Bone & Joint Open

Supplementary Material

10.1302/2633-1462.34.BJO-2022-0003.R1

Table i. The number and proportion of cases achieving each given threshold of International Hip Outcome Tool 12 improvement versus preoperative baseline score by labral procedure.

Threshold for iHOT-12 gain vs preop at 12 mths	Number of cases achieving iHOT-12 threshold by labral procedure (%)		
	Labral repair (n = 648)	Labral debridement (n = 590)	Overall (n = 1,238)
≥ 0	536 (82.7%)	478 (81.0%)	1,014 (81.9%)
≥ 1	531 (81.9%)	471 (79.8%)	1,002 (80.9%)
≥ 2	522 (80.6%)	464 (78.6%)	986 (79.6%)
≥ 3	516 (79.6%)	454 (76.9%)	970 (78.4%)
≥ 4	508 (78.4%)	438 (74.2%)	946 (76.4%)
≥ 5	503 (77.6%)	429 (72.7%)	932 (75.3%)
≥ 6	494 (76.2%)	425 (72.0%)	919 (74.2%)
≥ 7	485 (74.8%)	417 (70.7%)	902 (72.9%)
≥ 8	476 (73.5%)	409 (69.3%)	885 (71.5%)
≥ 9	469 (72.4%)	400 (67.8%)	869 (70.2%)
≥ 10	464 (71.6%)	393 (66.6%)	857 (69.2%)
≥ 11	456 (70.4%)	386 (65.4%)	842 (68.0%)
≥ 12	446 (68.8%)	378 (64.1%)	824 (66.6%)
≥ 13	440 (67.9%)	371 (62.9%)	811 (65.5%)
≥ 14	430 (66.4%)	366 (62.0%)	796 (64.3%)
≥ 15	423 (65.3%)	355 (60.2%)	778 (62.8%)
≥ 16	419 (64.7%)	348 (59.0%)	767 (62.0%)
≥ 17	409 (63.1%)	344 (58.3%)	753 (60.8%)
≥ 18	405 (62.5%)	336 (56.9%)	741 (59.9%)
≥ 19	397 (61.3%)	328 (55.6%)	725 (58.6%)
≥ 20	391 (60.3%)	317 (53.7%)	708 (57.2%)
≥ 21	380 (58.6%)	306 (51.9%)	686 (55.4%)
≥ 22	371 (57.3%)	300 (50.8%)	671 (54.2%)
≥ 23	367 (56.6%)	293 (49.7%)	660 (53.3%)
≥ 24	359 (55.4%)	289 (49.0%)	648 (52.3%)
≥ 25	349 (53.9%)	277 (46.9%)	626 (50.6%)
≥ 26	343 (52.9%)	269 (45.6%)	612 (49.4%)
≥ 27	337 (52.0%)	262 (44.4%)	599 (48.4%)
≥ 28	332 (51.2%)	259 (43.9%)	591 (47.7%)
≥ 29	328 (50.6%)	249 (42.2%)	577 (46.6%)
≥ 30	320 (49.4%)	247 (41.9%)	567 (45.8%)
≥ 31	310 (47.8%)	242 (41.0%)	552 (44.6%)
≥ 32	299 (46.1%)	235 (39.8%)	534 (43.1%)
≥ 33	290 (44.8%)	233 (39.5%)	523 (42.2%)
≥ 34	280 (43.2%)	223 (37.8%)	503 (40.6%)
≥ 35	274 (42.3%)	216 (36.6%)	490 (39.6%)
≥ 36	269 (41.5%)	210 (35.6%)	479 (38.7%)
≥ 37	267 (41.2%)	205 (34.7%)	472 (38.1%)
≥ 38	261 (40.3%)	191 (32.4%)	452 (36.5%)
≥ 39	255 (39.4%)	185 (31.4%)	440 (35.5%)
≥ 40	246 (38.0%)	180 (30.5%)	426 (34.4%)

≥ 41	237 (36.6%)	176 (29.8%)	413 (33.4%)
≥ 42	229 (35.3%)	171 (29.0%)	400 (32.3%)
≥ 43	224 (34.6%)	168 (28.5%)	392 (31.7%)
≥ 44	218 (33.6%)	161 (27.3%)	379 (30.6%)
≥ 45	210 (32.4%)	153 (25.9%)	363 (29.3%)
≥ 46	201 (31.0%)	145 (24.6%)	346 (27.9%)
≥ 47	197 (30.4%)	141 (23.9%)	338 (27.3%)
≥ 48	192 (29.6%)	135 (22.9%)	327 (26.4%)
≥ 49	186 (28.7%)	128 (21.7%)	314 (25.4%)
≥ 50	181 (27.9%)	123 (20.8%)	304 (24.6%)
≥ 55	141 (21.8%)	99 (16.8%)	240 (19.4%)
≥ 60	115 (17.7%)	73 (12.4%)	188 (15.2%)

iHOT-12, International Hip Outcome Tool 12 questionnaire.

