SUPPLEMENTARY DATA

Ancestry-Specific Genetic Risk Score Improves Prediction of Type 1 Diabetes

John S. Kaddis, Ph.D.; Daniel J. Perry, Ph.D.; Anh Nguyet Vu, B.S.; Stephen S. Rich, Ph.D.; Mark A. Atkinson, Ph.D.; Desmond A. Schatz, M.D.; Bart O. Roep, M.D., Ph.D; Todd M. Brusko, Ph.D.

Notebook Dependencies

Code in this notebook relies on the use of SAS Software, which is only accessible through a paid license.

-If you have SAS, then install the SAS Kernel for Jupyter Notebooks, found here: https://github.com/sassoftware/sas_kernel
-If you do not have access to SAS, there is a free version of it, currently called "SAS OnDemand for Academics: Studio" You can find out more about that here: https://www.sas.com/en_us/software/on-demand-for-academics/references/getting-started-with-sas-ondemand-for-academics-studio.html

Regardless of your experience with or access to SAS, all of the data files used for this analysis are provided here, including the data derived from SAS.

METHODS

A. Human Organ Donors

The Network for Pancreatic Organ Donors with Diabetes (nPOD) program coordinates with many organ procurement organizations in the United States to screen and identify potential donors using acceptance criteria posted here.(1) Following acquisition of informed research consent from next of kin, pancreata, related tissues, and blood were obtained from deceased organ donors in the United States. All donations were then centrally shipped to the nPOD organ processing and pathology core at the University of Florida for biobank sharing, as previously described.(2) All experimental data was acquired under an approval from the University of Florida Institutional Review Board.

B. DNA Isolation and genotyping

DNA from snap-frozen spleen or pancreas tissue was isolated, as previously described.(2) Donors were genotyped at 974,650 unique loci using a custom SNP array termed UFDIchip, as described elsewhere.(3) In brief, the base array consists of the AxiomTM Precision Medicine Research Array (ThermoFisher Scientific), to which all content from the ImmunoChip(4) was added, as well as all previously reported credible T1D risk variants.(5) UFDIchips were processed on an Affymetrix Gene Titan instrument with external sample handling on a BioMek FX dual arm robotic workstation. Data processing and quality control procedures were performed at the SNP, sample, and plate levels using Axiom™ Analysis Suite 3.0 (ThermoFisher Scientific) set to the default stringency thresholds as recommended. An analysis of X chromosome heterozygosity found all samples to be concordant with reported sex.

C. GRS Calculation

EUR GRS was calculated as previously described(6,7) using 26 SNP genotypes extracted from UFDIchip array data and 4 from imputed data. The 4 imputed SNPs were for IL2 (rs2069762, r2 = 0.9962), HLA-A*24 (rs1264813, r2 = 0.9961), INS (rs689, r2 = 0.9486), and UBASH3A (rs3788013, r2 = 0.9967). AFR GRS was calculated as previously described (8) using 4 SNP genotypes extracted from the UFDIchip array and 3 from imputed data. The 3 imputed SNPs were for rs9271594 (r2 = 0.9498), rs34303755 (r2 = 0.8325), INS (rs689, r2 = 0.9210). The resultant datafiles are provided below. The 1000 Genomes Phase 3 dataset (version 5) was used for imputation.

In [1]:

```
/**********************
Two files contain the EUR GRS and AFR GRS data
**********************/

%let location =F:\Manuscripts\2021_06_11_Diab_Care_GRS\submission;


PROC import out=eurgrs datafile = "&location\data\EUR_GRS_nPOD.xlsx"
        DBMS = xlsx replace;
RUN;


PROC import out=afrgrs datafile = "&location\data\AFR_GRS_nPOD.xlsx"
    DBMS = xlsx replace;
RUN;
```

```
SAS Connection established. Subprocess id is 14988
```

Out[1]:

```
4                                                              The SAS System
  22:09 Thursday, September 16, 2021

29         ods listing close;ods html5 (id=saspy_internal) file=_tomods1 options(bitmap_mode='inl
ine') device=svg style=HTMLBlue;
29       ! ods graphics on / outputfmt=png;
NOTE: Writing HTML5(SASPY_INTERNAL) Body file: _TOMODS1
30
31         /**********************
32         Two files contain the EUR GRS and AFR GRS data
33         **********************/
34
35         %let location =F:\Manuscripts\2021_06_11_Diab_Care_GRS\submission;
36
37
38         PROC import out=eurgrs datafile = "&location\data\EUR_GRS_nPOD.xlsx"
39              DBMS = xlsx replace;
40         RUN;

NOTE: The import data set has 295 observations and 2 variables.
NOTE: WORK.EURGRS data set was successfully created.
NOTE: PROCEDURE IMPORT used (Total process time):
      real time           0.33 seconds
      cpu time            0.06 seconds


41
42
43         PROC import out=afrgrs datafile = "&location\data\AFR_GRS_nPOD.xlsx"
44             DBMS = xlsx replace;
45         RUN;

NOTE: The import data set has 377 observations and 8 variables.
NOTE: WORK.AFRGRS data set was successfully created.
NOTE: PROCEDURE IMPORT used (Total process time):
      real time           0.22 seconds
      cpu time            0.09 seconds


46
47
48
49         ods html5 (id=saspy_internal) close;ods listing;
50
```

D. Ancestry Analysis

Ancestry analysis was performed using ADMIXTURE v1.3.(9) The UFDIchip data was first filtered to exclude markers with high linkage disequilibrium and missingness using recommended parameters. The 1000 Genomes Phase 3 data(10) was obtained and used as the reference, with all samples and the super population labels (EUR, EAS, AMR, SAS, AFR) given as reference input to ADMIXTURE supervised training over a total of five runs. Runs were compared and confirmed to have consistent results for ancestry proportions; results reported are representative of all runs. The SAS FASTCLUS and CANDISC procedures were then used to define clusters and group individuals together based on ancestry proportions.

D.1 Admixture runs

The ADMIXTURE pipeline developed for this analysis is available as a dockerized container on GitLab at: https://gitlab.com/kaddis-lab/admixture-project

Detailed documentation on how to use the dockerized container is available at: https://kaddis-lab.gitlab.io/admixture-project/

For those without a bioinformatics background, a web-based implementation of the ADMIXTURE pipeline used for this analysis is also available on the documentation site at: https://kaddis-lab.gitlab.io/admixture-project/usage/web-app/

In [2]:

```
/**********************
Below file contains the results from the admixture runs
**********************/

PROC import out=genetics datafile = "&location\data\npod_admix_results_v1.xlsx"
        DBMS = xlsx replace;
RUN;

DATA genetics1 (keep=EUR EAS AMR SAS AFR ID corelabel);
    set genetics;
RUN;
```

Out[2]:

```
5                                                              The SAS System
  22:09 Thursday, September 16, 2021

53         ods listing close;ods html5 (id=saspy_internal) file=_tomods1 options(bitmap_mode='inl
ine') device=svg style=HTMLBlue;
53       ! ods graphics on / outputfmt=png;
NOTE: Writing HTML5(SASPY_INTERNAL) Body file: _TOMODS1
54
55         /**********************
56         Below file contains the results from the admixture runs
57         **********************/
58
59         PROC import out=genetics datafile = "&location\data\npod_admix_results_v1.xlsx"
60              DBMS = xlsx replace;
61         RUN;

NOTE: The import data set has 377 observations and 9 variables.
NOTE: WORK.GENETICS data set was successfully created.
NOTE: PROCEDURE IMPORT used (Total process time):
      real time           0.19 seconds
      cpu time            0.07 seconds


62
63         DATA genetics1 (keep=EUR EAS AMR SAS AFR ID corelabel);
64             set genetics;
65         RUN;

NOTE: There were 377 observations read from the data set WORK.GENETICS.
NOTE: The data set WORK.GENETICS1 has 377 observations and 7 variables.
NOTE: DATA statement used (Total process time):
      real time           0.03 seconds
      cpu time            0.03 seconds


66
67
68         ods html5 (id=saspy_internal) close;ods listing;
69
```

