LOCATE: descriptive statistics

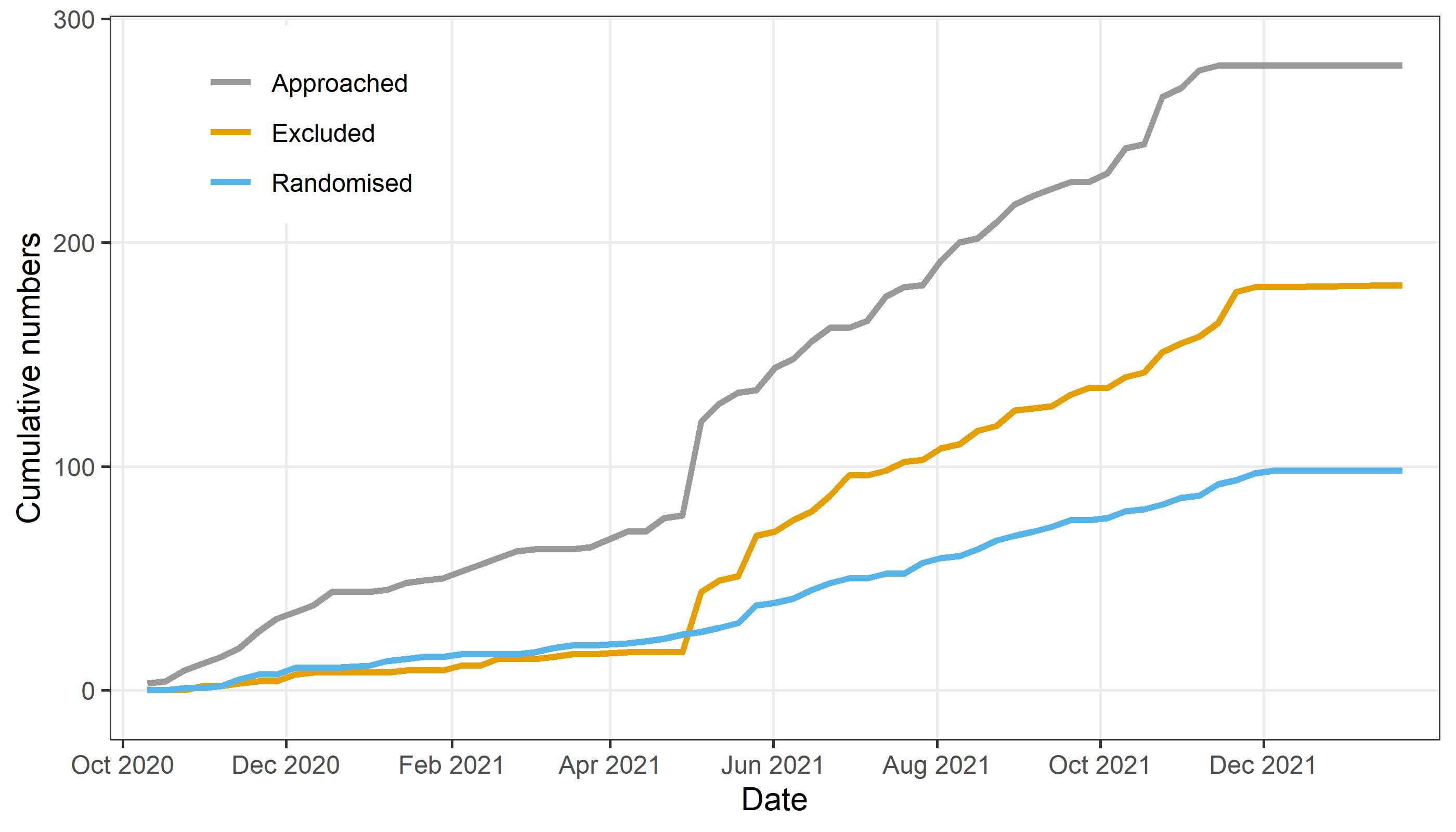
Adrian Barnett

25 May, 2023

# Recruitment

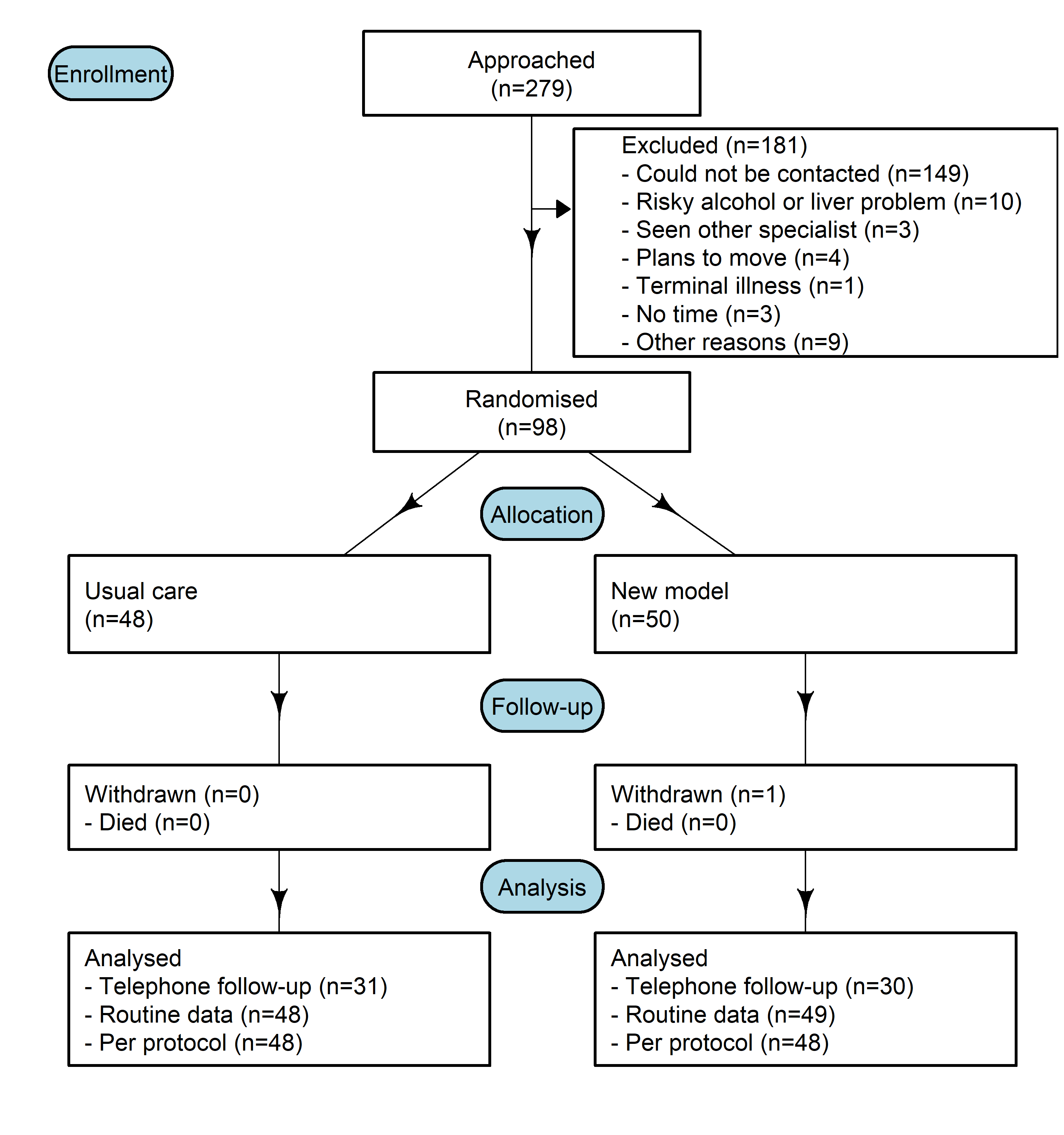
The target sample size was 156 participants, the final recruitment was 98 participants, hence we were 37% short of our target.

