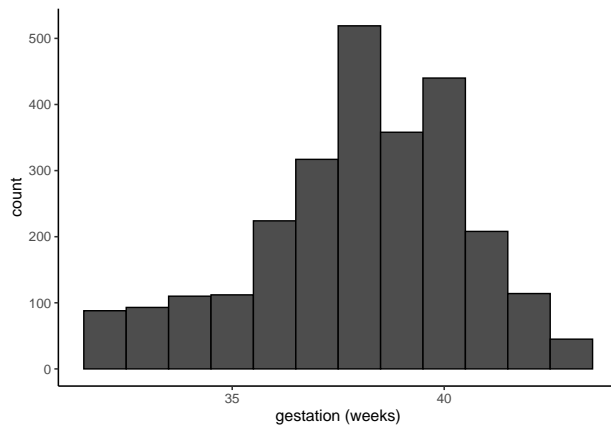


7-DescriptiveStatistics

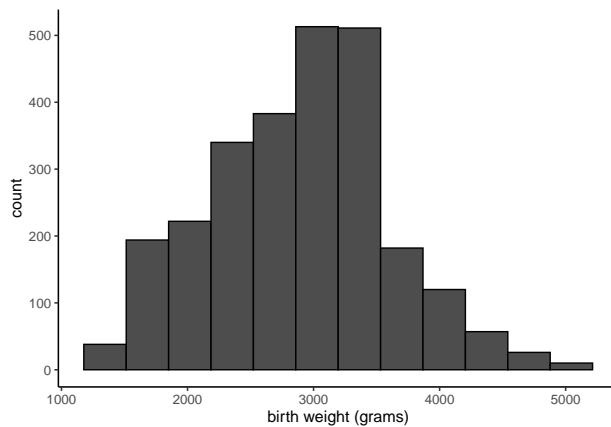
Descriptive analysis of included participants. Data are presented for the observed data only (i.e. before MICE) using pairwise deletion of missing values.