D.2 Cluster creation

In [3]:

```
/**********************
more info on fastclus found here: https://documentation.sas.com/?docsetId=statug&docsetVersion=15.1
&docsetTarget=statug_fastclus_overview.htm&locale=en
**********************/

/**********************
no standardization is needed prior to clustering, all vars measured on the same scale
**********************/

proc fastclus data=genetics1 out=Clust
              maxclusters=15 maxiter=100; /*tried clustering from 5 to 16; tried maxiter 100 and 10
00 no difference*/
   var EUR EAS AMR SAS AFR;
run;

proc candisc data=clust out=genetics2;
var EUR EAS AMR SAS AFR;
class cluster;
run;

proc sgplot data=genetics2;
   scatter y=can2 x=can1 / markerchar=cluster;
run;
```

Out[3]:

The SAS System

The FASTCLUS Procedure
Replace=FULL Radius=0 Maxclusters=15 Maxiter=100 Converge=0.02

Initial Seeds

Cluster	EUR	EAS	AMR	SAS	AFR
1	0.0000100000	0.0000100000	0.0000100000	0.0000100000	0.9999600000
2	0.0000100000	0.0000100000	0.9999590000	0.0000110000	0.0000100000
3	0.1625520000	0.0000100000	0.7935300000	0.0000100000	0.0438980000
4	0.4828870000	0.0000100000	0.3467170000	0.0000100000	0.1703760000
5	0.2980280000	0.0000100000	0.5496850000	0.0000100000	0.1522660000
6	0.0582950000	0.0000110000	0.0000100000	0.4705760000	0.4711070000
7	0.6276810000	0.0000100000	0.0000100000	0.0000100000	0.3722890000
8	0.9999600000	0.0000100000	0.0000100000	0.0000100000	0.0000100000
9	0.7404710000	0.0000100000	0.2594990000	0.0000100000	0.0000100000
10	0.0000100000	0.9999600000	0.0000100000	0.0000100000	0.0000100000
11	0.2348740000	0.0000100000	0.0000100000	0.0000100000	0.7650960000
12	0.7987390000	0.0000100000	0.0116320000	0.0119550000	0.1776650000
13	0.8075910000	0.1923790000	0.0000100000	0.0000100000	0.0000100000
14	0.3529260000	0.0000100000	0.2770860000	0.0000100000	0.3699680000
15	0.2670590000	0.1796030000	0.1042800000	0.0071950000	0.4418630000

Minimum Distance Between Initial Seeds =	0.248144

Iteration History

		Relative Change in Cluster Seeds														
Iteration	Criterion	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	0.0300	0.4133	0.0605	0.1713	0.1388	0.6162	0	0.2874	0.0295	0.1977	0.2331	0.2399	0.3745	0	0.3349	0.6998
2	0.0199	0.0396	0	0.0876	0	0.0997	0	0	0.00293	0.0565	0	0.0779	0	0	0	0.2348
3	0.0193	0.0198	0	0	0	0	0	0	0	0	0	0.0268	0	0	0	0
4	0.0193	0.0101	0	0	0	0	0	0	0	0	0	0.0434	0	0	0	0.0937
5	0.0192	0.00994	0	0	0	0	0	0	0	0	0	0.0113	0	0	0	0

Convergence criterion is satisfied.

Criterion Based on Final Seeds =	0.0192

Cluster Summary

Cluster	Frequency	RMS Std Deviation	Maximum Distance from Seed to Observation	Radius Exceeded	Nearest Cluster	Distance Between Cluster Centroids
1	29	0.0234	0.1025		11	0.1353
2	33	0.0149	0.0982		3	0.2322
3	7	0.0322	0.1032		2	0.2322
4	2	0.0218	0.0344		14	0.1720
5	6	0.0528	0.1556		3	0.2937
6	1	.	0		15	0.5638
7	10	0.0291	0.1396		15	0.2792
8	236	0.0108	0.1528		13	0.2680
9	11	0.0423	0.1402		12	0.2037
10	5	0.0278	0.0717		13	1.0975
11	23	0.0240	0.0969		1	0.1353
12	3	0.0459	0.1029		9	0.2037
13	1	.	0		12	0.2662
14	3	0.0374	0.0831		4	0.1720
15	7	0.0572	0.2310		7	0.2792

Statistics for Variables

Variable	Total STD	Within STD	R-Square	RSQ/(1-RSQ)
EUR	0.40556	0.02894	0.995098	**202.995072**
EAS	0.11003	0.01059	0.991082	**111.130912**
AMR	0.30207	0.01972	0.995898	**242.805892**
SAS	0.02627	0.01010	0.857839	**6.034272**
AFR	0.31206	0.02172	0.995336	**213.401794**
OVER-ALL	0.27052	0.01956	0.994969	**197.768896**

Pseudo F Statistic =	5113.74

Approximate Expected Over-All R-Squared =	0.82587

Cubic Clustering Criterion =	105.879

WARNING: The two values above are invalid for correlated variables.

Cluster Means

Cluster	EUR	EAS	AMR	SAS	AFR
1	0.0653712069	0.0007431724	0.0108333793	0.0013503793	0.9217018621
2	0.0081524545	0.0025689697	0.9876693636	0.0005494848	0.0010597576
3	0.1678804286	0.0000100000	0.8194165714	0.0023411429	0.0103518571
4	0.4698050000	0.0000100000	0.3316940000	0.0000100000	0.1984810000
5	0.3686113333	0.0000100000	0.6056186667	0.0003740000	0.0253860000
6	0.0582950000	0.0000110000	0.0000100000	0.4705760000	0.4711070000
7	0.5741653000	0.0000101000	0.0062287000	0.0005949000	0.4190008000
8	0.9941534195	0.0000347966	0.0018771483	0.0024323390	0.0015022797
9	0.7650419091	0.0003069091	0.2103951818	0.0026634545	0.0215925455
10	0.0172352000	0.9532676000	0.0000100000	0.0294772000	0.0000100000
11	0.1601333478	0.0000100000	0.0142990435	0.0004081739	0.8251495217
12	0.7339443333	0.0166126667	0.0752360000	0.0043060000	0.1699013333
13	0.8075910000	0.1923790000	0.0000100000	0.0000100000	0.0000100000
14	0.3431770000	0.0000100000	0.3381083333	0.0041300000	0.3145746667
15	0.3501734286	0.0274288571	0.0393080000	0.0030990000	0.5799907143

Cluster Standard Deviations

Cluster	EUR	EAS	AMR	SAS	AFR
1	0.0361894960	0.0025035835	0.0125288648	0.0047221647	0.0351678171
2	0.0211669169	0.0088175994	0.0234273593	0.0022365291	0.0060303982
3	0.0518950048	0.0000000000	0.0462351226	0.0039882841	0.0182117796
4	0.0185007418	0.0000000000	0.0212457303	0.0000000000	0.0397464722
5	0.0726515277	0.0000000000	0.0692152066	0.0008881894	0.0621582517
6	.	.	.	.	.
7	0.0470042170	0.0000003162	0.0114952333	0.0015927155	0.0433348784
8	0.0180627829	0.0002784832	0.0096866910	0.0113285148	0.0059394687
9	0.0692906149	0.0009847361	0.0519163660	0.0060712494	0.0378457466
10	0.0157775237	0.0482493518	0.0000000000	0.0358991471	0.0000000000
11	0.0369396702	0.0000000000	0.0140525194	0.0013502814	0.0361879884
12	0.0686210873	0.0269542354	0.0665269814	0.0066409874	0.0254236648
13	.	.	.	.	.
14	0.0323931174	0.0000000000	0.0597210782	0.0071360493	0.0480493102
15	0.0607694714	0.0672598086	0.0428366670	0.0035794826	0.0794337949

