BU Physical Therapy Data Analysis

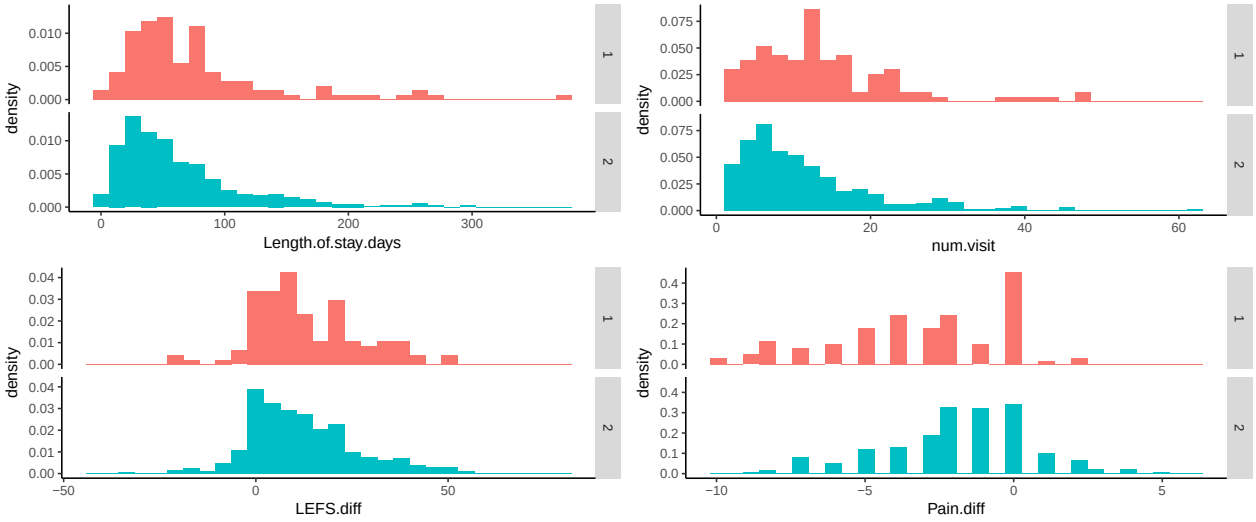
March 5, 2018

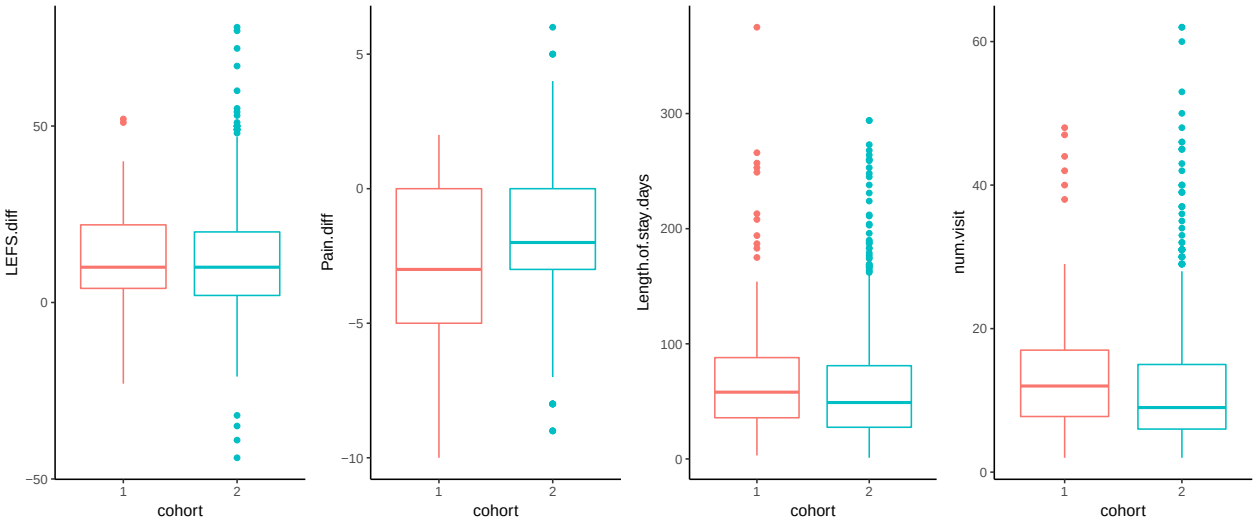
Introduction and Summary of Preliminary Analysis