## Cumulative participant numbers over time



The final numbers are: 279 approached, 181 excluded and 98 randomised. The first patient was randomised on 27-Oct-2020 and the last on 06-Dec-2021.

## CONSORT flow diagram



There were no participants who were pregnant (this was an exclusion criterion).

All but one patient randomised to the scan completed their scan, hence the per protocol analysis will be almost identical to the intention-to-treat analysis. This means a per protocol analysis will not likely be needed.

# Summary statistics

Results from now on are only for the 98 randomised patients.

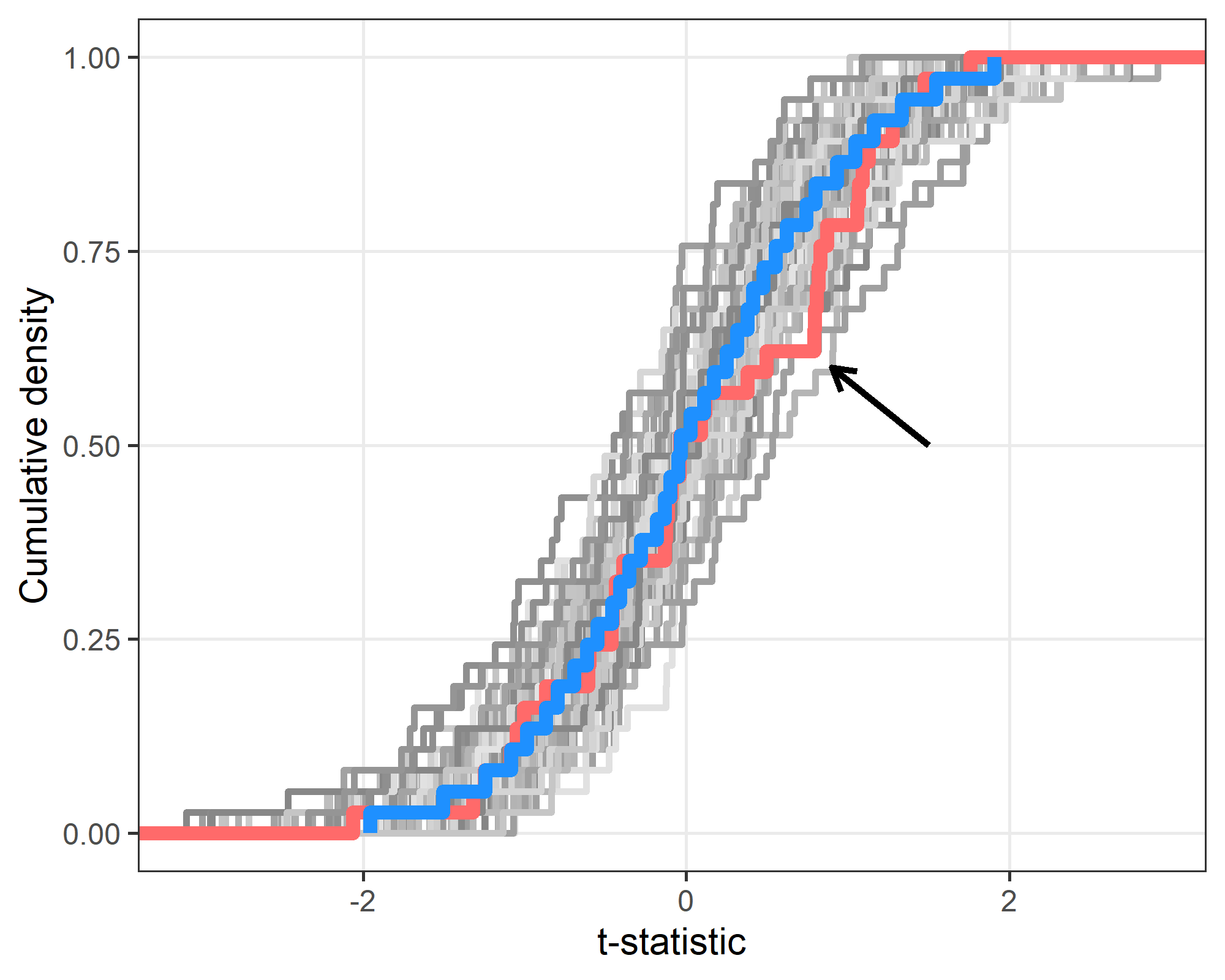
## Baseline summary table

Stratified by randomised  
 level New model of care Usual care   
 n 50 48   
 centrename (%) Metro South 35 (70.0) 34 (70.8)   
 Sunshine Coast 15 (30.0) 14 (29.2)   
 re\_referral (%) First referral 43 (86.0) 44 (91.7)   
 Re-referral 5 (10.0) 2 ( 4.2)   
 <NA> 2 ( 4.0) 2 ( 4.2)   
 referral\_cat (%) One 2 ( 4.0) 1 ( 2.1)   
 Two 44 (88.0) 44 (91.7)   
 Three 4 ( 8.0) 2 ( 4.2)   
 Unknown 0 ( 0.0) 1 ( 2.1)   
 age (mean (SD)) 55.7 (14.4) 53.3 (15.8)   
 sex (%) Female 28 (56.0) 32 (66.7)   
 Male 22 (44.0) 16 (33.3)   
 employment (%) Carer 1 ( 2.0) 3 ( 6.2)   
 Disability pension 1 ( 2.0) 6 (12.5)   
 Full-time 14 (28.0) 10 (20.8)   
 Other (please describe below) 5 (10.0) 6 (12.5)   
 Part-time 4 ( 8.0) 8 (16.7)   
 Retired 17 (34.0) 10 (20.8)   
 Unemployed 8 (16.0) 5 (10.4)   
 height (mean (SD)) 168.7 (11.6) 165.9 (10.9)   
 weight (mean (SD)) 97.8 (24.0) 98.5 (22.0)   
 bmi (mean (SD)) 34.9 (7.8) 35.9 (10.9)   
 ast (mean (SD)) 44.4 (30.9) 49.9 (54.8)   
 alt (mean (SD)) 67.5 (44.4) 68.4 (58.2)   
 albumin (mean (SD)) 44.0 (6.0) 42.2 (3.7)   
 glucose (mean (SD)) 6.2 (1.6) 7.6 (7.3)   
 hba1c (mean (SD)) 6.4 (1.2) 6.8 (1.6)   
 hbsag (%) Positive 0 ( 0.0) 3 ( 6.2)   
 Negative 38 (76.0) 33 (68.8)   
 <NA> 12 (24.0) 12 (25.0)   
 hepc\_igg (%) Positive 0 ( 0.0) 2 ( 4.2)   
 Negative 38 (76.0) 36 (75.0)   
 <NA> 12 (24.0) 10 (20.8)   
 platelet\_count (mean (SD)) 262.5 (56.5) 275.5 (64.4)   
 ferritin (mean (SD)) 194.3 (250.2) 196.8 (235.6)  
 transferrin (mean (SD)) 32.9 (22.1) 28.2 (8.9)   
 alcohol (%) Non-drinker 22 (44.0) 19 (39.6)   
 Moderate drinker 8 (16.0) 11 (22.9)   
 Not mentioned 5 (10.0) 9 (18.8)   
 Other 7 (14.0) 3 ( 6.2)   
 <NA> 8 (16.0) 6 (12.5)   
 audit (mean (SD)) 1.6 (2.0) 1.8 (2.1)   
 eq5d\_b (mean (SD)) 72.4 (24.5) 67.9 (26.8)   
 dietician\_baseline (%) No 32 (64.0) 33 (68.8)   
 Yes 15 (30.0) 15 (31.2)   
 <NA> 3 ( 6.0) 0 ( 0.0)