The SAS System

The CANDISC Procedure

Total Sample Size	377	DF Total	376
Variables	5	DF Within Classes	362
Classes	15	DF Between Classes	14

Number of Observations Read	377
Number of Observations Used	377

Class Level Information

CLUSTER	Variable Name	Frequency	Weight	Proportion
1	1	29	29.0000	0.076923
2	2	33	33.0000	0.087533
3	3	7	7.0000	0.018568
4	4	2	2.0000	0.005305
5	5	6	6.0000	0.015915
6	6	1	1.0000	0.002653
7	7	10	10.0000	0.026525
8	8	236	236.0000	0.625995
9	9	11	11.0000	0.029178
10	10	5	5.0000	0.013263
11	11	23	23.0000	0.061008
12	12	3	3.0000	0.007958
13	13	1	1.0000	0.002653
14	14	3	3.0000	0.007958
15	15	7	7.0000	0.018568

The SAS System

The CANDISC Procedure

Multivariate Statistics and F Approximations

S=4 M=4.5 N=178.5

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.00000002	2311.27	56	1398.6	<.0001
Pillai's Trace	3.83698268	608.61	56	1448	<.0001
Hotelling-Lawley Trace	614.37874811	3923.92	56	1081.5	<.0001
Roy's Greatest Root	274.97780564	7110.14	14	362	<.0001

NOTE: F Statistic for Roy's Greatest Root is an upper bound.

The SAS System

The CANDISC Procedure

					Eigenvalues of Inv(E)*H = CanRsq/(1-CanRsq)				Test of H0: The canonical correlations in the curren				
	Canonical Correlation	Adjusted Canonical Correlation	Approximate Standard Error	Squared Canonical Correlation	Eigenvalue	Difference	Proportion	Cumulative	Likelihood Ratio	Approximate F Value	Num DF	Den DF	Pr > F
1	0.998187	.	0.000187	0.996377	274.9778	47.9413	0.4476	0.4476	0.00000002	2311.27	56	1398.6	<.0001
2	0.997805	.	0.000226	0.995615	227.0365	120.5353	0.3695	0.8171	0.00000594	1564.15	39	1066.8	<.0001
3	0.995338	.	0.000480	0.990698	106.5013	100.6382	0.1733	0.9905	0.00135539	787.05	24	722	<.0001
4	0.924280	0.922408	0.007514	0.854294	5.8631	5.8631	0.0095	1.0000	0.14570636	192.95	11	362	<.0001
5	0.000000	.	0.051571	0.000000	0.0000		0.0000	1.0000	1.00000000	.	.	.	.

The SAS System

The CANDISC Procedure

Total Canonical Structure

Variable	Label	Can1	Can2	Can3	Can4
EUR	EUR	0.064097	0.989769	-0.114118	-0.056799
EAS	EAS	-0.339952	-0.078090	0.936012	-0.047068
AMR	AMR	0.749465	-0.639681	0.170594	0.002863
SAS	SAS	-0.094547	-0.007013	0.128952	0.987109
AFR	AFR	-0.680954	-0.638991	-0.357731	0.004535

Between Canonical Structure

Variable	Label	Can1	Can2	Can3	Can4
EUR	EUR	0.064138	0.990026	-0.113865	-0.052627
EAS	EAS	-0.340859	-0.078269	0.935831	-0.043700
AMR	AMR	0.749645	-0.639590	0.170148	0.002652
SAS	SAS	-0.101896	-0.007555	0.138578	0.985067
AFR	AFR	-0.681310	-0.639081	-0.356897	0.004202

Pooled Within Canonical Structure

Variable	Label	Can1	Can2	Can3	Can4
EUR	EUR	0.055108	0.936141	-0.157202	-0.309661
EAS	EAS	-0.216692	-0.054759	0.955955	-0.190252
AMR	AMR	0.704428	-0.661429	0.256909	0.017063
SAS	SAS	-0.015095	-0.001232	0.032986	0.999341
AFR	AFR	-0.600197	-0.619591	-0.505201	0.025348

The SAS System

The CANDISC Procedure

Total-Sample Standardized Canonical Coefficients

Variable	Label	Can1	Can2	Can3	Can4
EUR	EUR	12.61201386	14.88227606	3.72840173	0.08577537
EAS	EAS	-2.14352359	1.75045241	10.40889247	-0.37379322
AMR	AMR	19.76984723	-0.36011049	4.49056441	0.14994615
SAS	SAS	0.56615389	0.94084348	0.69783697	2.59076846
AFR	AFR	0.00000000	0.00000000	0.00000000	0.00000000

Pooled Within-Class Standardized Canonical Coefficients

Variable	Label	Can1	Can2	Can3	Can4
EUR	EUR	0.899941416	1.061937985	0.266043407	0.006120577
EAS	EAS	-0.206302849	0.168471820	1.001801048	-0.035975628
AMR	AMR	1.290389595	-0.023504624	0.293101789	0.009787074
SAS	SAS	0.217552733	0.361532573	0.268153842	0.995539857
AFR	AFR	0.000000000	0.000000000	0.000000000	0.000000000

Raw Canonical Coefficients

Variable	Label	Can1	Can2	Can3	Can4
EUR	EUR	31.09809400	36.69599676	9.19331273	0.21150076
EAS	EAS	-19.48039144	15.90815159	94.59625314	-3.39704139
AMR	AMR	65.44779567	-1.19214063	14.86594907	0.49639459
SAS	SAS	21.54908335	35.81060757	26.56123604	98.61044277
AFR	AFR	0.00000000	0.00000000	0.00000000	0.00000000

Class Means on Canonical Variables

CLUSTER	Can1	Can2	Can3	Can4
1	-27.21820785	-23.60826277	-8.91579993	-0.36952974
2	34.88133414	-26.87212324	5.23120501	0.01808568
3	28.92525364	-20.78671315	4.00391650	0.15371761
4	6.34393115	-9.20933580	-0.53277344	-0.25440299
5	21.13261047	-13.23625993	2.61873997	-0.10393661
6	-18.02098570	-7.06342119	3.25219818	45.89663440
7	-11.69904677	-4.97078295	-4.39616201	-0.33621302
8	1.11609780	10.51248006	-0.54862898	-0.06843852
9	7.63789164	1.86902720	0.47677526	0.00847748
10	-47.37295374	-9.20161701	81.33309590	-0.84724987
11	-24.05048698	-20.18040748	-8.08748305	-0.43818781
12	-2.45730140	1.20721442	-0.23229238	0.04138883
13	-8.60697815	6.64170734	15.83918806	-1.00109260
14	2.91462790	-13.71618356	-1.49211680	0.12827415
15	-16.97996684	-12.70396803	-3.30341046	-0.21337932