Our clients, James Camarinos & Lee Marinko, have provded patient level outcome data from the BU Physical Therapy and Athletic Training Center. Outcomes are measured in several ways: the difference in the intake versus discharge pain score, the difference in the intake versus discharge level of disability measurement (which is based on a survey depending on the patients' injury type), and the length of stay. The data can be separated into 2 cohorts (one from 2014-2015 and one from 2016-2017). The distinguishing feature between the cohorts is that patients in the more recent cohort (cohort 2) were measured more frequently than patients in the first cohort on pain and disability (at each visit versus sporadically). The first question we address is whether patient outcomes are different across cohorts. This investigation seeks to guage whether the increased level of data collection for the second cohort is associated with different outcomes. Second, using the data from cohort 2, we investigate whether outcomes differ depending on certain factors including gender, age, and body region of injury.

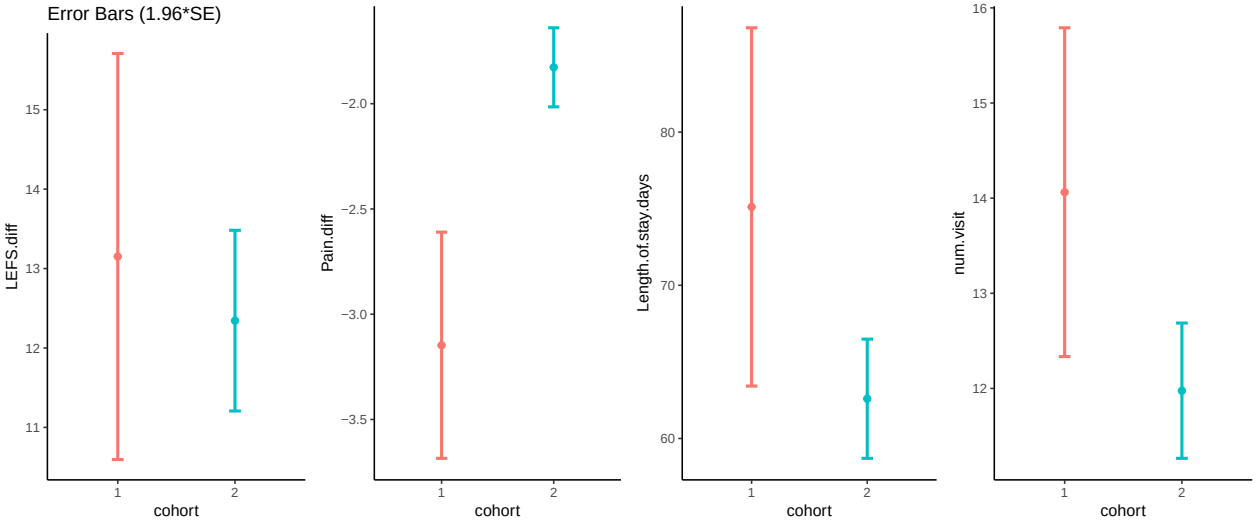
Question 1 - Are Patient Outcomes Different Across Cohorts?

First we visualize the distributions of each raw response variable for cohort 1 and cohort 2 using histograms and boxplots. Note that the cohort 1 sample consists of 112 patients and the cohort 2 sample consists of 671 patients. Since the only survey represented in cohort 1 is the LEFS, we have restricted cohort 2 to LEFS patients.