The table shows numbers and percentages in round brackets for categorical variables, and the mean and standard deviation for continuous variables. We do not compare the two randomised groups using statistical tests as this is not good practice (Altman 1985). Instead we present a graphical summary of the differences between groups.

### Test of randomisation

The plot below is an overall test of the differences at baseline between the randomised groups (Barnett 2022). It examines the distribution of t-statistics and checks if it is over- or under-dispersed compared the expected distribution when the two groups were randomised.



The plot shows the cumulative distribution function (CDF) of t-statistics. The trial CDF is in red and 100 simulated trial CDFs in grey. The simulated trials were generated following the null hypothesis of no dispersion. A median of the simulations is in blue. If the groups in the trial were correctly randomised, then the trial CDF should be similar to the CDFs from the simulated data. If the trial CDF is outside the simulated CDFs then this will indicate if the trial summary statistics are under- or over-dispersed.

The plot shows no strong concerns about the trial. There is a small shortage of observed t-statistics between around 0.3 to 0.8 (highlighted by arrow), but this is a moderate departure from the expected distribution.

## Ultrasound (from referral letter at baseline)

Stratified by randomised  
 level New model of care Usual care   
 n 50 48   
 spleen (%) Yes 2 ( 4.0) 1 ( 2.1)   
 No 33 (66.0) 36 (75.0)   
 Unclear 10 (20.0) 3 ( 6.2)   
 <NA> 5 (10.0) 8 (16.7)   
 fatty\_liver (%) Yes 41 (82.0) 39 (81.2)   
 No 2 ( 4.0) 0 ( 0.0)   
 Unclear 2 ( 4.0) 1 ( 2.1)   
 <NA> 5 (10.0) 8 (16.7)   
 liver\_outline (%) Nodular 1 ( 2.0) 1 ( 2.1)   
 Smooth 8 (16.0) 9 (18.8)   
 Not stated 34 (68.0) 29 (60.4)   
 <NA> 7 (14.0) 9 (18.8)   
 liver\_size (mean (SD)) 16.3 (2.4) 18.1 (3.5)

The table shows numbers and percentages in round brackets.

## GP referral (from referral letter at baseline)

Stratified by randomised  
 level New model of care Usual care   
 n 50 48   
 re\_referral (%) First referral 43 (86.0) 44 (91.7)   
 Re-referral 5 (10.0) 2 ( 4.2)   
 <NA> 2 ( 4.0) 2 ( 4.2)   
 referral\_cat (%) One 2 ( 4.0) 1 ( 2.1)   
 Two 44 (88.0) 44 (91.7)   
 Three 4 ( 8.0) 2 ( 4.2)   
 Unknown 0 ( 0.0) 1 ( 2.1)   
 referral\_time (mean (SD)) 56.4 (34.6) 52.8 (29.4)   
 previous\_hepatology (%) Yes 47 (94.0) 45 (93.8)   
 <NA> 3 ( 6.0) 3 ( 6.2)

The table shows numbers and percentages in round brackets for categorical variables, and the mean and standard deviation for the continuous variable. referral\_time is the time in days between the GP referral and the randomisation into the study.

Every patient had a mention of previous evaluation in a specialist hepatology clinic in the previous 12 months?