In [4]:

```
/**********************
post-cluster coding
**********************/

DATA genetics3;
    set genetics2;
    member="        ";
    if cluster=1  then member="AFR";
    if cluster=2  then member="AMR";
    if cluster=3  then member="AMRp";
    if cluster=4  then member="MIX";
    if cluster=5  then member="AMRp";
    if cluster=6  then member="MIX";
    if cluster=7  then member="EURp";
    if cluster=8  then member="EUR";
    if cluster=9  then member="EURp";
    if cluster=10 then member="EAS";
    if cluster=11 then member="AFRp";
    if cluster=12 then member="EURp";
    if cluster=13 then member="EURp";
    if cluster=14 then member="MIX";
    if cluster=15 then member="AFRp";


    alpha=.;
    if cluster=1  then alpha=1;
    if cluster=2  then alpha=1;
    if cluster=3  then alpha=.50;
    if cluster=4  then alpha=1;
    if cluster=5  then alpha=0.25;
    if cluster=6  then alpha=1;
    if cluster=7  then alpha=0.25;
    if cluster=8  then alpha=1;
    if cluster=9  then alpha=0.50;
    if cluster=10 then alpha=1;
    if cluster=11 then alpha=0.50;
    if cluster=12 then alpha=0.50;
    if cluster=13 then alpha=0.50;
    if cluster=14 then alpha=1;
    if cluster=15 then alpha=0.25;
RUN;

Data genetics3;
    set genetics3;
    if corelabel^=member then flag=1;
run;


proc export
  data=genetics3
  dbms=xlsx
  outfile="&location\data\npod_admix_results_v2.xlsx"
  replace;
run;



PROC means data=genetics3 noprint nway n;
class CLUSTER member alpha;
var Can1;
output out=genetics3_can1 mean=Can1_mean;
run;

PROC means data=genetics3 noprint nway n;
class CLUSTER member alpha;
var Can2;
output out=genetics3_can2 mean=Can2_mean;
run;


DATA genetics4;
merge genetics3_can1 genetics3_can2;
by cluster;
rename _FREQ_=count;
run;


proc export
  data=genetics4
  dbms=xlsx
  outfile="&location\data\npod_admix_results_v2_sum.xlsx"
  replace;
run;
```

Out[4]:

```
7                                                                    The SAS System
  22:09 Thursday, September 16, 2021

101        ods listing close;ods html5 (id=saspy_internal) file=_tomods1 options(bitmap_mode='inl
ine') device=svg style=HTMLBlue;
101      ! ods graphics on / outputfmt=png;
NOTE: Writing HTML5(SASPY_INTERNAL) Body file: _TOMODS1
102
103        /**********************
104        post-cluster coding
105        **********************/
106
107        DATA genetics3;
108            set genetics2;
109            member="        ";
110            if cluster=1  then member="AFR";
111            if cluster=2  then member="AMR";
112            if cluster=3  then member="AMRp";
113            if cluster=4  then member="MIX";
114            if cluster=5  then member="AMRp";
115            if cluster=6  then member="MIX";
116            if cluster=7  then member="EURp";
117            if cluster=8  then member="EUR";
118            if cluster=9  then member="EURp";
119            if cluster=10 then member="EAS";
120            if cluster=11 then member="AFRp";
121            if cluster=12 then member="EURp";
122            if cluster=13 then member="EURp";
123            if cluster=14 then member="MIX";
124            if cluster=15 then member="AFRp";
125
126
127            alpha=.;
128            if cluster=1  then alpha=1;
129            if cluster=2  then alpha=1;
130            if cluster=3  then alpha=.50;
131            if cluster=4  then alpha=1;
132            if cluster=5  then alpha=0.25;
133            if cluster=6  then alpha=1;
134            if cluster=7  then alpha=0.25;
135            if cluster=8  then alpha=1;
136            if cluster=9  then alpha=0.50;
137            if cluster=10 then alpha=1;
138            if cluster=11 then alpha=0.50;
139            if cluster=12 then alpha=0.50;
140            if cluster=13 then alpha=0.50;
141            if cluster=14 then alpha=1;
142            if cluster=15 then alpha=0.25;
143        RUN;

NOTE: There were 377 observations read from the data set WORK.GENETICS2.
NOTE: The data set WORK.GENETICS3 has 377 observations and 16 variables.
NOTE: DATA statement used (Total process time):
      real time           0.05 seconds
      cpu time            0.06 seconds



144
145        Data genetics3;
146            set genetics3;
147            if corelabel^=member then flag=1;
148        run;

NOTE: There were 377 observations read from the data set WORK.GENETICS3.
NOTE: The data set WORK.GENETICS3 has 377 observations and 17 variables.
NOTE: DATA statement used (Total process time):
      real time           0.03 seconds
      cpu time            0.03 seconds



149
150
151        proc export
152          data=genetics3
153          dbms=xlsx
154          outfile="&location\data\npod_admix_results_v2.xlsx"
155          replace;
156        run;

NOTE: The export data set has 377 observations and 17 variables.
NOTE: "F:\Manuscripts\2021_06_11_Diab_Care_GRS\submission\data\npod_admix_results_v2.xlsx" file w
as successfully created.
NOTE: PROCEDURE EXPORT used (Total process time):
      real time           2.03 seconds
      cpu time            0.14 seconds



157
158
159
160        PROC means data=genetics3 noprint nway n;
161        class CLUSTER member alpha;
162        var Can1;
163        output out=genetics3_can1 mean=Can1_mean;
164        run;

NOTE: There were 377 observations read from the data set WORK.GENETICS3.
NOTE: The data set WORK.GENETICS3_CAN1 has 15 observations and 6 variables.
NOTE: PROCEDURE MEANS used (Total process time):
      real time           0.07 seconds
      cpu time            0.06 seconds



165
166        PROC means data=genetics3 noprint nway n;
167        class CLUSTER member alpha;
168        var Can2;
169        output out=genetics3_can2 mean=Can2_mean;
170        run;

NOTE: There were 377 observations read from the data set WORK.GENETICS3.
NOTE: The data set WORK.GENETICS3_CAN2 has 15 observations and 6 variables.
NOTE: PROCEDURE MEANS used (Total process time):
      real time           0.05 seconds
      cpu time            0.06 seconds



171
172
173        DATA genetics4;
174        merge genetics3_can1 genetics3_can2;
175        by cluster;
176        rename _FREQ_=count;
177        run;

NOTE: There were 15 observations read from the data set WORK.GENETICS3_CAN1.
NOTE: There were 15 observations read from the data set WORK.GENETICS3_CAN2.
NOTE: The data set WORK.GENETICS4 has 15 observations and 7 variables.
NOTE: DATA statement used (Total process time):
      real time           0.03 seconds
      cpu time            0.04 seconds



178
179
180        proc export
181          data=genetics4
182          dbms=xlsx
183          outfile="&location\data\npod_admix_results_v2_sum.xlsx"
184          replace;
185        run;

NOTE: The export data set has 15 observations and 7 variables.
NOTE: "F:\Manuscripts\2021_06_11_Diab_Care_GRS\submission\data\npod_admix_results_v2_sum.xlsx" fi
le was successfully created.
NOTE: PROCEDURE EXPORT used (Total process time):
      real time           2.00 seconds
      cpu time            0.10 seconds



186
187
188
189        ods html5 (id=saspy_internal) close;ods listing;
190
```

D.3 Statistical Analysis

Statistical testing was performed for differences in GRSs between non-diabetic and T1D individuals within each ancestry using a two-sample t test with a pooled or Satterthwaite corrected p-value if parametric, or the Kruskal-Wallis test if non-parametric. Normality testing was performed using the Shapiro Wilks method. The Hodges-Lehmann estimation was used to obtain median differences and 95% CIs.

Testing was performed using both the EUR GRS and AFR GRS. Multiple comparison corrections are denoted with an * within the main text and are only signficant at a nominal alpha of <0.025. All reported p-values are 2-sided.