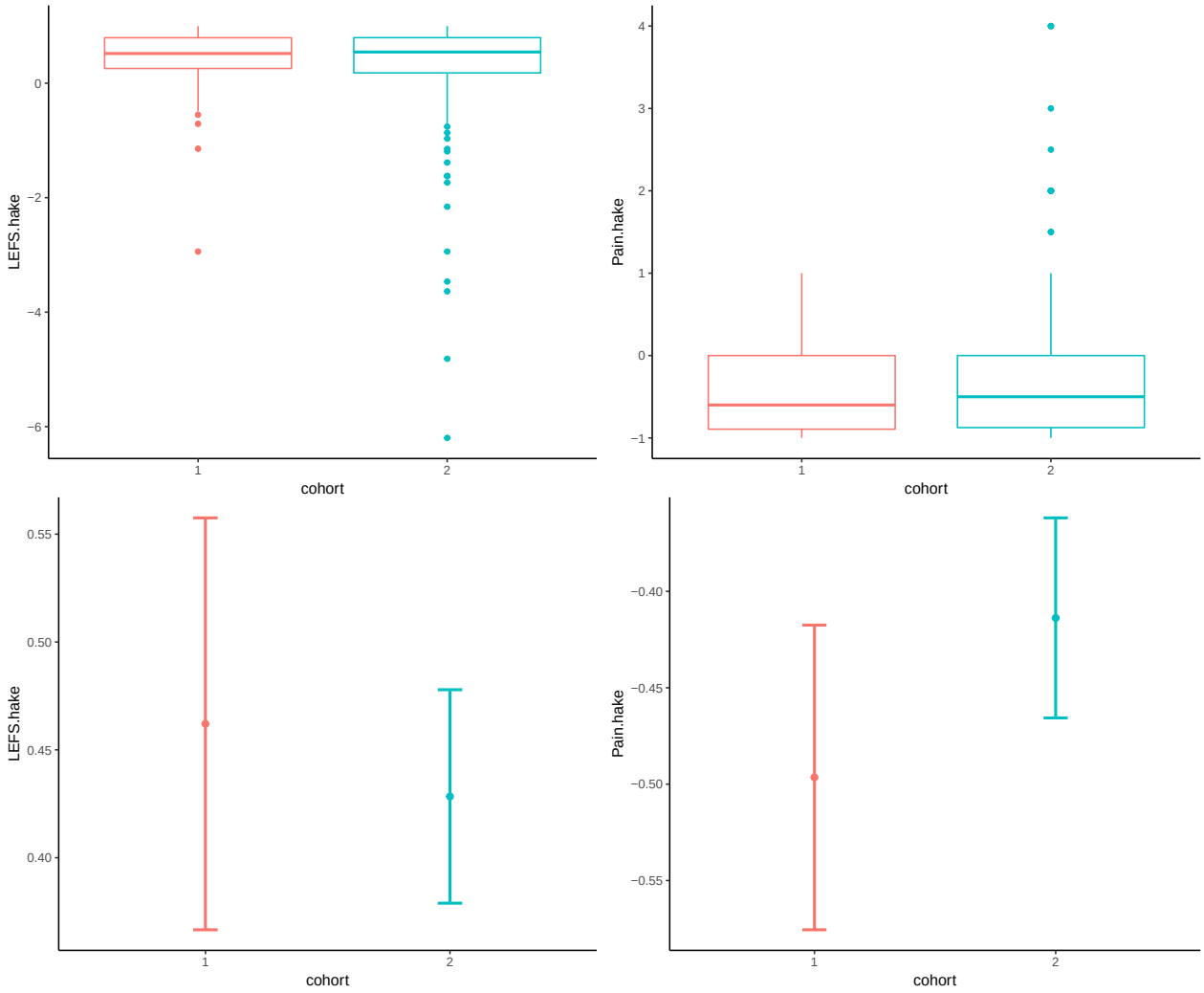


Below, we visualize the mean and standard error for each of the raw responses. We can see that the average decrease in pain is lower for cohort 2 as compared with cohort 1. This observation is validated in the t-tests below.