Table ii. Demographic data and outcomes by labral procedure and femoroacetabular pathology group (defined by index procedure performed at index arthroscopy concurrently with labral procedure).

	FAI and labral procedure performed											
	Cam			Pincer			Mixed			None		
Variable	**Repair**	**Debridement**	**p-value**	**Repair**	**Debridement**	**p-value**	**Repair**	**Debridement**	**p-value**	**Repair**	**Debridement**	**p-value**
No. of cases (%)	865 (46)	1,015 (54)		86 (41.7)	120 (58.3)		469 (59.6)	318 (40.4)		239 (29.5)	572 (70.5)	
Mean age, yrs (SD)	35.2 (10.3)	36.8 (10.3)	< 0.001	35.1 (10.0)	37.2 (9.6)	0.142	35.9 (10.5)	38.6 (9.9)	< 0.001	34.2 (10.6)	37.5 (10.9)	< 0.001
Sex, n (%)			< 0.001			0.118			0.045			< 0.001
Female	527 (60.9)	520 (51.2)		64 (74.4)	77 (64.2)		248 (52.9)	145 (45.6)		209 (87.4)	390 (68.2)	
Male	338 (39.1)	495 (48.8)		22 (25.6)	43 (35.8)		221 (47.1)	173 (54.4)		30 (12.6)	182 (31.8)	
Mean BMI, kg/m^2 (SD); n (%)	25.1 (4.6); 579 (66.9)	25.9 (4.5); 651 (64.1)	0.002	25.4 (4.0); 55 (64.0)	27.4 (6.7); 42 (35.0)	0.073	25.7 (4.6); 324 (69.1)	26.6 (4.7); 162 (50.9)	0.031	25.0 (4.2); 126 (52.7)	26.0 (5.1); 145 (25.3)	0.074
Mean outcome score (95% CI); n (%)†												
Preop iHOT-12	32.8 (31.6 to 34.1); 761 (88.0)	33.0 (31.7 to 34.2); 873 (86.0)	0.878	28.5 (23.7 to 33.2); 69 (80.2)	25.3 (21.7 to 28.9); 68 (56.7)	0.291	32.9 (31.2 to 34.6); 428 (91.3)	31.8 (29.5 to 34.0); 260 (81.8)	0.448	29.4 (26.9 to 31.9); 189 (79.1)	29.6 (27.1 to 32.1); 225 (39.3)	0.910
Preop EQ-5D Index	0.52 (0.51 to 0.54); 765 (88.4)	0.54 (0.52 to 0.55); 922 (90.8)	0.205	0.48 (0.42 to 0.54); 75 (87.2)	0.45 (0.39 to 0.51); 88 (73.3)	0.453	0.54 (0.52 to 0.56); 427 (91.0)	0.50 (0.47 to 0.53); 271 (85.2)	0.033	0.50 (0.47 to 0.54); 202 (84.5)	0.50 (0.47 to 0.53); 321 (56.1)	0.947
6-mth iHOT-12	60.0 (57.5 to 62.6); 420 (48.6)	58.1 (55.6 to 60.5); 445 (43.8)	0.274	54.6 (45.2 to 63.9); 37 (43.0)	42.5 (34.2 to 50.8); 42 (35.0)	0.054	58.7 (54.9 to 62.4); 211 (45.0)	57.1 (52.4 to 61.9); 144 (45.3)	0.616	56.0 (50.9 to 61.1); 102 (42.7)	51.0 (45.8 to 56.3); 111 (19.4)	0.179
6-mth EQ-5D Index	0.68 (0.66 to 0.71); 434 (50.2)	0.67 (0.65 to 0.69); 464 (45.7)	0.321	0.65 (0.57 to 0.73); 42 (48.8)	0.58 (0.51 to 0.66); 57 (47.5)	0.227	0.66 (0.63 to 0.70); 222 (47.3)	0.65 (0.61 to 0.69); 150 (47.2)	0.608	0.66 (0.62 to 0.70); 113 (47.3)	0.67 (0.64 to 0.70); 232 (40.6)	0.843
12-mth iHOT-12	64.1 (61.3 to 67.0); 382 (44.2)	59.7 (56.9 to 62.5); 383 (37.7)	0.028	48.4 (39.4 to 57.5); 37 (43.0)	45.7 (37.6 to 53.8); 42 (35.0)	0.651	60.9 (56.9 to 64.9); 201 (42.9)	58.3 (53.5 to 63.0); 139 (43.7)	0.411	60.2 (54.2 to 66.2); 97 (40.6)	52.9 (47.4 to 58.4); 105 (18.4)	0.076