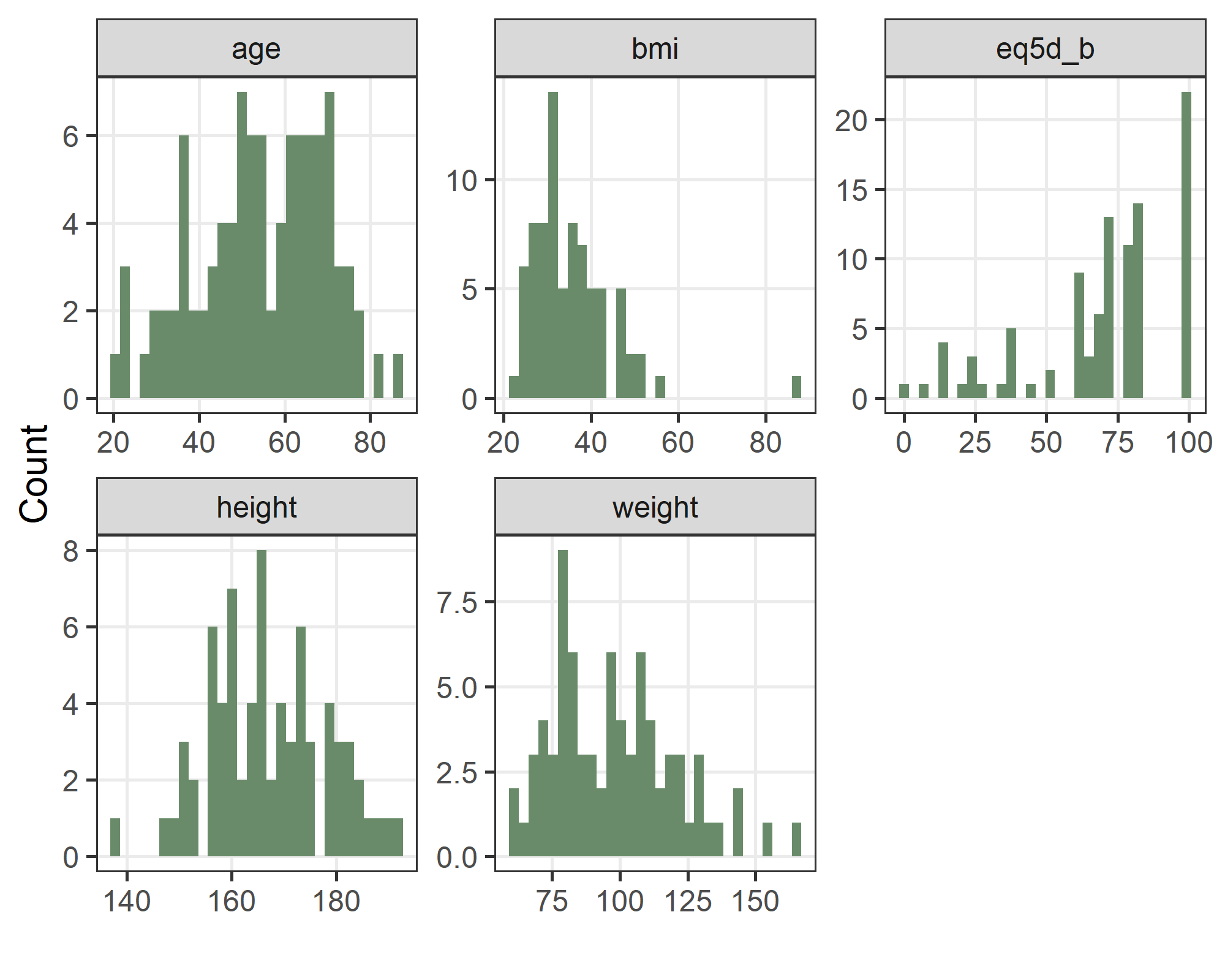
If “yes” do not approach this participant

## Was the information dated within six months of the referral date?

| **Variable** | **No** | **Yes** | **Missing** |
| --- | --- | --- | --- |
| Height/Weight/BMI | 68 (69) | 16 (16) | 14 (14) |
| Full blood count | 79 (81) | 10 (10) | 9 (9) |
| Hepatitis blood tests | 64 (65) | 16 (16) | 18 (18) |
| Iron studies | 48 (49) | 14 (14) | 36 (37) |
| Liver function tests | 88 (90) | 6 (6) | 4 (4) |
| Ultrasound | 74 (76) | 10 (10) | 14 (14) |

The table shows the counts and percentages in round brackets. Much of the information is older than six months.

## Histograms of continuous variables



eq5d\_b is EQ-5D at baseline. Height and weight are from the GP referral letter.

# Scan data

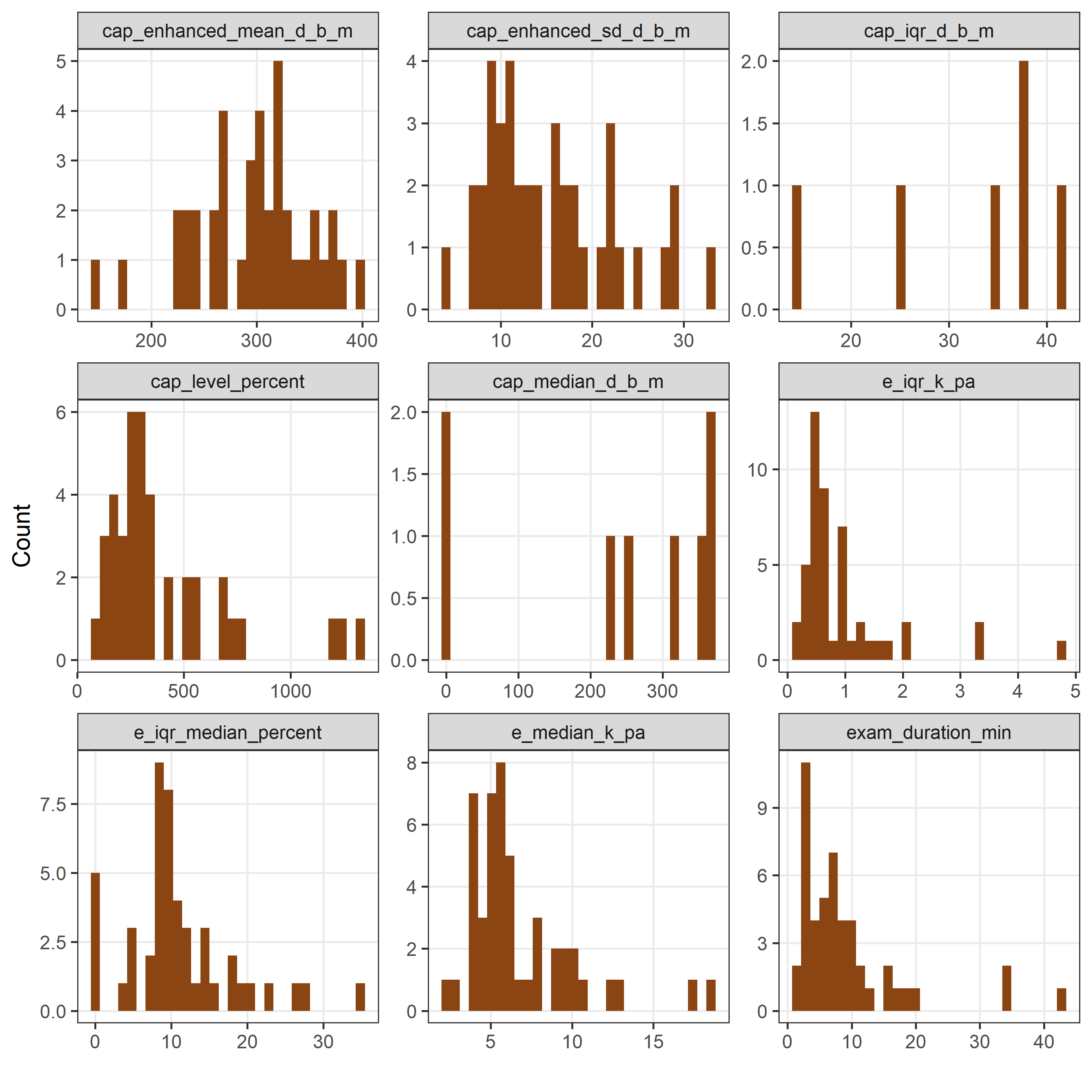
This section gives summary statistics on the Fibroscan data for the new model of care group.