In [5]:

```
/**********************
add demographics, GRS data analysis
**********************/

data genetics3;
    set genetics3;
    caseid=ID*1;
    id1=put(ID, 4.);
RUN;

PROC import out=demographics datafile = "&location\data\Demographics_2021-05-20_13-42-23.xlsx"
        DBMS = xlsx replace;
RUN;

DATA demographics;
    set demographics;
    id1=put('nPOD CaseID'n, 4.);
    rename 'Donor Type'n=donortype;
RUN;


DATA eurgrs1;
    set eurgrs;
    id1=put('nPOD CaseID'n, 4.);
RUN;

DATA eurgrs1 (keep = id1 grs1);
    set eurgrs1;
RUN;


DATA afrgrs1;
    set afrgrs;
    id1=put(FID, 4.);
RUN;

DATA afrgrs1 (keep = id1 grs);
set afrgrs1;
RUN;


PROC sort data=genetics3; by id1; run;
PROC sort data=demographics; by id1; run;
PROC sort data=eurgrs1; by id1; run;
PROC sort data=afrgrs1; by id1; run;

DATA all;
 merge genetics3 (in=a) demographics(in=b) eurgrs1 (in=c) afrgrs1(in=d);
 by id1;
 if a;
RUN;


proc export
  data=all
  dbms=xlsx
  outfile="&location\data\data_for_figures_analysis_all.xlsx"
  replace;
run;


DATA all2;
    SET ALL;
    if donortype^="T1D" and donortype^="No Diabetes" then delete;
    if member^="AFR" and member^="EUR" and member^="AMR" then delete;
run;

proc export
  data=all2
  dbms=xlsx
  outfile="&location\data\data_for_figures_analysis_analyzed.xlsx"
  replace;
run;
```

Out[5]:

```
8                                                                    The SAS System
  22:09 Thursday, September 16, 2021

193        ods listing close;ods html5 (id=saspy_internal) file=_tomods1 options(bitmap_mode='inl
ine') device=svg style=HTMLBlue;
193      ! ods graphics on / outputfmt=png;
NOTE: Writing HTML5(SASPY_INTERNAL) Body file: _TOMODS1
194
195        /**********************
196        add demographics, GRS data analysis
197        **********************/
198
199        data genetics3;
200            set genetics3;
201            caseid=ID*1;
202            id1=put(ID, 4.);
203        RUN;

NOTE: There were 377 observations read from the data set WORK.GENETICS3.
NOTE: The data set WORK.GENETICS3 has 377 observations and 19 variables.
NOTE: DATA statement used (Total process time):
      real time           0.02 seconds
      cpu time            0.03 seconds



204
205        PROC import out=demographics datafile = "&location\data\Demographics_2021-05-20_13-42-
23.xlsx"
206             DBMS = xlsx replace;
207        RUN;

NOTE: The import data set has 653 observations and 11 variables.
NOTE: WORK.DEMOGRAPHICS data set was successfully created.
NOTE: PROCEDURE IMPORT used (Total process time):
      real time           0.46 seconds
      cpu time            0.14 seconds



208
209        DATA demographics;
210            set demographics;
211            id1=put('nPOD CaseID'n, 4.);
212            rename 'Donor Type'n=donortype;
213        RUN;

NOTE: There were 653 observations read from the data set WORK.DEMOGRAPHICS.
NOTE: The data set WORK.DEMOGRAPHICS has 653 observations and 12 variables.
NOTE: DATA statement used (Total process time):
      real time           0.02 seconds
      cpu time            0.03 seconds



214
215
216        DATA eurgrs1;
217            set eurgrs;
218            id1=put('nPOD CaseID'n, 4.);
219        RUN;

NOTE: There were 295 observations read from the data set WORK.EURGRS.
NOTE: The data set WORK.EURGRS1 has 295 observations and 3 variables.
NOTE: DATA statement used (Total process time):
      real time           0.04 seconds
      cpu time            0.04 seconds



220
221        DATA eurgrs1 (keep = id1 grs1);
222            set eurgrs1;
223        RUN;

NOTE: There were 295 observations read from the data set WORK.EURGRS1.
NOTE: The data set WORK.EURGRS1 has 295 observations and 2 variables.
NOTE: DATA statement used (Total process time):
      real time           0.03 seconds
      cpu time            0.03 seconds



224
225
226        DATA afrgrs1;
227            set afrgrs;
228            id1=put(FID, 4.);
229        RUN;

NOTE: There were 377 observations read from the data set WORK.AFRGRS.
NOTE: The data set WORK.AFRGRS1 has 377 observations and 9 variables.
NOTE: DATA statement used (Total process time):
      real time           0.03 seconds
      cpu time            0.03 seconds



230
231        DATA afrgrs1 (keep = id1 grs);
232        set afrgrs1;
233        RUN;

NOTE: There were 377 observations read from the data set WORK.AFRGRS1.
NOTE: The data set WORK.AFRGRS1 has 377 observations and 2 variables.
NOTE: DATA statement used (Total process time):
      real time           0.03 seconds
      cpu time            0.03 seconds



234
235
236        PROC sort data=genetics3; by id1; run;

NOTE: There were 377 observations read from the data set WORK.GENETICS3.
NOTE: The data set WORK.GENETICS3 has 377 observations and 19 variables.
NOTE: PROCEDURE SORT used (Total process time):
      real time           0.02 seconds
      cpu time            0.03 seconds


237        PROC sort data=demographics; by id1; run;

NOTE: There were 653 observations read from the data set WORK.DEMOGRAPHICS.
NOTE: The data set WORK.DEMOGRAPHICS has 653 observations and 12 variables.
NOTE: PROCEDURE SORT used (Total process time):
      real time           0.02 seconds
      cpu time            0.03 seconds


238        PROC sort data=eurgrs1; by id1; run;

NOTE: There were 295 observations read from the data set WORK.EURGRS1.
NOTE: The data set WORK.EURGRS1 has 295 observations and 2 variables.
NOTE: PROCEDURE SORT used (Total process time):
      real time           0.01 seconds
      cpu time            0.01 seconds


239        PROC sort data=afrgrs1; by id1; run;

NOTE: There were 377 observations read from the data set WORK.AFRGRS1.
NOTE: The data set WORK.AFRGRS1 has 377 observations and 2 variables.
NOTE: PROCEDURE SORT used (Total process time):
      real time           0.01 seconds
      cpu time            0.01 seconds



240
241        DATA all;
242         merge genetics3 (in=a) demographics(in=b) eurgrs1 (in=c) afrgrs1(in=d);
243         by id1;
244         if a;
245        RUN;

NOTE: There were 377 observations read from the data set WORK.GENETICS3.
NOTE: There were 653 observations read from the data set WORK.DEMOGRAPHICS.
NOTE: There were 295 observations read from the data set WORK.EURGRS1.
NOTE: There were 377 observations read from the data set WORK.AFRGRS1.
NOTE: The data set WORK.ALL has 377 observations and 32 variables.
NOTE: DATA statement used (Total process time):
      real time           0.03 seconds
      cpu time            0.03 seconds



246
247
248        proc export
249          data=all
250          dbms=xlsx
251          outfile="&location\data\data_for_figures_analysis_all.xlsx"
252          replace;
253        run;

NOTE: The export data set has 377 observations and 32 variables.
NOTE: "F:\Manuscripts\2021_06_11_Diab_Care_GRS\submission\data\data_for_figures_analysis_all.xls
x" file was successfully created.
NOTE: PROCEDURE EXPORT used (Total process time):
      real time           2.03 seconds
      cpu time            0.17 seconds



254
255
256        DATA all2;
257            SET ALL;
258            if donortype^="T1D" and donortype^="No Diabetes" then delete;
259            if member^="AFR" and member^="EUR" and member^="AMR" then delete;
260        run;

NOTE: There were 377 observations read from the data set WORK.ALL.
NOTE: The data set WORK.ALL2 has 207 observations and 32 variables.
NOTE: DATA statement used (Total process time):
      real time           0.05 seconds
      cpu time            0.06 seconds



261
262        proc export
263          data=all2
264          dbms=xlsx
265          outfile="&location\data\data_for_figures_analysis_analyzed.xlsx"
266          replace;
267        run;

NOTE: The export data set has 207 observations and 32 variables.
NOTE: "F:\Manuscripts\2021_06_11_Diab_Care_GRS\submission\data\data_for_figures_analysis_analyze
d.xlsx" file was successfully
      created.
NOTE: PROCEDURE EXPORT used (Total process time):
      real time           2.11 seconds
      cpu time            0.18 seconds



268
269
270        ods html5 (id=saspy_internal) close;ods listing;
271
```