	FAI and labral procedure performed											
	Cam			Pincer			Mixed			None		
Variable	Repair	Debridement	p-value	Repair	Debridement	p-value	Repair	Debridement	p-value	Repair	Debridement	p-value
12-mth EQ-5D Index	0.71 (0.68 to 0.73); 393 (45.4)	0.68 (0.65 to 0.70); 398 (39.2)	0.104	0.61 (0.55 to 0.68); 39 (45.3)	0.65 (0.57 to 0.73); 46 (38.3)	0.449	0.69 (0.66 to 0.72); 213 (45.4)	0.67 (0.63 to 0.72); 146 (45.9)	0.507	0.66 (0.60 to 0.71); 108 (45.2)	0.68 (0.64 to 0.71); 205 (35.8)	0.534
At 6 mths												
Change vs preop ‡												
Change iHOT-12 – 6 mths (95% CI); n (%)	+27.0 (24.4 to 29.7); 372 (43.0); p=<0.0001	+26.4 (23.8 to 29.1); 382 (37.6); p=<0.0001	0.752	+25.8 (17.0 to 34.6); 30 (34.9); p=<0.0001	+18.5 (9.1 to 27.8); 32 (26.7); p=<0.001	0.251	+27.4 (23.2 to 31.6); 194 (41.4); p=<0.0001	+26.8 (22.0 to 31.7); 126 (39.6); p=<0.0001	0.867	+28.0 (22.4 to 33.7); 83 (34.7); p=<0.0001	+19.1 (14.0 to 24.2); 98 (17.1); p=<0.0001	0.020
Change EQ-5D Index - 6-mth (95% CI); n (%)	+0.15 (0.13 to 0.17); 388 (44.9); p=<0.0001	+0.15 (0.12 to 0.17); 418 (41.2); p=<0.0001	0.726	+0.22 (0.13 to 0.31); 36 (41.9); p=<0.0001	+0.15 (0.09 to 0.21); 46 (38.3); p=<0.0001	0.210	+0.13 (0.10 to 0.17); 202 (43.1); p=<0.0001	+0.13 (0.08 to 0.17); 135 (42.5); p=<0.0001	0.804	+0.17 (0.11 to 0.22); 96 (40.2); p=<0.0001	+0.12 (0.08 to 0.16); 150 (26.2); p=<0.0001	0.180
Achieving MCID iHOT-12 at 6 mths, n (%)	Yes = 252 of 372 (67.7); No = 120 of 372 (32.3)	Yes = 255 of 382 (66.8); No = 127 of 382 (33.2)	0.773	Yes = 21 of 30 (70.0); No = 9 of 30 (30.0)	Yes = 19 of 32 (59.4); No = 13 of 32 (40.6)	0.382	Yes = 127 of 194 (65.5); No = 67 of 194 (34.5)	Yes = 81 of 126 (64.3); No = 45 of 126 (35.7)	0.829	Yes = 57 of 83 (68.7); No = 26 of 83 (31.3)	Yes = 59 of 98 (60.2); No = 39 of 98 (39.8)	0.237
Achieving SCB iHOT-12 at 6 mths, n (%)	Yes = 181 of 372 (48.7); No = 191 of 372 (51.3)	Yes = 178 of 382 (46.6); No = 204 of 382 (53.4)	0.571	Yes = 15 of 30 (50.0); No = 15 of 30 (50.0)	Yes = 9 of 32 (28.1); No = 23 of 32 (71.9)	0.077	Yes = 91 of 194 (46.9); No = 103 of 194 (53.1)	Yes = 61 of 126 (48.4); No = 65 of 126 (51.6)	0.792	Yes = 38 of 83 (45.8); No = 45 of 83 (54.2)	Yes = 37 of 98 (37.8); No = 61 of 98 (62.2)	0.275
At 12 mths												
Change vs preop ‡												
Change iHOT-12 - 12-mth (95% CI); n (%)	+29.1 (26.1 to 32.2); 344 (39.8); p=<0.0001	+26.5 (23.6 to 29.4); 338 (33.3); p=<0.0001	0.218	+19.6 (10.6 to 28.7); 34 (39.5); p=<0.001	+22.3 (12.7 to 31.9); 34 (28.3); p=<0.0001	0.681	+27.8 (23.6 to 32.1); 187 (39.9); p=<0.0001	+23.2 (18.4 to 28.1); 124 (39.0); p=<0.0001	0.164	+32.3 (25.9 to 38.8); 83 (34.7); p=<0.0001	+21.2 (14.9 to 27.5); 94 (16.4); p=<0.0001	0.016
Change EQ-5D Index - 12-mth (95% CI); n (%)	+0.17 (0.14 to 0.19);	+0.14 (0.12 to 0.17); 372	0.210	+0.12 (0.03 to 0.21); 37	+0.19 (0.11 to 0.26); 36 (30.0); p=<0.0001	0.282	+0.15 (0.11 to 0.19);	+0.13 (0.08 to 0.18); 134 (42.1); p=<0.0001	0.475	+0.16 (0.10 to 0.21); 99	+0.12 (0.08 to 0.17); 134	0.333