## Frequencies for categorical variables

| **Variable** | **value** | **n** | **percent** |
| --- | --- | --- | --- |
| fasting | Yes | 37 | 77 |
| Missing | 8 | 17 |
| No | 3 | 6 |
| probe size | M | 24 | 50 |
| XL | 24 | 50 |

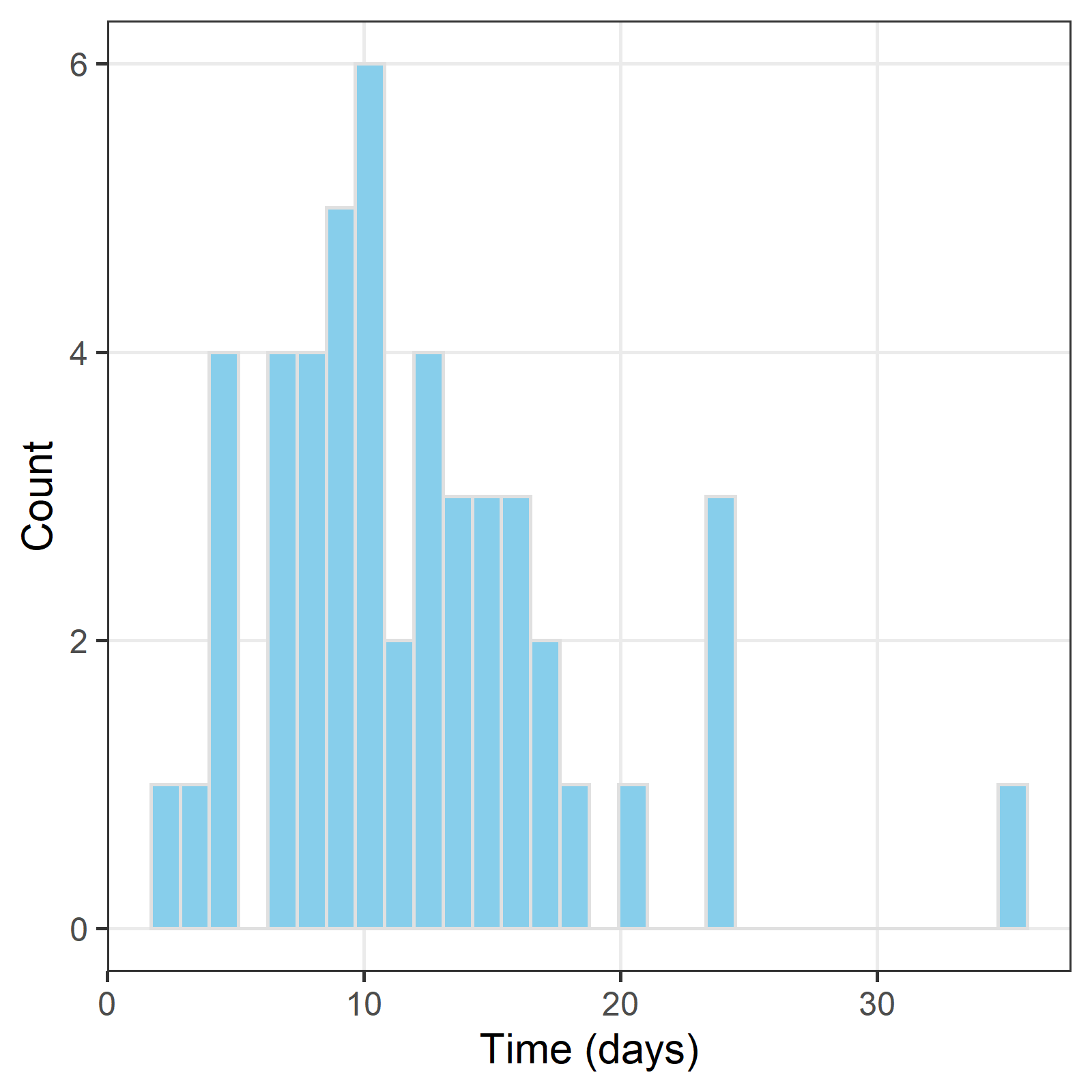
The table shows frequencies and percentages. Results are combined across the two randomised groups.

## Histograms of continuous variables from scans



## Time from randomisation to exam

#### Histogram of times



*This is only for participants randomised to the new model of care.*

#### Summary statistics of times

| **n** | **Missing** | **Min** | **Q1** | **Median** | **Q3** | **Max** |
| --- | --- | --- | --- | --- | --- | --- |
| 50 | 2 | 2 | 8 | 10 | 15 | 35 |

The statistics are in days.