Next, we utilize the same set of visualizations for the hake score with respect to pain and LEFS. Note the highest possible score for the LEFS is 80. For pain and LEFS, we calculate the hake score as:

hake pain = (discharge pain − intake pain)/(intake pain)
hake LEFS = (discharge LEFS − intake LEFS)/(80 − intake LEFS)



Statistical Tests

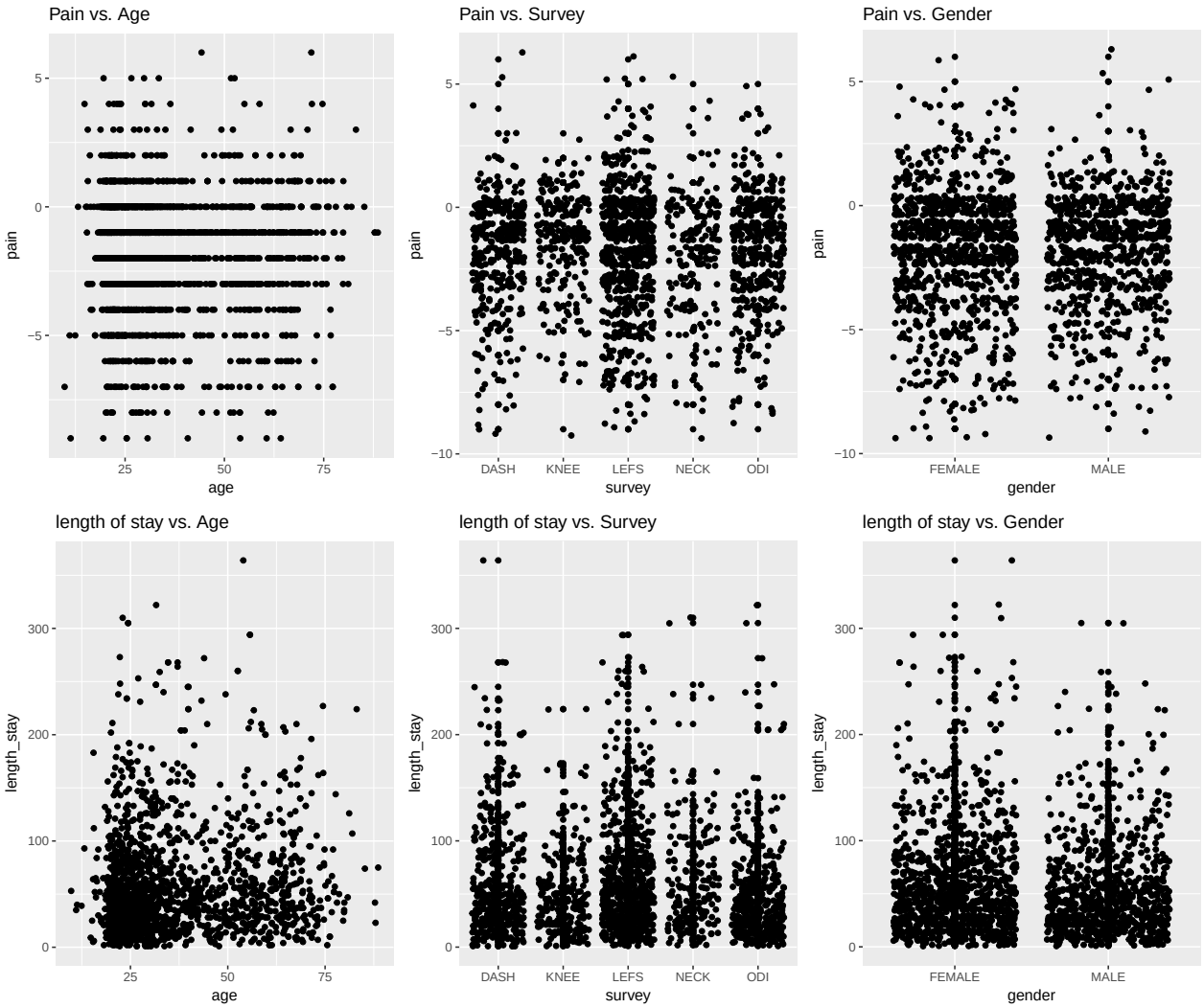
We carry out t-tests for each response variable to test the null hypothesis that the mean response for cohort 1 is equal to the mean response for cohort 2. The first table shows the results of the t-test for the raw outcome scores whereas the second table shows the results of the t-tests for the normalized hake scores.