0.1 Distribution of continuous variables

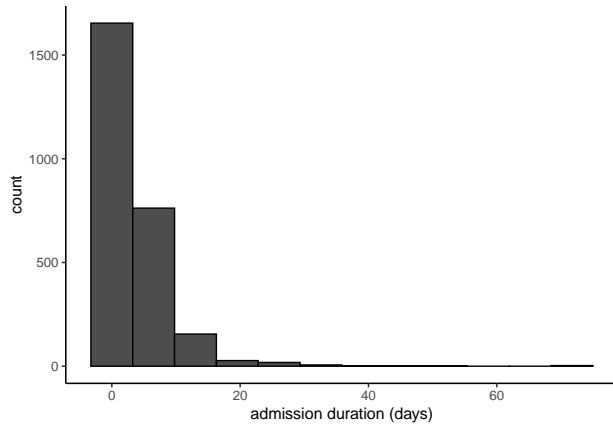
\$pi_gest



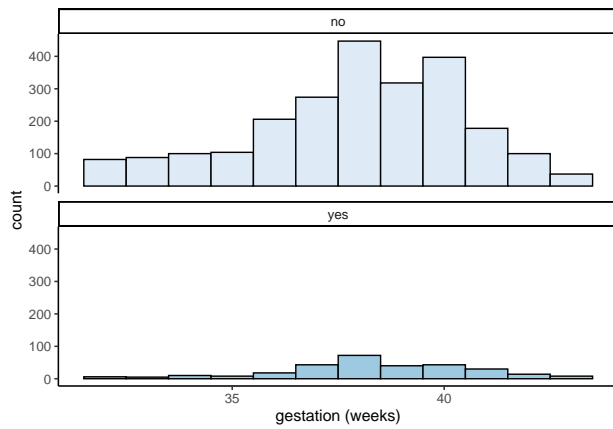
\$et_bw



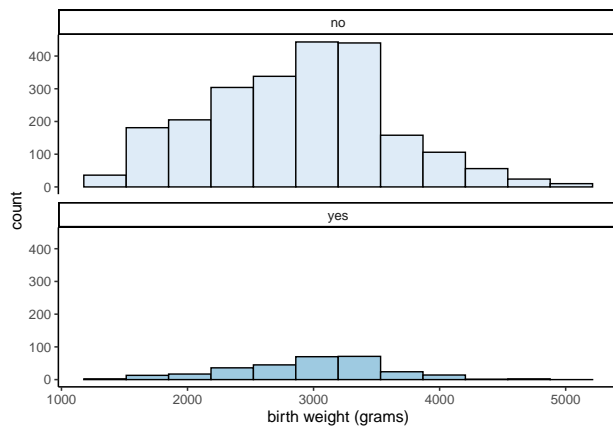
\$adm_dur



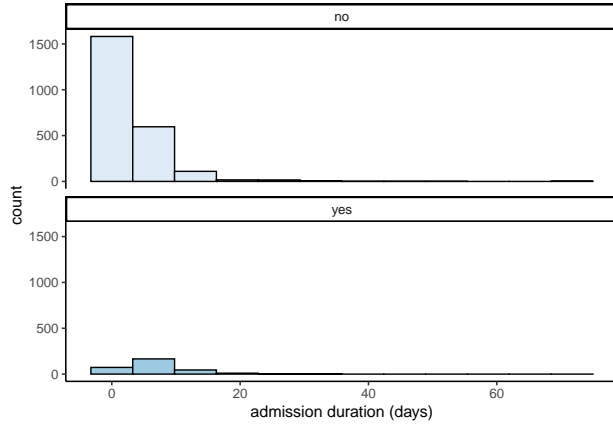
\$pi_gest



\$et_bw



\$adm_dur



“Gestation” and “birth weight” are approximately normally distributed, while “admission duration” is very right-skewed.

0.2 Table 1

Table summarising the characteristics of included participants:

Characteristic	Overall, N = 2,628	no, N = 2,331	yes, N = 297	p-value
Sex, n (%)				0.7
f	1,122 (43%)	990 (42%)	132 (44%)	
m	1,503 (57%)	1,338 (57%)	165 (56%)	
u	3 (0.1%)	3 (0.1%)	0 (0%)	
Gestational age, mean weeks (SD)	38.00 (2.50)	37.96 (2.52)	38.36 (2.29)	0.005
Birth weight, mean grams (SD)	2,889 (703)	2,881 (716)	2,950 (595)	0.067
Chronological age, n (%)				<0.001
fnb	1,001 (38%)	901 (39%)	100 (34%)	
dol1	1,257 (48%)	1,136 (49%)	121 (41%)	
dol2	235 (9.0%)	181 (7.8%)	54 (18%)	
dol3	110 (4.2%)	91 (3.9%)	19 (6.5%)	
Type of birth, n (%)				0.032
singleton	2,496 (95%)	2,205 (95%)	291 (98%)	
twin1	127 (4.8%)	121 (5.2%)	6 (2.0%)	
triplet1	2 (<0.1%)	2 (<0.1%)	0 (0%)	
Mode of delivery, n (%)				0.074
svd	1,889 (72%)	1,663 (71%)	226 (76%)	
electiveCS	136 (5.2%)	124 (5.3%)	12 (4.0%)	
emergencyCS	561 (21%)	510 (22%)	51 (17%)	
instrumental	42 (1.6%)	34 (1.5%)	8 (2.7%)	
Admission duration, median days [Q1-Q3]	2.3 [1.3-4.9]	2.1 [1.2-4.1]	6.0 [3.5-8.8]	<0.001
Death, n (%)	221 (8.4%)	184 (7.9%)	37 (12%)	0.008

p-values are from Welch’s two-sample t-test for gestational age and birth weight; the Wilcoxon-Mann-Whitney U test for admission duration; Pearson’s chi-squared test for age at admission and death; and Fisher’s exact test for sex, type of birth and mode of delivery.