In [6]:

```
/**********************
differences in EUR GRS across ancestries
GRS is the AFR GRS
GRS1 is the EUR GRS
**********************/

proc sort data=ALL2; by member; run;
proc freq data=ALL2; tables donortype*member; run;
 proc NPAR1WAY data=all2 wilcoxon hl alpha=0.05; /*HL for hodges-lehmann estimates and alpha to set
 the CIs*/
        class donortype;
        var GRS;
     by member;
RUN;
```

Out[6]:

The SAS System

The FREQ Procedure

Table of donortype by member

Frequency Percent Row Pct Col Pct	member			
donortype(Donor Type)	AFR	AMR	EUR	Total
No Diabetes	11 5.31 10.19 64.71	13 6.28 12.04 68.42	84 40.58 77.78 49.12	108 52.17
T1D	6 2.90 6.06 35.29	6 2.90 6.06 31.58	87 42.03 87.88 50.88	99 47.83
Total	17 8.21	19 9.18	171 82.61	207 100.00

The SAS System

The NPAR1WAY Procedure

member=AFR

Wilcoxon Scores (Rank Sums) for Variable GRS Classified by Variable donortype

donortype	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
T1D	6	76.0	54.0	9.949874	12.666667
No Diabetes	11	77.0	99.0	9.949874	7.000000

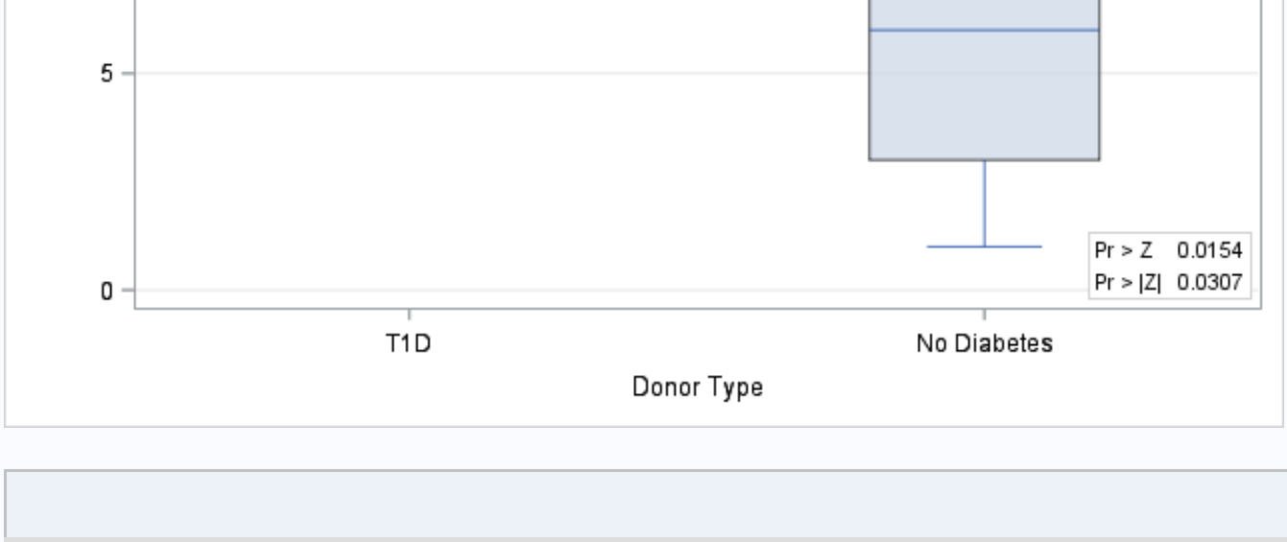
Wilcoxon Two-Sample Test

Statistic	76.0000
Normal Approximation	
Z	2.1608
One-Sided Pr > Z	0.0154
Two-Sided Pr > |Z|	0.0307
t Approximation	
One-Sided Pr > Z	0.0231
Two-Sided Pr > |Z|	0.0462

Z includes a continuity correction of 0.5.

Kruskal-Wallis Test

Chi-Square	4.8889
DF	1
Pr > Chi-Square	0.0270



Hodges-Lehmann Estimation

Location Shift (T1D - No Diabetes) 2.5370

95% Confidence Limits		Interval Midpoint	Asymptotic Standard Error
0.2640	4.9610	2.6125	1.1982

The SAS System

The NPAR1WAY Procedure

member=AMR

Wilcoxon Scores (Rank Sums) for Variable GRS Classified by Variable donortype

donortype	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
No Diabetes	13	107.0	130.0	11.396752	8.230769
T1D	6	83.0	60.0	11.396752	13.833333

Average scores were used for ties.

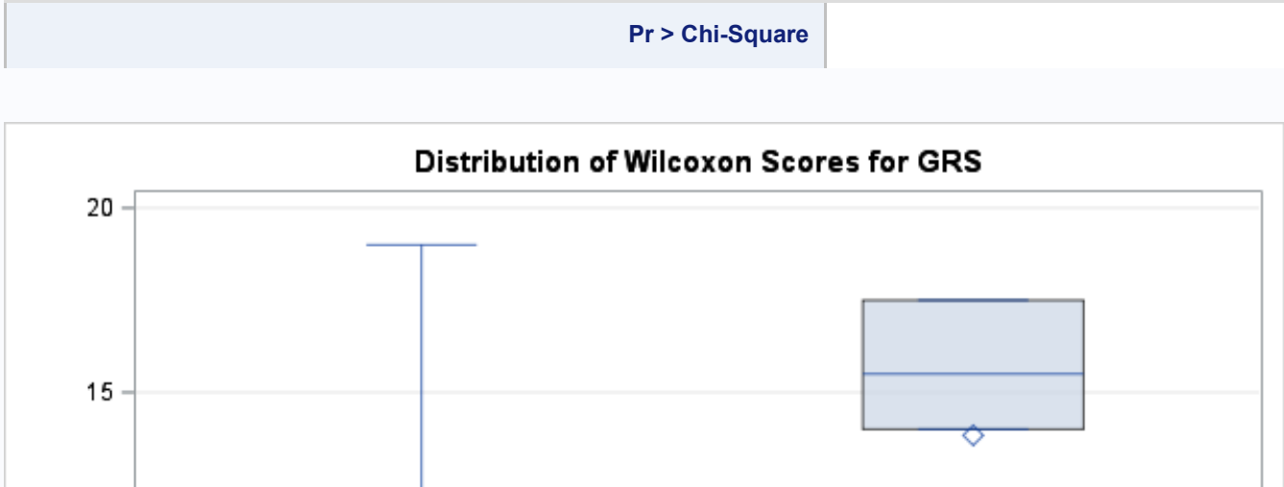
Wilcoxon Two-Sample Test

Statistic	83.0000
Normal Approximation	
Z	1.9742
One-Sided Pr > Z	0.0242
Two-Sided Pr > |Z|	0.0484
t Approximation	
One-Sided Pr > Z	0.0320
Two-Sided Pr > |Z|	0.0639

Z includes a continuity correction of 0.5.