	FAI and labral procedure performed											
	Cam			Pincer			Mixed			None		
Variable	Repair	Debridement	p-value	Repair	Debridement	p-value	Repair	Debridement	p-value	Repair	Debridement	p-value
	355 (41.0); p=<0.0001	(36.7); p=<0.0001		(43.0); p = 0.008			199 (42.4); p=<0.0001			(41.4); p=<0.0001	(23.4); p=<0.0001	
Achieving MCID iHOT-12 at 12 mths, n (%)	Yes = 238 of 344 (69.2); No = 106 of 344 (30.8)	Yes = 221 of 338 (65.4); No = 117 of 338 (34.6)	0.290	Yes = 20 of 34 (58.8); No = 14 of 34 (41.2)	Yes = 20 of 34 (58.8); No = 14 of 34 (41.2)	1.000	Yes = 124 of 187 (66.3); No = 63 of 187 (33.7)	Yes = 78 of 124 (62.9); No = 46 of 124 (37.1)	0.538	Yes = 58 of 83 (69.9); No = 25 of 83 (30.1)	Yes = 52 of 94 (55.3); No = 42 of 94 (44.7)	0.046
Achieving SCB iHOT-12 at 12 mths, n (%)	Yes = 179 of 344 (52.0); No = 165 of 344 (48.0)	Yes = 154 of 338 (45.6); No = 184 of 338 (54.4)	0.091	Yes = 10 of 34 (29.4); No = 24 of 34 (70.6)	Yes = 11 of 34 (32.4); No = 23 of 34 (67.6)	0.793	Yes = 94 of 187 (50.3); No = 93 of 187 (49.7)	Yes = 57 of 124 (46.0); No = 67 of 124 (54.0)	0.458	Yes = 49 of 83 (59.0); No = 34 of 83 (41.0)	Yes = 37 of 94 (39.4); No = 57 of 94 (60.6)	0.009

†The values are given as the mean score (95% confidence interval (CI)); number (%) of cases available for follow-up.
‡For cases with pre- and postoperative follow-up data, the values are given as the mean score improvement (95% CI); number (%) of cases available for follow-up; p-values derived with the paired *t*-test.
EQ-5D, EuroQol five-dimension questionnaire; FAI, femoroacetabular impingement; iHOT-12, International Hip Outcome Tool 12 questionnaire; MCID, minimal clinically important difference; SCB, significant clinical benefit; SD, standard deviation.

Multivariable model

A large proportion of patients were lost to follow-up at one year and, furthermore, there were substantial differences in the 12-month iHOT-12 questionnaire follow-up rates between labral repair (12-month follow-up rate = 39.1%) and labral debridement (12-month iHOT-12 follow-up rate = 29.1%) groups. Considering the overall cohort including both surgical groups, we found that patients who returned both pre- and 12-month postoperative iHOT-12 questionnaires were, on average, older (37.1 vs 36.0 years; p = 0.003, independent-samples *t*-test), more likely to be female, and more likely to be missing data for BMI than those who did not return questionnaires. There were also differences in the observed proportions of femoroacetabular impingement (FAI) pathology types (Table iii).

The differences in follow-up rates between the two surgical groups may therefore confound our primary outcome measure of 12-month iHOT-12 gain by virtue of the fact that responders are inherently different to non-responders. Thus, we developed a novel methodology to attempt to create a synthetic cohort of cases in both surgical groups to better balance the demographic differences that may have arisen from differences in follow-up rates.

Methodology

We used a combination of random sampling and propensity score matching in order to attempt to account for the impact of the above, building upon a similar methodology we have applied in previous work.[1]

1. Patients were classified according to whether they did ('responders') or did not ('non-responders') return iHOT-12 questionnaires both preoperatively and at 12 months postoperatively.

2. We randomly sampled 200 cases from the 1,659 cases in the overall labral repair group (overall meaning in this instance to include both responders and non-responders). This random sample of 200 cases would, on average, be expected to select 78 responders (39.1% (the labral repair iHOT-12 responder rate) multiplied by the random sample size (200) = 78.2).