Construct	Confidence interval for diff.	Reject H0?
Length.of.stay.days	(-0.04,25.24)	No
num.visit	(0.22,4.04)	Yes
Pain.diff	(-1.71,-0.55)	Yes
LEFS.diff	(-2.57,3.16)	No

Response	Confidence interval for diff.	Reject H0?
Pain.hake	(-0.18,0.01)	Yes
LEFS.hake	(-0.07,0.14)	No

Question 2 - Do Patient Outcomes Depend on Gender/ Age/ Body Region?

Next, we investigate whether outcomes depend on gender, age, and body region. Before fitting linear regression models, we visually explore the bivariate relationships between pain difference and the explanatory variables - age, survey type, and gender. We can see that there does not appear to be any discernable difference in this particular response across the range of the explanatory variables. This lack of effect is further validated in the regression models fit below.



Next, we fit a series of linear regression models to test whether outcomes depend on gender, age, and body region. We split our models into 2 catgories, depending on whether the response variables are survey specific (change in survey outcome score) or not (pain, length of stay, number of visits). The first set of models we present are not survey specific.

Table 3: Fitting linear model: pain ~ age + survey + gender + body

	Estimate	Std. Error	t value	Pr(>|t|)
age	-0.001092	0.003388	-0.3224	0.7472
surveyKNEE	0.1929	0.3845	0.5018	0.6159
surveyLEFS	-0.5363	0.339	-1.582	0.1139
surveyNECK	-0.2579	0.3048	-0.8461	0.3976
surveyODI	-0.3622	0.3294	-1.1	0.2717
genderMALE	0.02214	0.1101	0.2011	0.8407

	Estimate	Std. Error	t value	Pr(>|t|)
bodyBalance	2.123	2.361	0.899	0.3688
bodyCervical	-0.2184	0.462	-0.4726	0.6366
bodyElbow	-0.125	0.5887	-0.2124	0.8318
bodyFoot/Ankle	0.3712	0.4572	0.812	0.4169
bodyHand	-0.4864	1.051	-0.4629	0.6435
bodyHip	0.4368	0.478	0.9138	0.361
bodyKnee	-0.01738	0.4568	-0.03805	0.9697
bodyLumbar	0.1605	0.4385	0.3661	0.7144
bodyShoulder	-0.3169	0.4774	-0.6637	0.507
bodyThoracic	-0.1905	0.4865	-0.3917	0.6954
bodyWrist	0.2432	0.5884	0.4133	0.6795
(Intercept)	-1.494	0.4795	-3.115	0.001868

Table 4: Fitting linear model: HakePain ~ age + survey + gender + body

	Estimate	Std. Error	t value	Pr(>|t|)
age	0.001021	0.0009185	1.112	0.2664
surveyKNEE	-0.08203	0.1035	-0.7922	0.4283
surveyLEFS	-0.01315	0.09064	-0.145	0.8847
surveyNECK	0.001979	0.08069	0.02453	0.9804
surveyODI	0.007464	0.08804	0.08479	0.9324
genderMALE	-0.01969	0.02978	-0.6611	0.5087
bodyCervical	0.09735	0.1287	0.7562	0.4497
bodyElbow	-0.05134	0.1643	-0.3125	0.7547
bodyFoot/Ankle	0.05446	0.1281	0.4252	0.6707
bodyHand	0.06386	0.3	0.2129	0.8315
bodyHip	0.06152	0.1335	0.4609	0.6449
bodyKnee	0.04211	0.1278	0.3295	0.7418
bodyLumbar	0.09255	0.1221	0.7579	0.4486
bodyShoulder	-0.02845	0.1349	-0.2109	0.833
bodyThoracic	0.1091	0.1354	0.8061	0.4203
bodyWrist	0.106	0.1626	0.652	0.5145
(Intercept)	-0.4864	0.136	-3.577	0.0003576