Kruskal-Wallis Test

Chi-Square	4.0728
DF	1
Pr > Chi-Square	0.0436



Hodges-Lehmann Estimation

Location Shift (T1D - No Diabetes) 2.9280

95% Confidence Limits		Interval Midpoint	Asymptotic Standard Error
-0.1410	7.0570	3.4580	1.8363

The SAS System

The NPAR1WAY Procedure

member=EUR

Wilcoxon Scores (Rank Sums) for Variable GRS Classified by Variable donortype

donortype	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
No Diabetes	84	4884.0	7224.0	323.529481	58.142857
T1D	87	9822.0	7482.0	323.529481	112.896552

Average scores were used for ties.

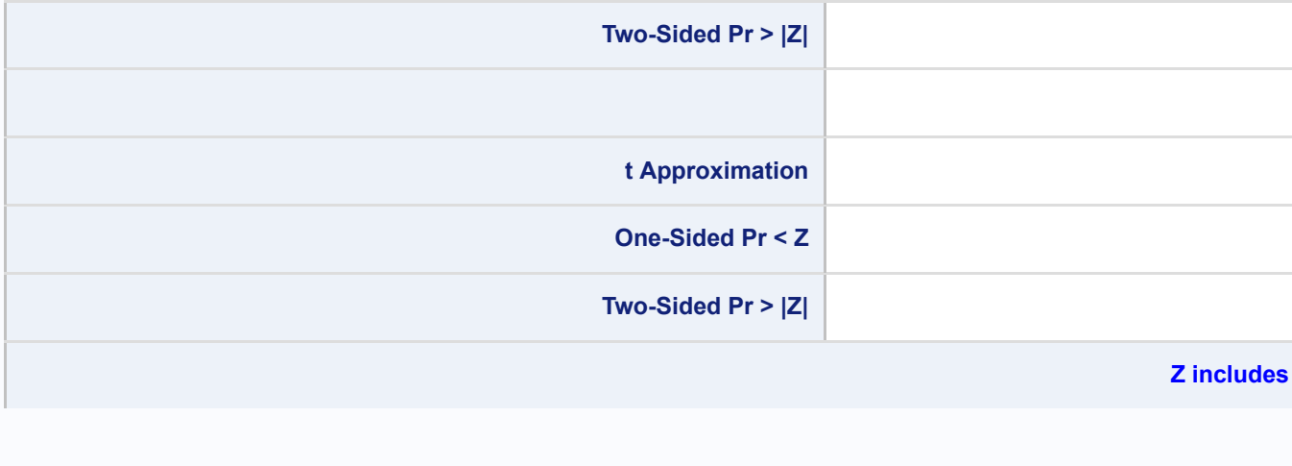
Wilcoxon Two-Sample Test

Statistic	4884.0000
Normal Approximation	
Z	-7.2312
One-Sided Pr < Z	<.0001
Two-Sided Pr > |Z|	<.0001
t Approximation	
One-Sided Pr < Z	<.0001
Two-Sided Pr > |Z|	<.0001

Z includes a continuity correction of 0.5.

Kruskal-Wallis Test

Chi-Square	52.3123
DF	1
Pr > Chi-Square	<.0001



Hodges-Lehmann Estimation

Location Shift (No Diabetes - T1D) -3.1770

95% Confidence Limits		Interval Midpoint	Asymptotic Standard Error
-4.0160	-2.5250	-3.2705	0.3804

In [13]:

```
proc freq data=all2;
tables member*donortype;
run;
```

Out[13]:

The SAS System

The FREQ Procedure

Table of member by donortype

Frequency Percent Row Pct Col Pct	donortype(Donor Type)		
member	No Diabetes	T1D	Total
AFR	11 5.31 64.71 10.19	6 2.90 35.29 6.06	17 8.21
AMR	13 6.28 68.42 12.04	6 2.90 31.58 6.06	19 9.18
EUR	84 40.58 49.12 77.78	87 42.03 50.88 87.88	171 82.61
Total	108 52.17	99 47.83	207 100.00

In [7]:

```
proc sort data=ALL2; by member; run;
proc freq data=ALL2; tables donortype*member; run;
 proc NPAR1WAY data=all2 wilcoxon hl alpha=0.05; /*HL for hodges-lehmann estimates and alpha to set
 the CIs*/
        class donortype;
        var GRS1;
     by member;
RUN;
```

Out[7]:

The SAS System

The FREQ Procedure

Table of donortype by member

Frequency Percent Row Pct Col Pct	member			
donortype(Donor Type)	AFR	AMR	EUR	Total
No Diabetes	11 5.31 10.19 64.71	13 6.28 12.04 68.42	84 40.58 77.78 49.12	108 52.17
T1D	6 2.90 6.06 35.29	6 2.90 6.06 31.58	87 42.03 87.88 50.88	99 47.83
Total	17 8.21	19 9.18	171 82.61	207 100.00

The SAS System

The NPAR1WAY Procedure

member=AFR

Wilcoxon Scores (Rank Sums) for Variable GRS1 Classified by Variable donortype

donortype	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
T1D	6	69.0	54.0	9.949874	11.500000
No Diabetes	11	84.0	99.0	9.949874	7.636364

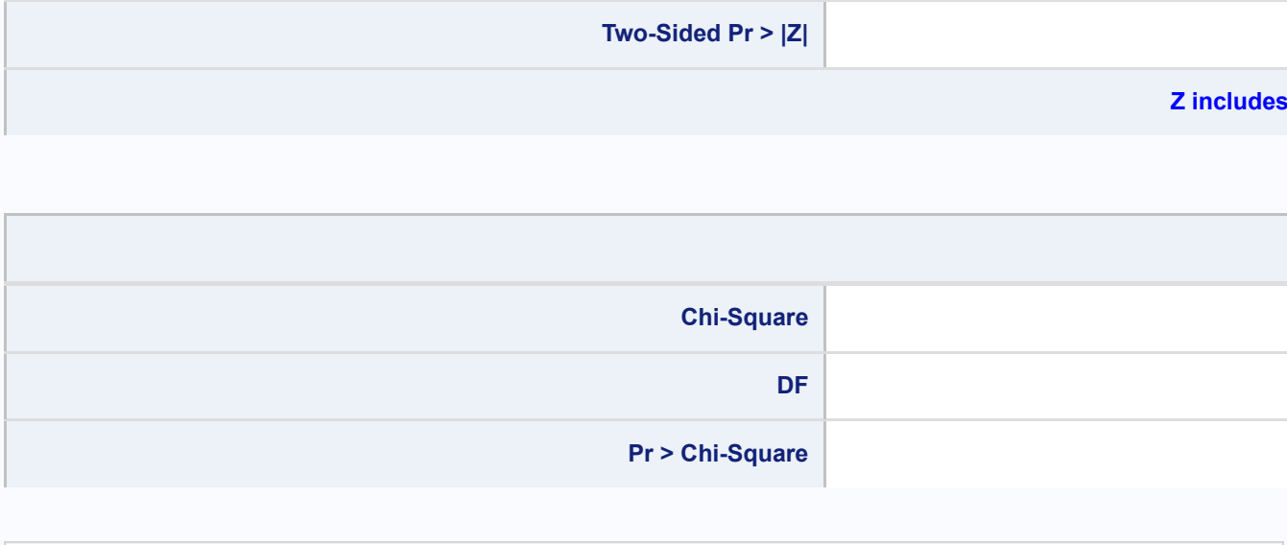
Wilcoxon Two-Sample Test

Statistic	69.0000
Normal Approximation	
Z	1.4573
One-Sided Pr > Z	0.0725
Two-Sided Pr > |Z|	0.1450
t Approximation	
One-Sided Pr > Z	0.0822
Two-Sided Pr > |Z|	0.1644

Z includes a continuity correction of 0.5.