3. Propensity score matching (1 to 1 matching without replacement, with a random match order and including the following co-variates – age, sex, BMI, and FAI pathology type) was then used to match these 200 randomly selected patients to their closest match in the group of labral repair 'responders' – with matches being allowed to be drawn from the entire sample of labral repair responders (n = 648). The expectation was that the 78 randomly selected responders would match back to themselves and the remaining randomly selected non-responders (on average 122 cases: 200 minus 78 responders) would each match to one of the remaining responders available, and with the closest propensity score (i.e. one of the 648 responder cases which had not already matched to a responder).

This yields a dataset of 200 matched responders. The propensity score matching in effect allows the creation of a 'synthetic' cohort of responder patients who are demographically more similar to the original cohort comprising both responders and non-responders.

4. Steps 2 to 3 were repeated over 1,000 iterations.

5. The resultant 1,000 matched responder datasets of 200 patients each were then combined, and for each unique responder we counted the number of times the case had been sampled over the 1,000 iterations. On average, over 1,000 iterations, we would expect each unique labral repair responder case to have been selected $200 \div 648 \times 1{,}000 = 308.6$ times (95% binomial confidence interval (CI) 273.2 to 345.8) which is confirmed in Figure a.

6. In order to select only those responders who were most similar to the overall cohort of responders and non-responders, we excluded those responders who were likely to be less representative of the overall cohort on the basis that they were selected fewer times than the lower 95% binomial CI of the probability of selection if the process were completely random.

We therefore selected only those labral repair responders who were selected more than 273.2 times which yielded 401 cases.

7. Steps 2 to 6 were then repeated for the labral debridement cohort.

The number of selections of each labral debridement case over 1,000 iterations is shown in Figure b. According to the same calculation described in step 6 we only included labral debridement responders who were selected more than 300.8 times, which yielded 246 cases.

8. Propensity score matching was then used to match labral repair and debridement cases (1 to 1 matching without replacement, with a random match order and including the following covariates – age, sex, BMI, and FAI pathology type). The resultant matched patients were then used as the basis of a linear regression model predicting iHOT-12 12-month gain (including the aforementioned covariates).

9. Step 8 was then repeated over 1,000 iterations to generate 1,000 regression models. Pooled regression estimates for the resultant 1,000 models were then derived using Rubin's rules. For non-parametric variables (age and preoperative iHOT-12 score), empirical 95% CIs for estimates were derived by selecting estimates at the 2.5th and 97.5th percentiles.

Results

Results of the pooled regression analyses are discussed and shown in Table iv in the main manuscript. For comparison, the results of a conventional multivariable linear regression model for the study cohort are shown in Table iv in supplementary materials below.

Table iii. Demographic data for the entire cohort, stratified by those who returned pre- and 12-month postoperative International Hip Outcome Tool 12 questionnaires (responders) and those who did not (non-responders).

Variable	Non-responder	Responder	Overall	p-value
No. of cases (%)	2,446 (66.4)	1,238 (33.6)	3,684 (100)	
Mean age, yrs (SD)	36.0 (10.4)	37.1 (10.4)	36.4 (10.4)	0.003*
Sex, n (%)				< 0.001†
Female	1,391 (56.9)	789 (63.7)	2,180 (59.2)	
Male	1,055 (43.1)	449 (36.3)	1,504 (40.8)	
Mean BMI, kg/m^2 (SD)	25.7 (4.7); n = 1,231	25.7 (4.5); n = 853	25.7 (4.6); n = 2,084	0.917*
BMI group, n (%)				< 0.001†
< 25	619 (25.3)	419 (33.8)	1,038 (28.2)	
25 to 30	407 (16.6)	304 (24.6)	711 (19.3)	
≥ 30	205 (8.4)	130 (10.5)	335 (9.1)	
Missing	1,215 (49.7)	385 (31.1)	1,600 (43.4)	
FAI type, n (%)				< 0.001†
Cam	1,198 (49.0)	682 (55.1)	1,880 (51.0)	
Pincer	138 (5.6)	68 (5.5)	206 (5.6)	
Mixed	476 (19.5)	311 (25.1)	787 (21.4)	
None (no FAI procedure)	634 (25.9)	177 (14.3)	811 (22.0)	

*Independent-samples *t*-test.
†Chi-squared test.
FAI, femoroacetbaular impingement; SD, standard deviation.

Table iv. Results of conventional multivariable linear regression model predicting iHOT-12 gain at 12-month follow-up versus preoperative score. Age is modelled as a categorical variable here to provide clinical context, but modelled as a continuous variable in the novel method in order to allow more appropriate propensity score matching.