Table 5: Fitting linear model: visit ~ age + survey + gender + body

	Estimate	Std. Error	t value	Pr(>|t|)
age	0.05409	0.01246	4.34	1.504e-05
surveyKNEE	-3.482	1.415	-2.462	0.01392
surveyLEFS	-0.02899	1.247	-0.02325	0.9815
surveyNECK	-1.415	1.121	-1.262	0.2071
surveyODI	-1.36	1.212	-1.122	0.2619
genderMALE	-1.973	0.4052	-4.868	1.224e-06
bodyBalance	-9.971	8.688	-1.148	0.2512
bodyCervical	1.496	1.7	0.8802	0.3789
bodyElbow	1.051	2.166	0.485	0.6277
bodyFoot/Ankle	-0.8809	1.682	-0.5237	0.6005

	Estimate	Std. Error	t value	Pr(>|t|)
bodyHand	-3.758	3.866	-0.972	0.3312
bodyHip	1.037	1.759	0.5894	0.5557
bodyKnee	2.208	1.681	1.314	0.1891
bodyLumbar	-0.6069	1.613	-0.3761	0.7069
bodyShoulder	1.037	1.757	0.5903	0.555
bodyThoracic	-0.5957	1.79	-0.3328	0.7393
bodyWrist	-2.482	2.165	-1.146	0.2517
(Intercept)	10.39	1.764	5.89	4.58e-09

Table 6: Fitting linear model: length_stay ~ age + survey + gender + body

	Estimate	Std. Error	t value	Pr(>|t|)
age	0.3246	0.07128	4.553	5.628e-06
surveyKNEE	-19.15	8.09	-2.367	0.01805
surveyLEFS	2.436	7.134	0.3415	0.7327
surveyNECK	-3.74	6.414	-0.5831	0.5599
surveyODI	-7.072	6.931	-1.02	0.3077
genderMALE	-9.94	2.318	-4.289	1.886e-05
bodyBalance	-10.24	49.69	-0.206	0.8368
bodyCervical	6.458	9.722	0.6643	0.5066
bodyElbow	-2.01	12.39	-0.1622	0.8711
bodyFoot/Ankle	-8.869	9.62	-0.922	0.3567
bodyHand	-14.81	22.11	-0.6698	0.5031
bodyHip	1.445	10.06	0.1436	0.8858
bodyKnee	6.951	9.612	0.7232	0.4697
bodyLumbar	-6.802	9.228	-0.7371	0.4611
bodyShoulder	1.232	10.05	0.1226	0.9024
bodyThoracic	0.02626	10.24	0.002565	0.998
bodyWrist	-20.81	12.38	-1.68	0.09303
(Intercept)	54.14	10.09	5.366	9.07e-08

Next, for the survey specific models, we fit separate models for each of 5 survey types - first for the raw differences and then for the hake score changes. The survey order of these models is: DASH, LEFS, ODI, NECK, KNEE.

DASH score as response

Table 7: Fitting linear model: outcome ~ age + gender + body

	Estimate	Std. Error	t value	Pr(>|t|)
age	0.04302	0.05716	0.7526	0.4522
genderMALE	1.24	1.971	0.6288	0.5299
bodyCervical	-2.199	6.977	-0.3151	0.7528
bodyElbow	-10.9	7.179	-1.519	0.1297
bodyFoot/Ankle	6.867	14.38	0.4776	0.6332
bodyHand	-15.45	9.846	-1.569	0.1175
bodyHip	-18.62	19.3	-0.9648	0.3353
bodyKnee	-20.1	11.14	-1.804	0.07205