Data are presented for the observed data only (i.e. before MICE) using pairwise deletion of missing values. The number of missing values for each variable in the above table are as follows:

A tibble: 9 x 3

```
##   variable      n_miss pct_miss
##   <chr>         <int>   <dbl>
## 1 et_bw         32      1.22
## 2 pi_age        25      0.951
## 3 pi_type        3      0.114
## 4 pi_sex         0       0
## 5 pi_gest        0       0
## 6 oh_delivery    0       0
## 7 adm_dur        0       0
## 8 outcome        0       0
## 9 sepsis         0       0
```

0.3 Distribution of candidate predictors

Characteristic	Overall, N = 2,628	no, N = 2,331	yes, N = 297	p-value ¹
pi_gest	38.00 [37.00-40.00]	38.00 [37.00-40.00]	38.00 [37.00-40.00]	0.032
et_bw	2,950 [2,400-3,350]	2,900 [2,400-3,350]	3,000 [2,600-3,350]	0.035
oh_matfever	14 (0.5%)	8 (0.3%)	6 (2.0%)	0.003
oh_offliquor	163 (6.2%)	131 (5.6%)	32 (11%)	0.001
co_prom	303 (12%)	257 (11%)	46 (15%)	0.027
et_grunt	750 (29%)	654 (28%)	96 (32%)	0.13
et_cyanosis	69 (2.6%)	60 (2.6%)	9 (3.0%)	0.6
et_seizures	14 (0.5%)	10 (0.4%)	4 (1.3%)	0.064
et_rr	56 [48-68]	56 [48-68]	60 [50-72]	<0.001
et_hr	138 [126-146]	138 [126-146]	139 [127-150]	0.011
et_temp	36.50 [36.00-37.00]	36.50 [36.00-36.90]	36.90 [36.20-38.00]	<0.001
oe_fontanelle				0.9
flat	2,608 (99%)	2,312 (99%)	296 (100%)	
sunken	10 (0.4%)	9 (0.4%)	1 (0.3%)	
bulging	10 (0.4%)	10 (0.4%)	0 (0%)	
oe_activity				<0.001
alert	2,152 (82%)	1,933 (83%)	219 (74%)	
lethargic	382 (15%)	327 (14%)	55 (19%)	
irritable	62 (2.4%)	45 (1.9%)	17 (5.7%)	
seizures	14 (0.5%)	9 (0.4%)	5 (1.7%)	
coma	18 (0.7%)	17 (0.7%)	1 (0.3%)	
oe_nasalflare	912 (35%)	791 (34%)	121 (41%)	0.023
oe_retractions	986 (38%)	848 (36%)	138 (46%)	<0.001
oe_grunt	421 (16%)	360 (15%)	61 (21%)	0.029
oe_wob				<0.001
normal	1,405 (54%)	1,263 (55%)	142 (48%)	
mild	413 (16%)	378 (16%)	35 (12%)	
moderate	614 (24%)	529 (23%)	85 (29%)	
severe	170 (6.5%)	139 (6.0%)	31 (11%)	
oe_colour				0.11
pink	2,507 (95%)	2,220 (95%)	287 (97%)	
pale	10 (0.4%)	7 (0.3%)	3 (1.0%)	
blue	62 (2.4%)	58 (2.5%)	4 (1.3%)	
yellow	49 (1.9%)	46 (2.0%)	3 (1.0%)	
oe_abdodist	28 (1.1%)	26 (1.1%)	2 (0.7%)	0.8
oe_omphalitis	6 (0.2%)	4 (0.2%)	2 (0.7%)	0.14
oe_abskin	27 (1.0%)	23 (1.0%)	4 (1.3%)	0.5

hx_vomit				0.3
no	2,605 (99%)	2,309 (99%)	296 (100%)	
yellow	7 (0.3%)	7 (0.3%)	0 (0%)	
bilious	13 (0.5%)	13 (0.6%)	0 (0%)	
bloody	3 (0.1%)	2 (<0.1%)	1 (0.3%)	

¹Data are presented as median [Q1-Q3] for continuous predictors or n (%) for categorical predictors. p-values are from the Wilcoxon-Mann-Whitney U test for continuous predictors and Fisher's exact test for categorical predictors.