Variable	Predictors of iHOT-12 score improvement at 12 mths vs preop baseline (Coefficient, 95%CI, p value)
Labral procedure	
Repair	4.47 (1.39 to 7.55), 0.004
Debridement	Reference
Age group	
< 40 yrs	Reference
≥ 40 yrs	-1.46, (-4.55 to 1.62), 0.353
Sex	
Female	Reference
Male	0.86, (-2.50 to 4.23), 0.616
BMI group, n (%)	
< 25 kg/m^2	Reference
25 to 30 kg/m^2	-1.86, (-5.91 to 2.20), 0.369
≥ 30 kg/m^2	-1.28, (-6.67 to 4.12), 0.643
Missing	-1.71, (-5.53 to 2.10), 0.379
FAI type	
Cam	Reference
Pincer	-10.08, (-16.88 to -3.28), 0.004
Mixed	-1.85, (-5.49 to 1.80), 0.321
None (no FAI procedure recorded)	-2.67, (-7.23 to 1.89), 0.251
Preoperative scores	
iHOT-12	-0.50, (-0.59 to -0.41), 0.000
EQ-5D Index	Not applicable
N	1,238
r2	0.11

CI, confidence interval; EQ-5D, EuroQol five-dimension questionnaire; FAI, femoroacetabular impingement; iHOT-12, International Hip Outcome Tool 12 questionnaire.