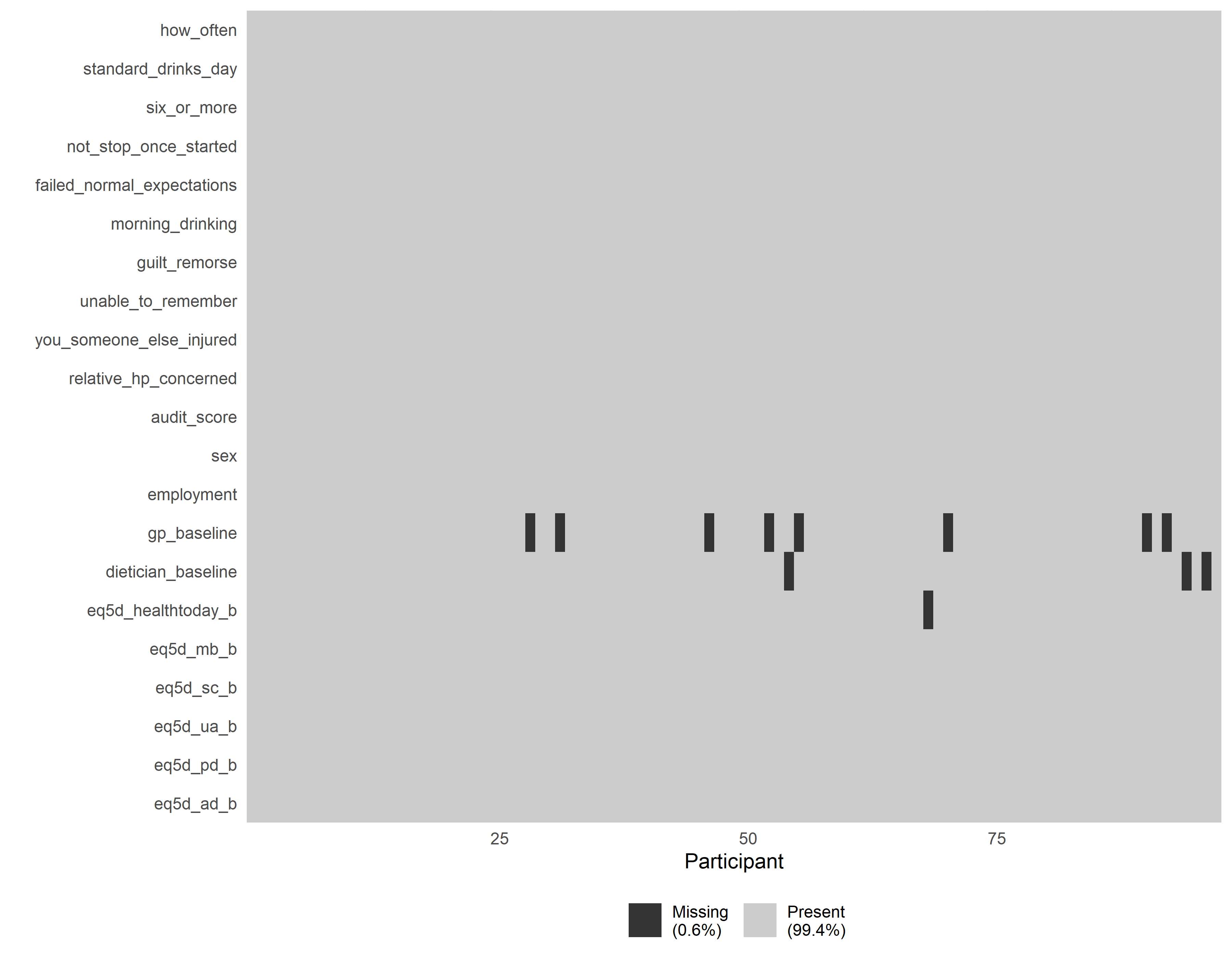
# Notes data

In this section examine the data extracted from the patients’ notes at the one year follow-up.

Stratified by randomised  
 level New model of care Usual care   
 n 50 48   
 appointment\_booked (%) No 6 ( 12.0) 9 ( 18.8)   
 Yes, completed 40 ( 80.0) 36 ( 75.0)   
 Yes, did not attend 3 ( 6.0) 3 ( 6.2)   
 <NA> 1 ( 2.0) 0 ( 0.0)   
 additional\_appointments (%) Yes 34 ( 68.0) 37 ( 77.1)   
 <NA> 16 ( 32.0) 11 ( 22.9)   
 fibroscan (%) Yes (successful) 35 ( 70.0) 37 ( 77.1)   
 Yes (not successful) 0 ( 0.0) 1 ( 2.1)   
 No 14 ( 28.0) 10 ( 20.8)   
 <NA> 1 ( 2.0) 0 ( 0.0)   
 ed\_presentations (%) Yes - once 12 ( 24.0) 8 ( 16.7)   
 Yes - more than once 3 ( 6.0) 5 ( 10.4)   
 No 34 ( 68.0) 34 ( 70.8)   
 <NA> 1 ( 2.0) 1 ( 2.1)   
 hcc (%) Yes 1 ( 2.0) 0 ( 0.0)   
 No 48 ( 96.0) 48 (100.0)   
 <NA> 1 ( 2.0) 0 ( 0.0)   
 variceal\_bleeding (%) Yes 0 ( 0.0) 2 ( 4.2)   
 No 49 ( 98.0) 46 ( 95.8)   
 <NA> 1 ( 2.0) 0 ( 0.0)   
 other\_diagnosis (%) Yes 3 ( 6.0) 1 ( 2.1)   
 No 46 ( 92.0) 47 ( 97.9)   
 <NA> 1 ( 2.0) 0 ( 0.0)   
 medications (%) Yes 4 ( 8.0) 3 ( 6.2)   
 No 43 ( 86.0) 45 ( 93.8)   
 Unclear 2 ( 4.0) 0 ( 0.0)   
 <NA> 1 ( 2.0) 0 ( 0.0)   
 studies (%) Yes 43 ( 86.0) 43 ( 89.6)   
 <NA> 7 ( 14.0) 5 ( 10.4)   
 Dead (%) No 50 (100.0) 48 (100.0)

