical data and reviewed medical records from each admission in detail for the occurrence of all events (death, red blood cell transfusion, and ICU transfer). At each hospital, we sampled the following numbers of patient records, based on extracted data.

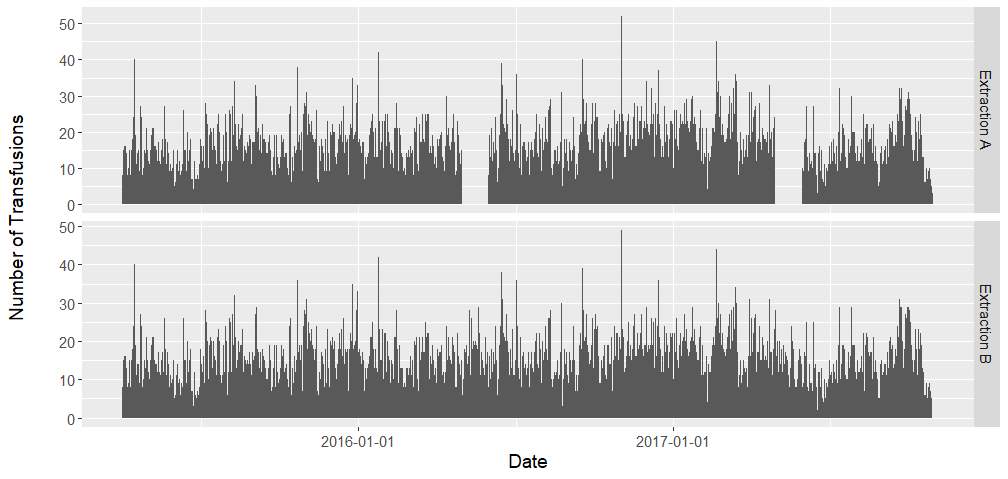
* 100 admissions for patients who died without a blood transfusion or ICU transfer
* 100 admissions for patients who had an ICU transfer without a transfusion or death
* 100 admissions for patients who had a blood transfusion without death or ICU transfer
* 500 admissions with no death, ICU transfer, or blood transfusion

**Appendix 2. Examples of data quality issues identified by computational quality checks.**



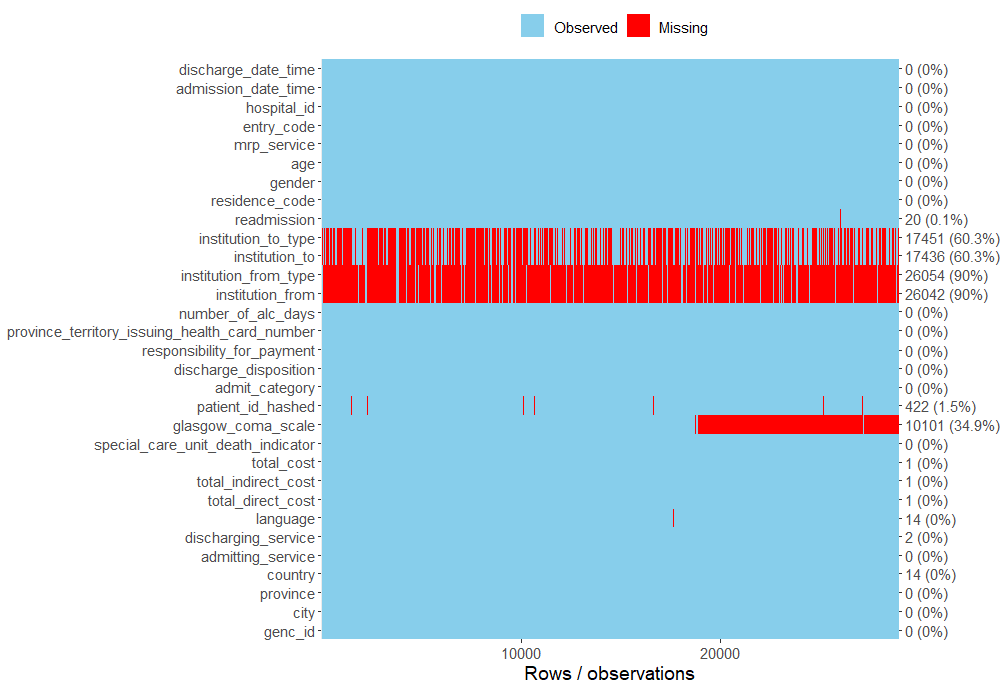
Appendix Figure 1. Patient admissions in ED Case Mix Group data table*.*

Appendix Figure 1 Description. This is an example of Computational Data Quality Check 1: Patient admissions over time. Each histogram bar represents the number of patient admissions occurring on the discharge date (x-axis) from the Admission data table that are present or missing in the ED Case Mix Group data table. In the first extraction (Extraction A, top panel), patient admissions that occurred before April 1, 2015 were missing (red histogram bars) from the data table. After identifying this problem, the data were re-extracted (Extraction B, bottom panel), which demonstrates that the issue was resolved (as indicated by those bars are green). ED: Emergency Department.