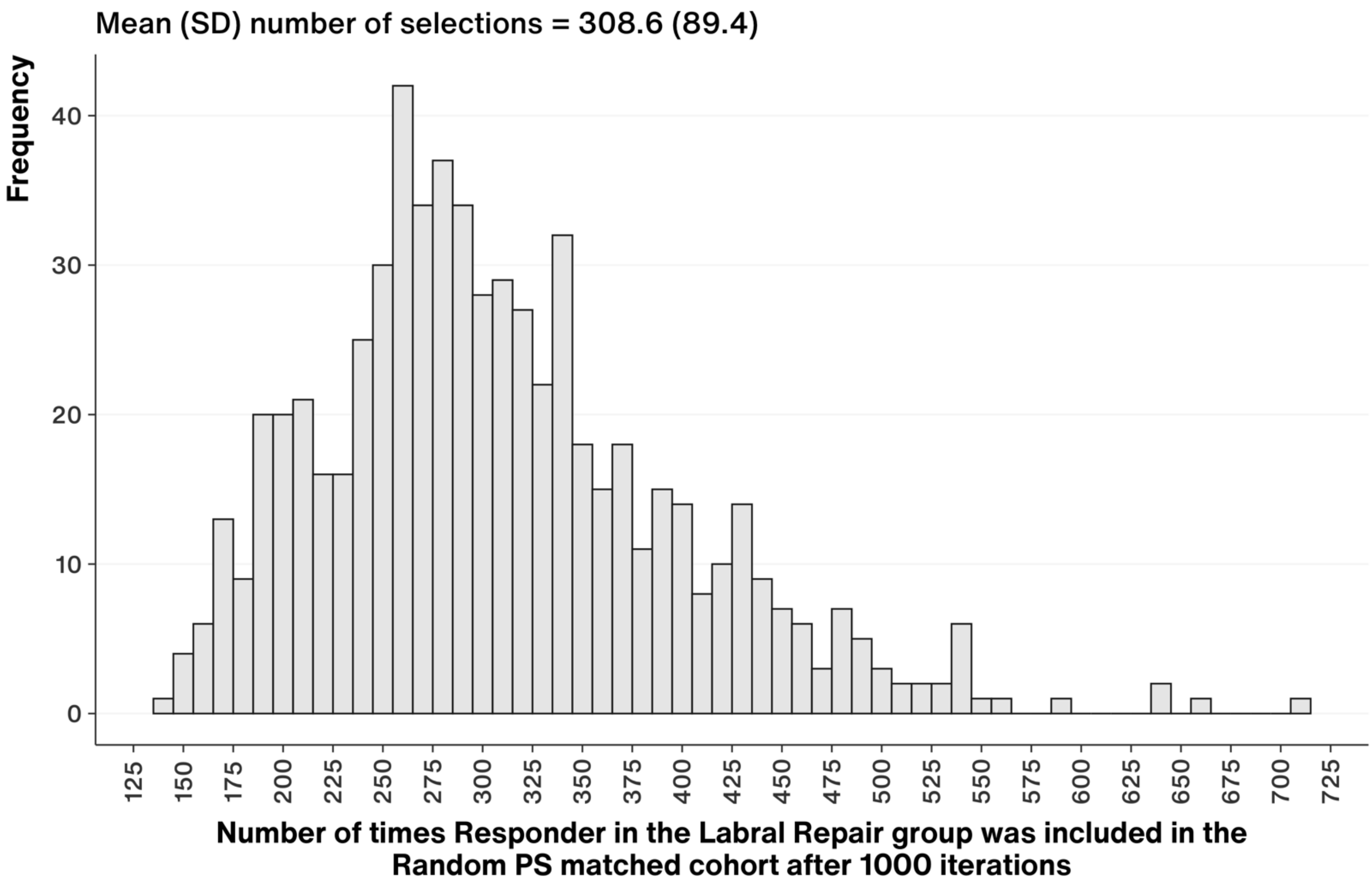

Fig. a. Number of times each unique responder in the labral repair group was included in the randomly selected responder cohort after 1,000 iterations. PS, propensity score; SD, standard deviation.

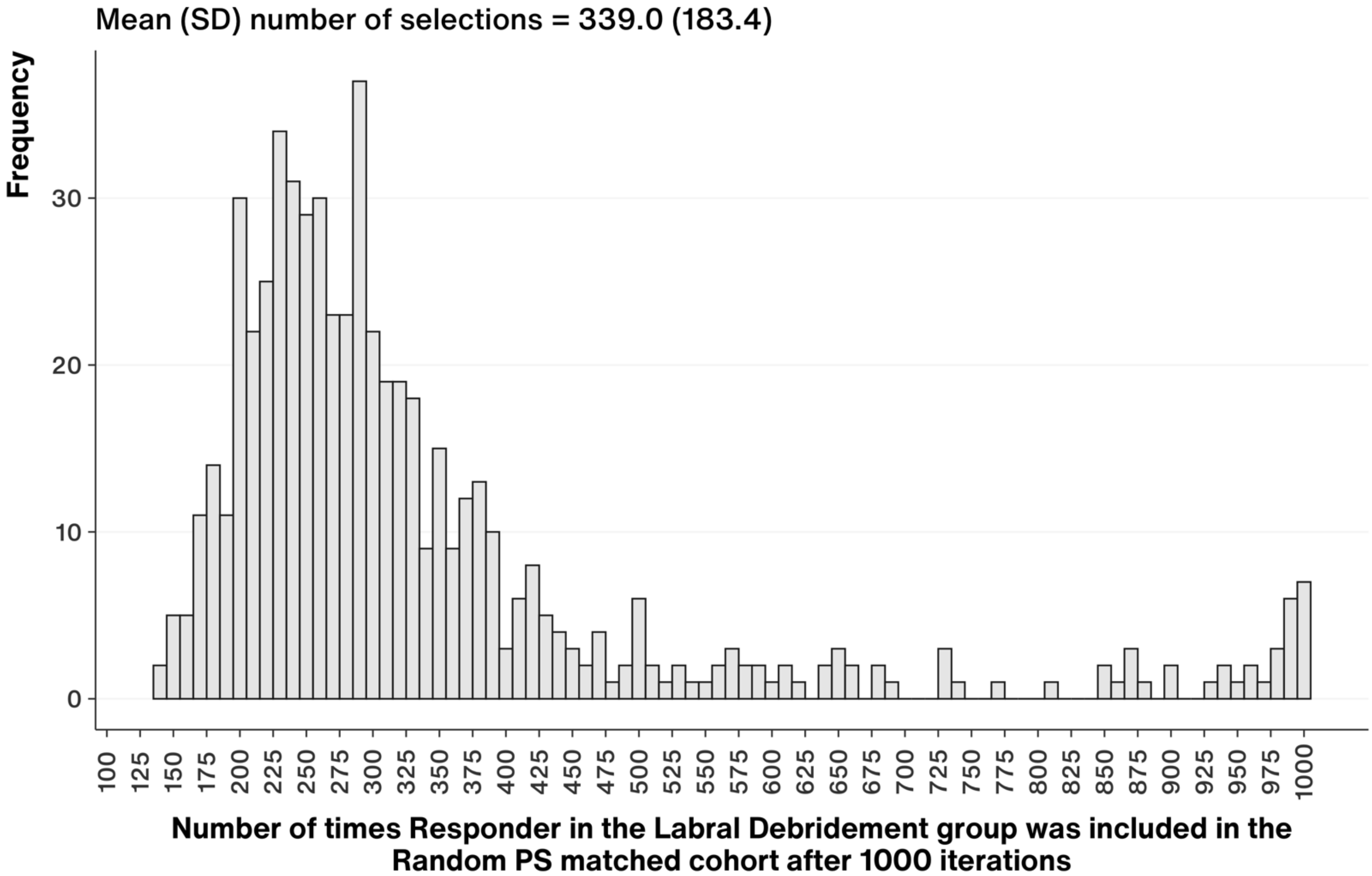

Fig. b. Number of times each unique responder in the labral debridement group was included in the randomly selected responder cohort after 1,000 iterations. PS, propensity score; SD, standard deviation.