	Estimate	Std. Error	t value	Pr(>|t|)
bodyLumbar	-9.494	14.46	-0.6567	0.5118
bodyShoulder	-3.766	6.554	-0.5746	0.5659
bodyThoracic	4.793	8.081	0.5931	0.5535
bodyWrist	-6.359	7.091	-0.8968	0.3704
(Intercept)	-12.71	6.806	-1.867	0.06268

LEFS score as response

Table 8: Fitting linear model: outcome ~ age + gender + body

	Estimate	Std. Error	t value	Pr(>|t|)
age	-0.06717	0.03664	-1.834	0.06717
genderMALE	0.6358	1.208	0.5262	0.5989
bodyBalance	-3.384	15.87	-0.2132	0.8312
bodyCervical	-3.933	9.911	-0.3968	0.6916
bodyElbow	4.164	9.912	0.4201	0.6746
bodyFoot/Ankle	2.309	4.86	0.4752	0.6348
bodyHip	0.7018	4.954	0.1417	0.8874
bodyKnee	1.803	4.884	0.3691	0.7122
bodyLumbar	-1.285	5.106	-0.2517	0.8013
bodyShoulder	3.473	8.929	0.389	0.6974
bodyThoracic	-7.926	9.909	-0.7998	0.4241
bodyWrist	8.656	15.8	0.5479	0.5839
(Intercept)	13.11	5.017	2.613	0.009176

ODI score as response

Table 9: Fitting linear model: outcome ~ age + gender + body

	Estimate	Std. Error	t value	Pr(>|t|)
age	0.01842	0.04501	0.4094	0.6825
genderMALE	-1.158	1.315	-0.881	0.3789
bodyCervical	-5.761	5.214	-1.105	0.2699
bodyFoot/Ankle	-17.94	10.07	-1.782	0.07554
bodyHip	6.656	7.254	0.9176	0.3594
bodyKnee	1.775	8.601	0.2064	0.8366
bodyLumbar	-4.607	4.369	-1.054	0.2924
bodyThoracic	-3.9	4.73	-0.8246	0.4101
(Intercept)	-4.086	4.499	-0.9081	0.3644

NECK score as response

Table 10: Fitting linear model: outcome ~ age + gender + body

	Estimate	Std. Error	t value	Pr(>|t|)
age	0.02421	0.06621	0.3656	0.7151
genderMALE	1.673	1.824	0.9171	0.3603
bodyCervical	-0.07192	5.587	-0.01287	0.9897

	Estimate	Std. Error	t value	Pr(>|t|)
bodyHip	22.01	13.42	1.641	0.1025
bodyKnee	14.22	13.46	1.056	0.2924
bodyLumbar	-0.04491	7.431	-0.006044	0.9952
bodyShoulder	-6.228	13.37	-0.4659	0.6418
bodyThoracic	-0.2782	6.271	-0.04437	0.9647
(Intercept)	-10.62	6.216	-1.709	0.08915

KNEE score as response

Table 11: Fitting linear model: outcome ~ age + gender + body

	Estimate	Std. Error	t value	Pr(>|t|)
age	-0.07676	0.06197	-1.239	0.2168
genderMALE	-2.396	2.242	-1.069	0.2864
bodyFoot/Ankle	-22.6	18.32	-1.234	0.2186
bodyKnee	-10.87	16.58	-0.6557	0.5127
bodyLumbar	-17.09	18.94	-0.9024	0.3678
bodyShoulder	-6.564	19.1	-0.3436	0.7315
(Intercept)	34.7	17.06	2.034	0.04312