Appendix Figure 2. Number of transfusions over time.

Appendix Figure 2 Description. This is an example of Computational Data Quality Check 2: Data volume over time. In the first extraction (Extraction A, top panel), blood transfusion data were entirely missing for two periods (May 2016 and May 2017). Data were re-extracted (Extraction B, bottom panel) and the issue was resolved.



Appendix Figure 3. Variable presence over time heat map.

Appendix Figure 3 Description. This is an example of Computational Data Quality Check 5: Variable presence over time heat map. This tool was used to investigate clustering and temporal patterns in the missingness of data for all variables in a data table. The blue colour indicates data are present for a specific variable, and the red areas indicate data are missing. This computational data quality check identified how Glasgow Coma Scale data after 2015-04-01 were missing from the Admission data table.

**Appendix 3. Manual data validation results by hospital site.**

In the tables below, the following abbreviations are used:

AST=Aspartate transaminase; INR=international normalized ratio; CT=Computerized tomography; MRI=magnetic resonance imaging; MD=Physician; RBC=Red blood cell; N/A=not applicable given lack of true negatives (i.e. every individual should theoretically have an admitting and discharging physician, making the number of true negatives equal to 0). N refers to the number of sampled admissions for that specific variable.

Note: Values are rounded to the nearest percentage point.

Appendix Table 2. Site A Manual Data Validation.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Data table | Variable (N) | N | Sensitivity | Specificity | PPV | NPV |
| Laboratory | Haemoglobin | 200 | 100% | 100% | 100% | 100% |
| Sodium | 100 | 100% | 100% | 100% | 100% |
| Creatinine | 200 | 100% | 100% | 100% | 100% |
| Calcium | 200 | 100% | 100% | 100% | 100% |
| AST | 200 | 100% | 100% | 100% | 100% |
| INR | 200 | 100% | 100% | 100% | 100% |
| Troponin | 200 | 100% | 100% | 100% | 100% |
| Radiology | X-ray | 200 | 100% | 100% | 100% | 100% |
| Ultrasound\* | 200 | 97% | 100% | 100% | 99% |
| CT scan | 200 | 100% | 100% | 100% | 100% |
| MRI | 200 | 100% | 100% | 100% | 100% |
| Echocardiography | 200 | 100% | 100% | 100% | 100% |
| Physicians | Admitting MD | 200 | 94% | N/A | 99% | N/A |
| Discharging MD | 200 | 93% | N/A | 99% | N/A |
| Death | In-hospital mortality | 800 | 100% | 100% | 100% | 100% |
| Transfer | Critical care unit transfer | 800 | 100% | 100% | 100% | 100% |
| Transfusion | RBC transfusion | 800 | 100% | 100% | 100% | 100% |

Appendix Table 3. Site B Manual Data Validation.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Data table | Variable | N | Sensitivity | Specificity | PPV | NPV |
| Laboratory | Haemoglobin | 200 | 100% | 100% | 100% | 100% |
| Sodium | 200 | 100% | 100% | 100% | 100% |
| Creatinine | 110 | 100% | 100% | 100% | 100% |
| Calcium | 110 | 100% | 100% | 100% | 100% |
| AST | 110 | 100% | 100% | 100% | 100% |
| INR | 110 | 100% | 100% | 100% | 100% |
| Troponin | 110 | 100% | 100% | 100% | 100% |
| Radiology | X-ray | 110 | 100% | 100% | 100% | 100% |
| Ultrasound | 200 | 100% | 100% | 100% | 100% |
| CT scan | 200 | 100% | 100% | 100% | 100% |
| MRI | 200 | 100% | 100% | 100% | 100% |
| Echocardiography | 110 | 100% | 100% | 100% | 100% |
| Physicians | Admitting MD | 200 | 100% | N/A | 100% | N/A |
| Discharging MD | 200 | 98% | N/A | 100% | N/A |
| Death | In-hospital mortality | 90 | 100% | 100% | 100% | 100% |
| Transfer | Critical care unit transfer | No data | | | | |
| Transfusion | RBC transfusion | 90 | 95% | 97% | 90% | 99% |

Appendix Table 4. Site C Manual Data Validation.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Data table | Variable | N | Sensitivity | Specificity | PPV | NPV |
| Laboratory | Haemoglobin | 280 | 100% | 100% | 100% | 100% |
| Sodium | 280 | 100% | 100% | 100% | 100% |
| Creatinine | 100 | 100% | 100% | 100% | 100% |
| Calcium | 100 | 100% | 100% | 100% | 100% |
| AST | 100 | 100% | 100% | 100% | 100% |
| INR | 100 | 100% | 100% | 100% | 100% |
| Troponin | 100 | 100% | 100% | 100% | 100% |
| Radiology | X-ray | 100 | 100% | 100% | 100% | 100% |
| Ultrasound\* | 280 | 100% | 100% | 98% | 100% |
| CT scan | 280 | 100% | 100% | 100% | 100% |
| MRI | 280 | 100% | 100% | 100% | 100% |
| Echocardiography | 100 | 100% | 100% | 100% | 100% |
| Physicians | Admitting MD | 280 | 100% | N/A | 100% | N/A |
| Discharging MD | 280 | 100% | N/A | 100% | N/A |
| Death | In-hospital mortality | 980 | 100% | 100% | 100% | 100% |
| Transfer | Critical care unit transfer | 800 | 100% | 99% | 91% | 100% |
| Transfusion | RBC transfusion\* | 982 | 100% | 100% | 97% | 100% |

