

## IMPUTATION OF MISSING ISSUED COUNTS

**Table S1** outlines the year and starting variables theorised to be associated with the data to be imputed (following Van Buuren, Boshuizen and Knook<sup>1</sup>).

Since all requests with missing issued counts in 2019 were associated with a patient type of public, patient type was removed as a possible predictor. Furthermore, in 2015, 2016 and 2018, all patients with a patient type unspecified were not associated with any missing issued counts. Therefore, only public and private associated requests were included in the set. The *quickpred()* function in the MICE package<sup>[2]</sup> in R statistical software<sup>[3]</sup> was applied to the dataset containing the starting variables. This function proposed the variables that were the most probable predictors for the missing values. In all cases (except 2019) this resulted in ordered count, cancel count and patient admission class being said to be the best predictors for the missing issued counts. The predictor matrix determined by *quickpred()* was then used within the *mice()* function in the MICE package<sup>[2]</sup> with the random forest imputation method and 5 imputations with 5 iterations were generated for each missing value. Considering the general relationship between the observed issued counts and the difference between the ordered and cancel count indicated that a linear model would be a reasonable choice in the imputation model. The estimates generated from the imputations were pooled and the pooled coefficients were used to generate the total predicted set of issued counts. Thereafter, only the missing issued counts from the original dataset was replaced.

**Table S1. Theorised variables necessary for imputation of missing issued counts in dataset**

Year(s)	Theorised Associated Variables
2014 and 2017	Hospital Patient type Ordered product Ordered count Cancel count
2019	Hospital Ordered product Ordered count Cancel count
2015, 2016 and 2018	Patient type (excluding unspecified) Hospital Ordered product Ordered count Cancel count

The validity of the imputations was tested on two levels. Firstly, on a function-specific level – ensuring that each step of the algorithm performs well and generates the expected outcomes. Secondly, the predicted issued counts of the observed issued counts were compared to their actual values to ensure reasonable

imputations. Furthermore, the validity of the number of imputations were investigated. The default number of imputations, as well as number of iterations, namely 5, was selected for the multiple imputation of the missing issued counts. To ascertain whether an increase in the number of multiple imputations would result in a difference in the mean issued counts (following Scheffer<sup>4</sup>) the following investigation was conducted:

Two years, namely 2014 and 2017, were selected for the investigation since these years followed the same imputation model and were associated with a different percentage of missing issued counts. Varying the number of multiple imputations ( $m = 5, 10$  and  $20$ ) and for 2014, further adjusting the number of iterations for a fixed number of multiple imputations (namely 5), it was investigated whether a difference to the mean issued counts of the completed sets would result. In both cases it was found that no change came about in either the mean or standard deviation of the issued counts in the completed sets with either increasing the number of multiple imputations or the number of iterations. Therefore, keeping the number of multiple imputations and number of iterations as 5, the results obtained for increasing number of missing issued counts are outlined (Table S2).

**Table S2. Percentage change in mean issued count with change in percentage of missing values**

Year	Total issued counts, n	Missing issued counts, n (%)	Mean issued count prior to imputation	Standard deviation prior to imputation	Mean issued count after imputation	Standard deviation after imputation	Change in mean issued count (%)
2014	639 347	95 (0.0149)	1.500053	1.396429	1.499841	1.396439	0.0141
2017	672 853	75 (0.0111)	1.477536	1.404574	1.477378	1.404579	0.0107

Imputing the missing issued counts results in a decrease in the mean issued counts for 2014 and 2017 of 0.0141 and 0.0107 %, respectively. Furthermore, increasing the missing issued counts from 0.0111% to 0.0149 % results in a 1.5 % increase in the mean issued count and a decrease in the standard deviation of 0.58 %.

## TRANSFUSION EVENTS DEMOGRAPHICS DISTRIBUTION

Table S3. Public and private RBC transfusion events demographic distribution, 2014-2019\*

Variable	Public healthcare ( <i>n</i> = 1 553 159)	Private healthcare ( <i>n</i> = 803 282)
Male	520 772 (33.5)	373 942 (46.6)
0-4	94 763	29 647
5-9	13 693	3 354
10-14	12 667	3 756
15-19	14 955	4 908
20-24	26 859	6 349
25-29	37 903	9 513
30-34	46 477	14 721
35-39	41 439	18 806
40-44	37 009	23 806
45-49	32 641	28 113
50-54	31 019	34 878
55-59	31 501	39 473
60-64	30 960	36 445
65-69	25 333	34 524
70-74	17 510	31 921
75-79	12 005	26 170
80 +	9 970	27 451
Not allocated	4 068	107

\*Data reported as number of transfusion events (%)

## ACKNOWLEDGEMENT

This work is based on research supported by the Department of Science and Innovation and the National Research Foundation. Any opinion, finding, and conclusion or recommendation expressed in this material is that of the authors and the NRF does not accept any liability in this regard.

## REFERENCES

1. Van Buuren S, Boshuizen HC, Knook DL. Multiple imputation of missing blood pressure covariates in survival analysis. *Stat. Med.* 1999;18(6):681–694. [https://doi.org/10.1002/\(SICI\)1097-0258\(19990330\)18:6%3C681::AID-SIM71%3E3.0.CO;2-R](https://doi.org/10.1002/(SICI)1097-0258(19990330)18:6%3C681::AID-SIM71%3E3.0.CO;2-R)
2. Groothuis-Oudshoorn K, van Buuren S. mice : Multivariate Imputation by Chained. *J. Stat. Softw.* 2011;45(3):1–67. <http://dx.doi.org/10.18637/jss.v045.i03>
3. R Core Team. R: A language environment for statistical computing, 2019. <https://www.r-project.org/>
4. Scheffer J. Dealing with Missing Data. *Res. Lett. Inf. Math. Sci.* 2002; 3:153–160. <http://hdl.handle.net/10179/4355>