DASH Hake score as response

Table 12: Fitting linear model: HakeOutcome ~ age + gender + body

	Estimate	Std. Error	t value	Pr(>|t|)
age	0.002927	0.002873	1.019	0.3091
genderMALE	-0.02989	0.09837	-0.3038	0.7615
bodyCervical	0.02098	0.3454	0.06072	0.9516
bodyElbow	-0.09594	0.3554	-0.27	0.7873
bodyFoot/Ankle	0.2466	0.7117	0.3465	0.7292
bodyHand	-0.1889	0.4874	-0.3877	0.6985
bodyHip	-0.2719	0.9556	-0.2846	0.7761
bodyKnee	-0.1364	0.5515	-0.2473	0.8048
bodyLumbar	-0.3633	0.7158	-0.5076	0.612
bodyShoulder	0.1722	0.3246	0.5304	0.5962
bodyThoracic	0.1861	0.4	0.4653	0.642
bodyWrist	0.03955	0.351	0.1127	0.9103
(Intercept)	-0.5781	0.3373	-1.714	0.08738

LEFS Hake score as response

Table 13: Fitting linear model: HakeOutcome ~ age + gender + body

	Estimate	Std. Error	t value	Pr(>|t|)
age	-0.006217	0.002473	-2.514	0.01217
genderMALE	0.07182	0.08194	0.8766	0.3811

	Estimate	Std. Error	t value	Pr(>|t|)
bodyBalance	0.1287	1.075	0.1198	0.9047
bodyCervical	-0.7064	0.6822	-1.035	0.3009
bodyElbow	0.1775	0.6821	0.2602	0.7948
bodyFoot/Ankle	-0.07234	0.3619	-0.1999	0.8416
bodyHip	-0.1765	0.3679	-0.4796	0.6316
bodyKnee	-0.1202	0.3635	-0.3306	0.741
bodyLumbar	-0.2232	0.3774	-0.5916	0.5543
bodyShoulder	0.1059	0.6181	0.1714	0.864
bodyThoracic	-0.3436	0.6824	-0.5036	0.6147
bodyWrist	0.07956	1.069	0.07442	0.9407
(Intercept)	0.7088	0.3698	1.917	0.05569

ODI Hake score as response

Table 14: Fitting linear model: HakeOutcome ~ age + gender + body

	Estimate	Std. Error	t value	Pr(>|t|)
age	0.005173	0.002442	2.118	0.0348
genderMALE	-0.06839	0.07105	-0.9625	0.3364
bodyCervical	-0.05173	0.2801	-0.1847	0.8536
bodyFoot/Ankle	-0.665	0.5359	-1.241	0.2154
bodyHip	0.07534	0.4183	0.1801	0.8572
bodyKnee	-0.1246	0.458	-0.272	0.7858
bodyLumbar	-0.1358	0.2327	-0.5837	0.5597
bodyThoracic	-0.1811	0.2524	-0.7175	0.4735
(Intercept)	-0.3815	0.2399	-1.59	0.1126

NECK Hake score as response

Table 15: Fitting linear model: HakeOutcome ~ age + gender + body

	Estimate	Std. Error	t value	Pr(>|t|)
age	-0.002403	0.002952	-0.8138	0.4169
genderMALE	-0.008139	0.08201	-0.09924	0.9211
bodyCervical	0.1119	0.2492	0.4489	0.654
bodyHip	0.6779	0.5982	1.133	0.2587
bodyKnee	0.8124	0.6005	1.353	0.1778
bodyLumbar	0.1634	0.3314	0.493	0.6226
bodyShoulder	0.04098	0.5961	0.06876	0.9453
bodyThoracic	-0.05169	0.2797	-0.1848	0.8536
(Intercept)	-0.3566	0.2772	-1.286	0.2

KNEE Hake score as response

Table 16: Fitting linear model: HakeOutcome ~ age + gender + body

	Estimate	Std. Error	t value	Pr(>|t|)
age	-0.001003	0.001968	-0.5098	0.6107
genderMALE	-0.1992	0.0712	-2.798	0.005601
bodyFoot/Ankle	-0.06415	0.5817	-0.1103	0.9123
bodyKnee	-0.04574	0.5264	-0.08689	0.9308
bodyLumbar	-0.3629	0.6013	-0.6036	0.5467
bodyShoulder	0.2822	0.6065	0.4652	0.6422
(Intercept)	0.734	0.5416	1.355	0.1767

‘