Appendix Table 5. Site D Manual Data Validation.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Data table | Variable | N | Sensitivity | Specificity | PPV | NPV |
| Laboratory | Haemoglobin | 180 | 100% | 100% | 100% | 100% |
| Sodium | 180 | 100% | 100% | 100% | 100% |
| Creatinine | 100 | 100% | 100% | 100% | 100% |
| Calcium | 100 | 100% | 100% | 100% | 100% |
| AST | 100 | 100% | 100% | 100% | 100% |
| INR | 100 | 100% | 100% | 100% | 100% |
| Troponin | 100 | 100% | 100% | 100% | 100% |
| Radiology | X-ray | 100 | 100% | 100% | 100% | 100% |
| Ultrasound | 180 | 100% | 100% | 100% | 100% |
| CT scan | 180 | 100% | 100% | 100% | 100% |
| MRI | 180 | 100% | 100% | 100% | 100% |
| Echocardiography | 100 | \* | 100% | \* | 100% |
| Physicians | Admitting MD | 180 | 100% | N/A | 100% | N/A |
| Discharging MD | 180 | 95% | N/A | 100% | N/A |
| Death | In-hospital mortality | 280 | 100% | 100% | 100% | 100% |
| Transfer | Critical care unit transfer | 200 | 100% | 100% | 100% | 100% |
| Transfusion | RBC transfusion | 280 | 66% | 95% | 75% | 93% |

\*Cannot be calculated because no echocardiography tests done on this subset of patients.Appendix Table 6. Site E Manual Data Validation.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Data table | Variable | N | Sensitivity | Specificity | PPV | NPV |
| Laboratory | Haemoglobin | 169 | 100% | 100% | 100% | 100% |
| Sodium | 169 | 100% | 100% | 100% | 100% |
| Creatinine | 100 | 100% | 100% | 100% | 100% |
| Calcium | No data | | | | |
| AST | 50 | 100% | 100% | 100% | 100% |
| INR | 50 | 100% | 100% | 100% | 100% |
| Troponin | 50 | 100% | 100% | 100% | 100% |
| Radiology | X-ray | 100 | 100% | 100% | 100% | 100% |
| Ultrasound | 169 | 100% | 100% | 100% | 100% |
| CT scan | 169 | 100% | 100% | 100% | 100% |
| MRI | 169 | 100% | 100% | 100% | 100% |
| Echocardiography | 100 | 100% | 100% | 100% | 100% |
| Physicians | Admitting MD | 169 | 100% | N/A | 100% | N/A |
| Discharging MD | 169 | 98% | N/A | 99% | N/A |
| Death | In-hospital mortality | 769 | 100% | 100% | 100% | 100% |
| Transfer | Critical care unit transfer | 700 | 100% | 99% | 95% | 100% |
| Transfusion | RBC transfusion | 69 | 100% | 95% | 80% | 100% |

Appendix Table 7. Site F Manual Data Validation.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Data table | Variable | N | Sensitivity | Specificity | PPV | NPV |
| Laboratory | Haemoglobin | 195 | 100% | 100% | 100% | 100% |
| Sodium | 195 | 100% | 100% | 100% | 100% |
| Creatinine | 100 | 100% | 100% | 100% | 100% |
| Calcium | 100 | 100% | 100% | 100% | 100% |
| AST | 100 | 100% | 100% | 100% | 100% |
| INR | 100 | 100% | 100% | 100% | 100% |
| Troponin | 100 | 100% | 100% | 100% | 100% |
| Radiology | X-ray | 100 | 100% | 98% | 98% | 100% |
| Ultrasound | 195 | 97% | 100% | 100% | 99% |
| CT scan | 195 | 100% | 99% | 97% | 100% |
| MRI | 195 | 100% | 100% | 100% | 100% |
| Echocardiography | 100 | 100% | 100% | 100% | 100% |
| Physicians | Admitting MD | 195 | 100% | N/A | 100% | N/A |
| Discharging MD | 195 | 98% | N/A | 99% | N/A |
| Death | In-hospital mortality | 895 | 100% | 100% | 100% | 100% |
| Transfer | Critical care unit transfer | 800 | 100% | 100% | 98% | 100% |
| Transfusion | RBC transfusion | 895 | 96% | 99% | 92% | 99% |