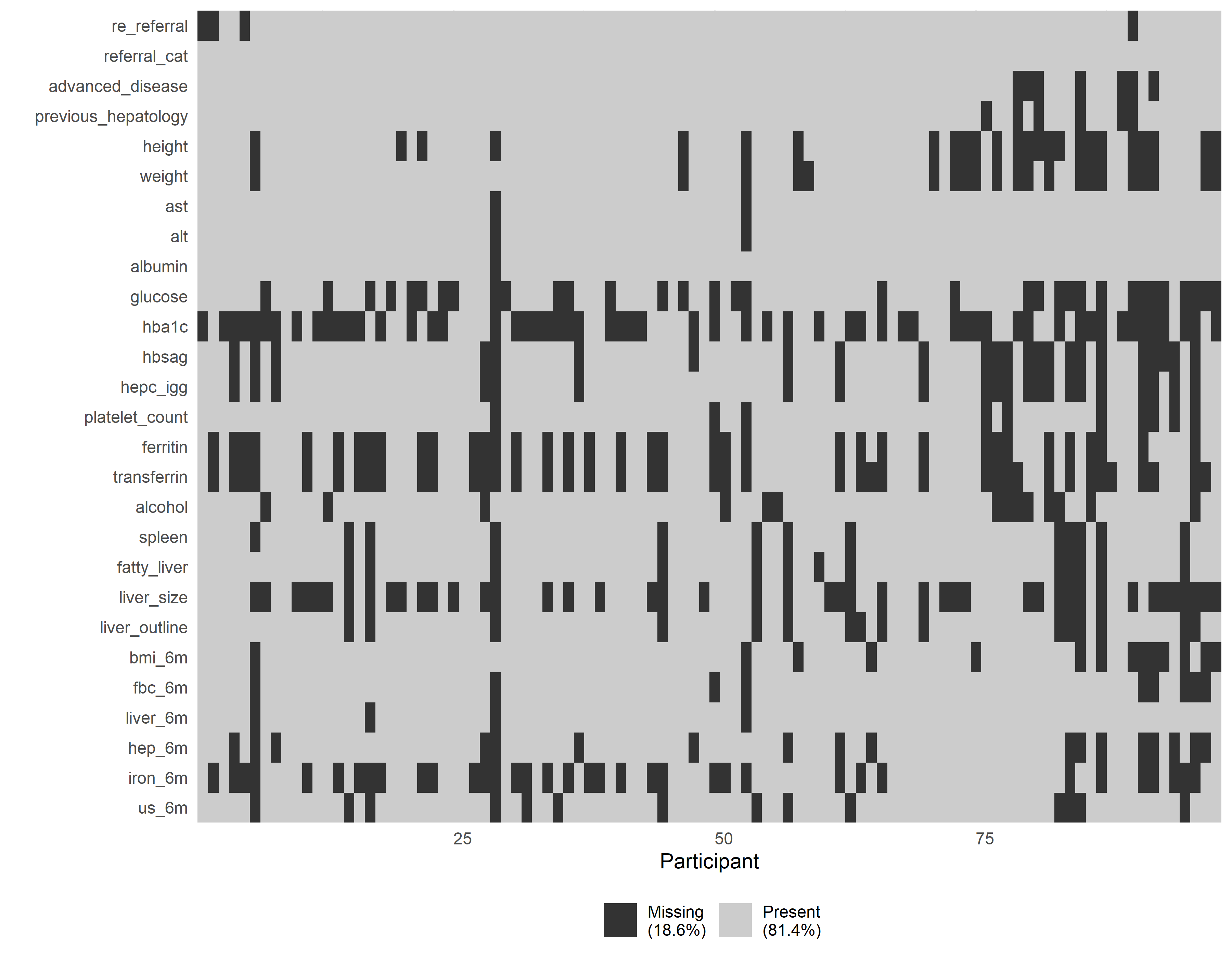
# Missing data

## Missing data in baseline questionnaire



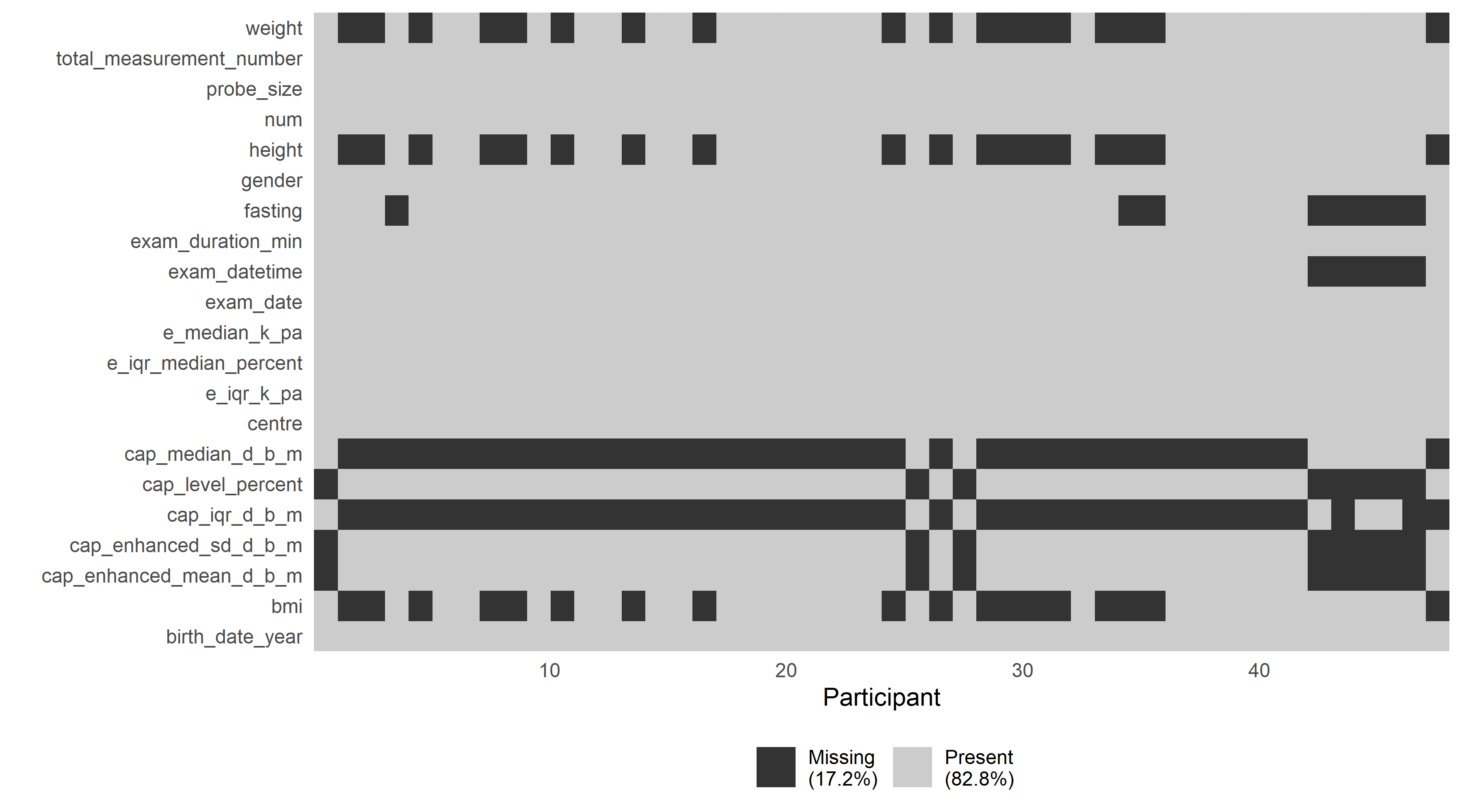
There was very little missing data in the baseline questionnaire. This questionnaire was completed by the participants.