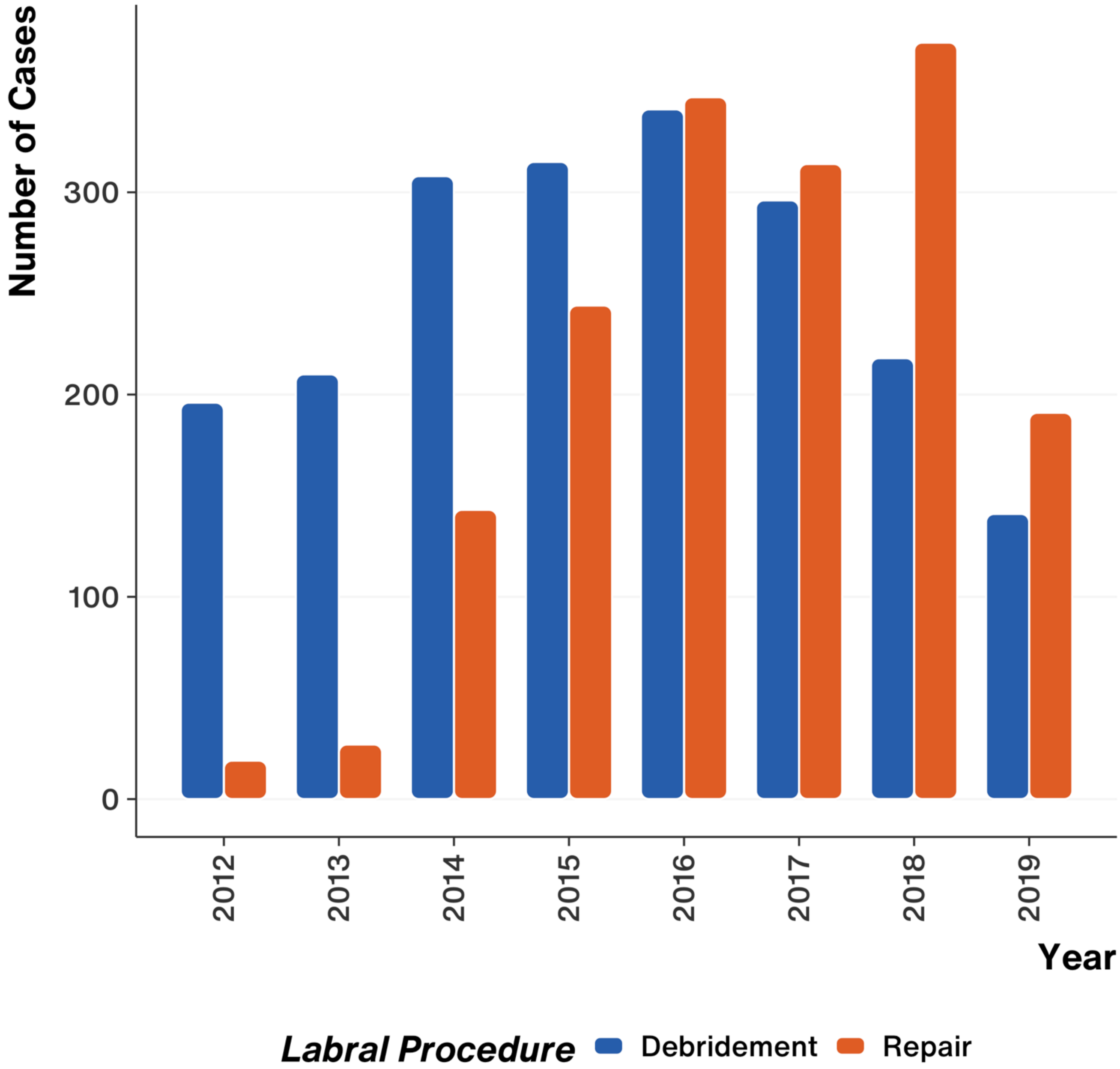

Fig. c. Frequency of labral repair versus debridement procedures recorded in the Non-Arthroplasty Hip Register over time (January 2012 to July 2019).

References

1. **Holleyman R, Sohatee MA, Witt J, et al.** Periacetabular osteotomy for developmental dysplasia of the hip and femoroacetabular impingement: a study using the U.K. Non-Arthroplasty Hip Registry (NAHR) data set. *J Bone Joint Surg Am*. 2020;102-A(15):1312-1320.