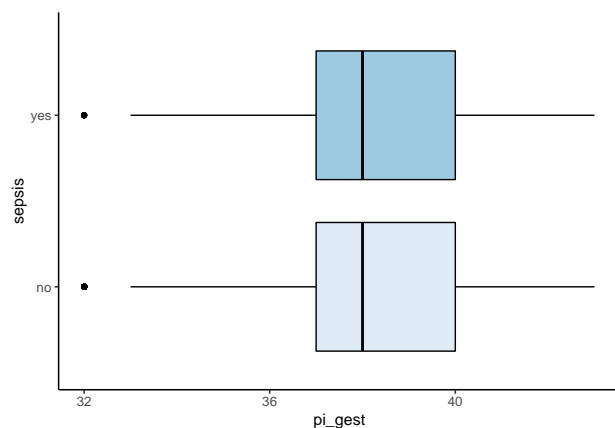
Data are presented for the observed data only (i.e. before MICE) using pairwise deletion of missing values. The number of missing values for each variable in the above table are as follows:

```
## # A tibble: 22 x 3
##   variable      n_miss pct_miss
##   <chr>        <int>    <dbl>
## 1 et_temp         814    31.0
## 2 et_bw           32     1.22
## 3 oe_wob          26     0.989
## 4 et_rr           22     0.837
## 5 et_hr            3     0.114
## 6 oe_abdodist      2     0.0761
## 7 oh_matfever       1     0.0381
## 8 oh_offliquor       1     0.0381
## 9 et_grunt          1     0.0381
## 10 et_cyanosis       1     0.0381
## # ... with 12 more rows
```

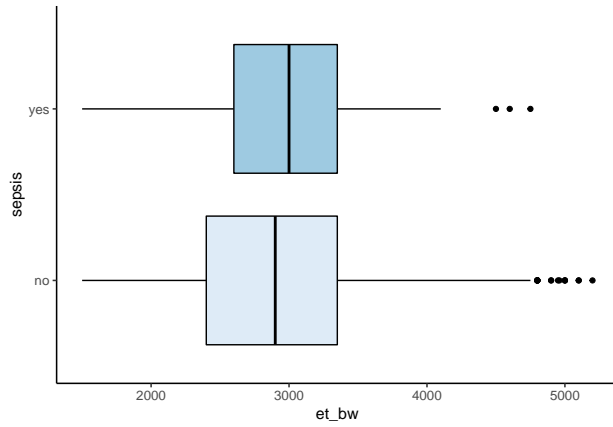
0.3.1 Box plots of continuous candidate predictors

The box plots below show the distribution of the continuous candidate predictors between participants with and without sepsis.

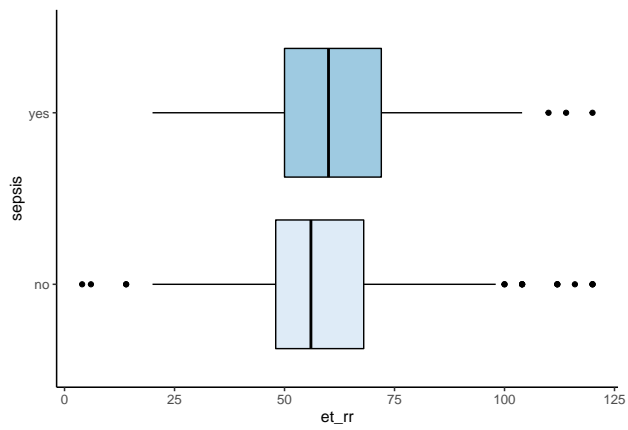
```
## $pi_gest
```



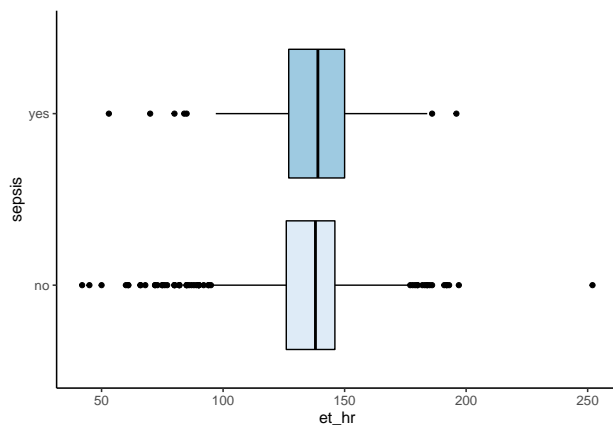
```
##
## $et_bw
```



```
##
## $et_rr
```



```
##
## $et_hr
```



```
##
## $et_temp
```