## Missing data in referral letter

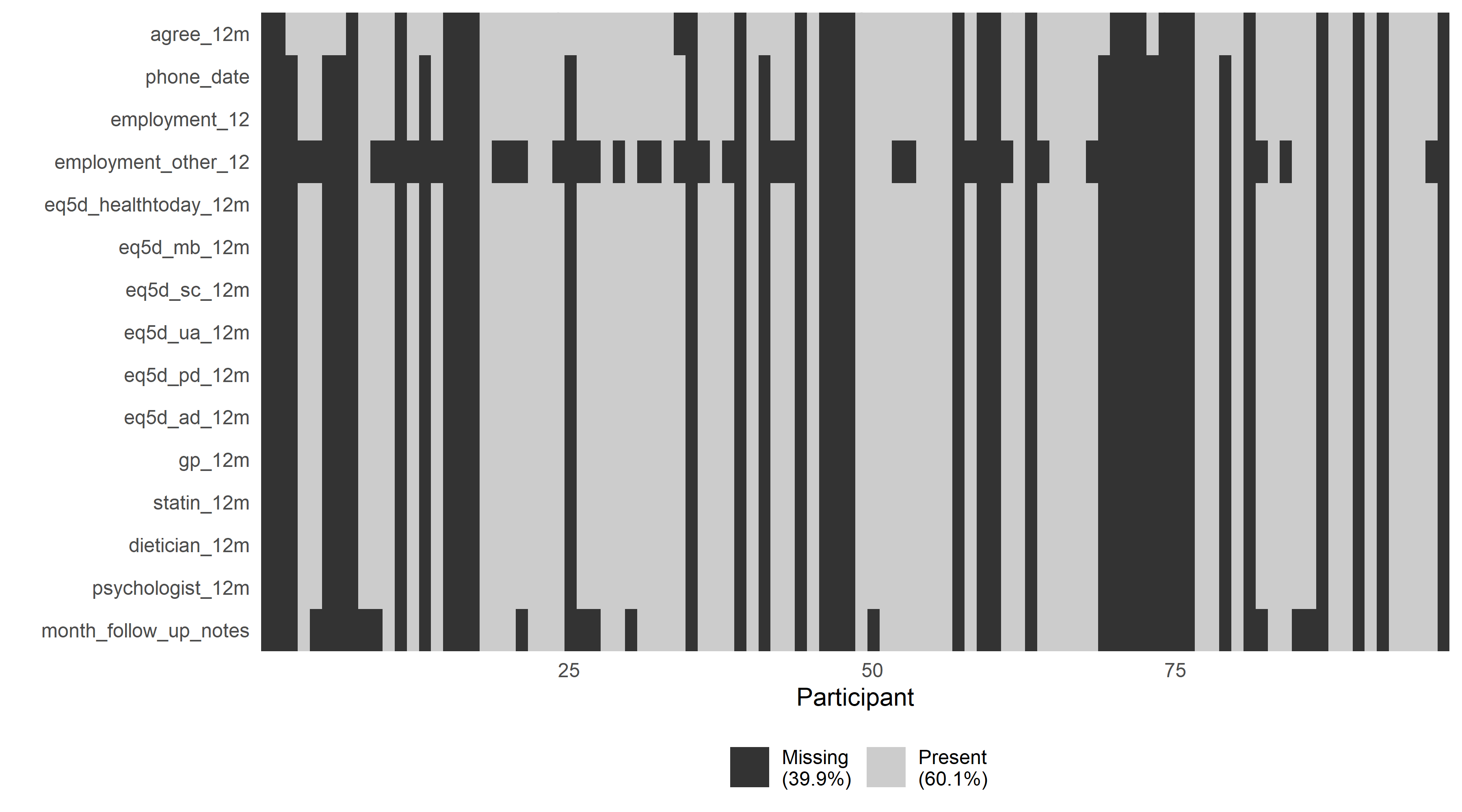


There was some missing data in the referral data. These date were extracted by the study team from routinely collected records.

## Missing scan data

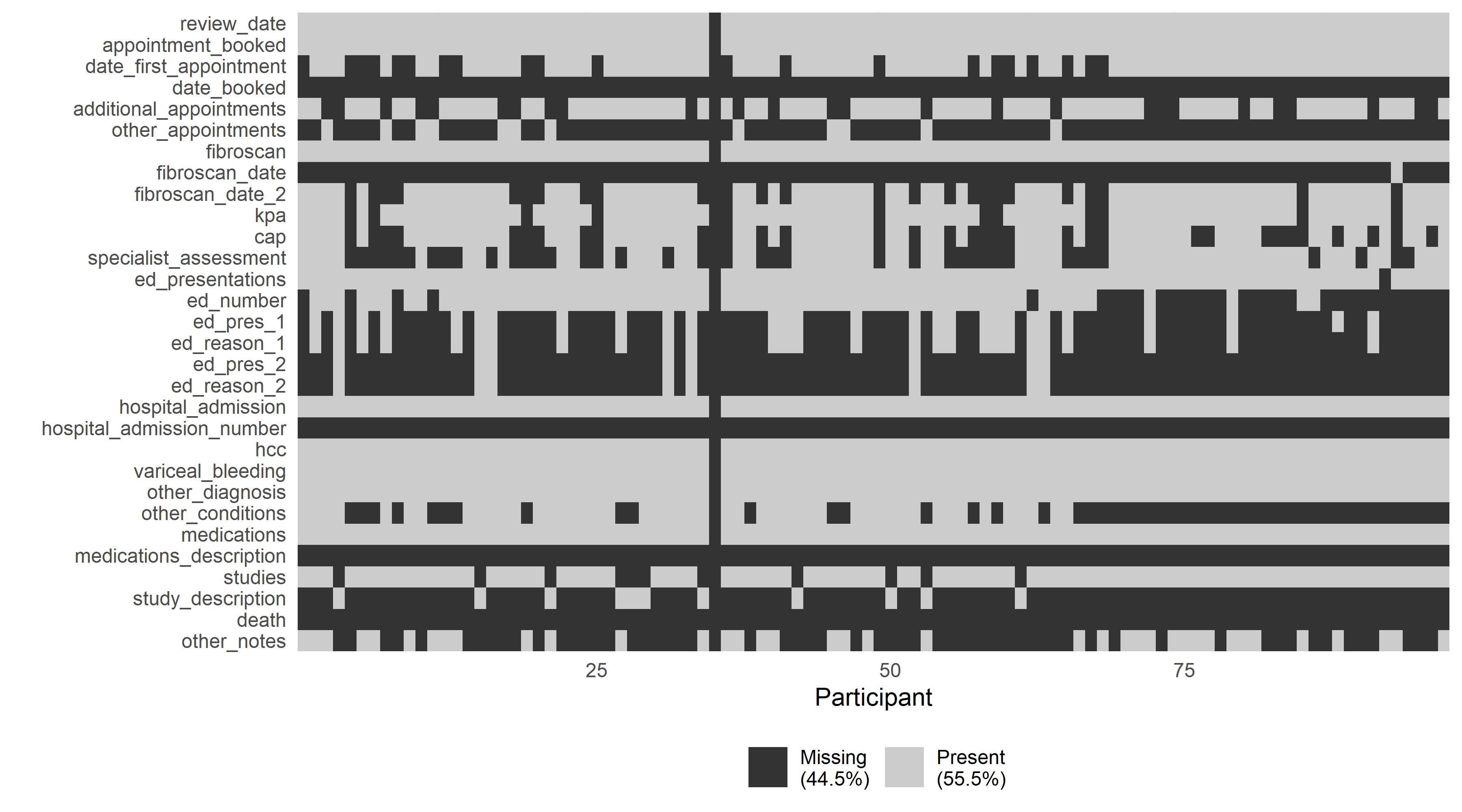


## Missing data at 12 month telephone follow-up



The vertical black ‘stripes’ show those people who had complete missing data, likely because they could not be contacted on the phone.

## Missing data from notes follow-up



# References

Altman, Douglas G. 1985. “Comparability of Randomised Groups.” *The Statistician* 34 (1): 125. <https://doi.org/10.2307/2987510>.

Barnett, Adrian. 2022. “Automated Detection of over- and Under-Dispersion in Baseline Tables in Randomised Controlled Trials.” <https://doi.org/10.17605/OSF.IO/J8EGW>.