Kruskal-Wallis Test

Chi-Square	2.2727
DF	1
Pr > Chi-Square	0.1317



Hodges-Lehmann Estimation

Location Shift (T1D - No Diabetes) 0.0146

95% Confidence Limits		Interval Midpoint	Asymptotic Standard Error
-0.0052	0.0279	0.0114	0.0084

The SAS System

The NPAR1WAY Procedure

member=AMR

Wilcoxon Scores (Rank Sums) for Variable GRS1 Classified by Variable donortype

donortype	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
No Diabetes	12	92.0	114.0	10.677078	7.666667
T1D	6	79.0	57.0	10.677078	13.166667

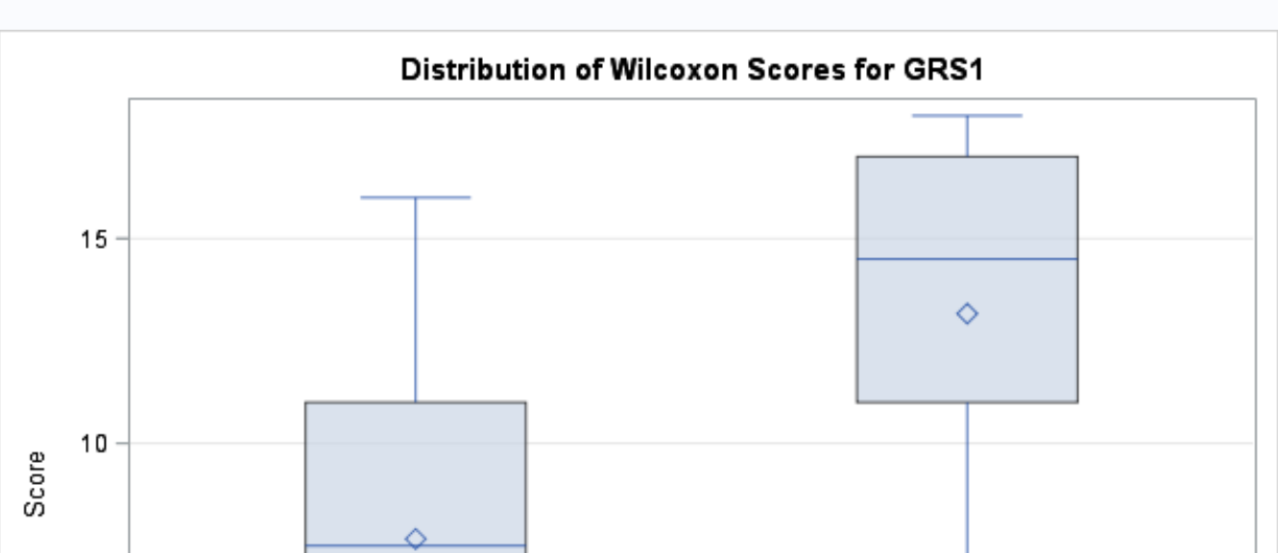
Wilcoxon Two-Sample Test

Statistic	79.0000
Normal Approximation	
Z	2.0137
One-Sided Pr > Z	0.0220
Two-Sided Pr > |Z|	0.0440
t Approximation	
One-Sided Pr > Z	0.0301
Two-Sided Pr > |Z|	0.0602

Z includes a continuity correction of 0.5.

Kruskal-Wallis Test

Chi-Square	4.2456
DF	1
Pr > Chi-Square	0.0394



Hodges-Lehmann Estimation

Location Shift (T1D - No Diabetes) 0.0399

95% Confidence Limits		Interval Midpoint	Asymptotic Standard Error
0.0012	0.0836	0.0424	0.0210

The SAS System

The NPAR1WAY Procedure

member=EUR

Wilcoxon Scores (Rank Sums) for Variable GRS1 Classified by Variable donortype

donortype	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
No Diabetes	83	4303.0	7096.50	320.779130	51.843373
T1D	87	10232.0	7438.50	320.779130	117.609195

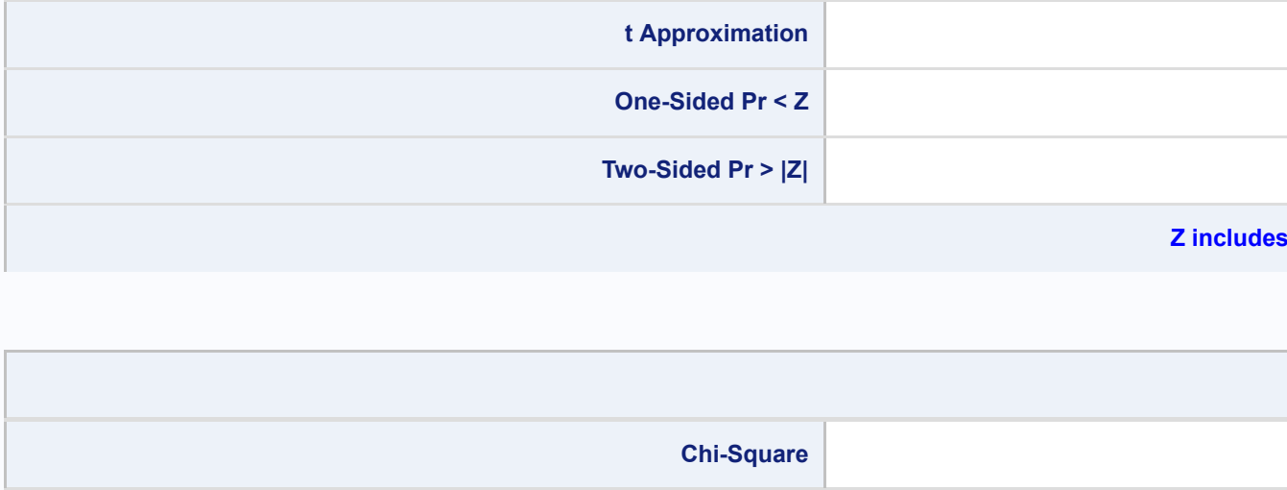
Wilcoxon Two-Sample Test

Statistic	4303.0000
Normal Approximation	
Z	-8.7069
One-Sided Pr < Z	<.0001
Two-Sided Pr > |Z|	<.0001
t Approximation	
One-Sided Pr < Z	<.0001
Two-Sided Pr > |Z|	<.0001

Z includes a continuity correction of 0.5.

Kruskal-Wallis Test

Chi-Square	75.8377
DF	1
Pr > Chi-Square	<.0001



Hodges-Lehmann Estimation

Location Shift (No Diabetes - T1D) -0.0541

95% Confidence Limits		Interval Midpoint	Asymptotic Standard Error
-0.0654	-0.0443	-0.0549	0.0054

In [8]:
```
DATA special_comp;
set all2;
if member='AMR' then delete;
if donortype="No Diabetes" then delete;
run;

proc NPAR1WAY data=special_comp wilcoxon hl alpha=0.05; /*HL for hodges-lehmann estimates and alpha to set the CIs*/
        class member;
        var GRS1;
RUN;


proc NPAR1WAY data=special_comp wilcoxon hl alpha=0.05; /*HL for hodges-lehmann estimates and alpha to set the CIs*/
        class member;
        var GRS;
RUN;
```

Out[8]:

The SAS System

The NPAR1WAY Procedure

Wilcoxon Scores (Rank Sums) for Variable GRS1
Classified by Variable member

member	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
AFR	6	47.0	282.0	63.945289	7.833333
EUR	87	4324.0	4089.0	63.945289	49.701149

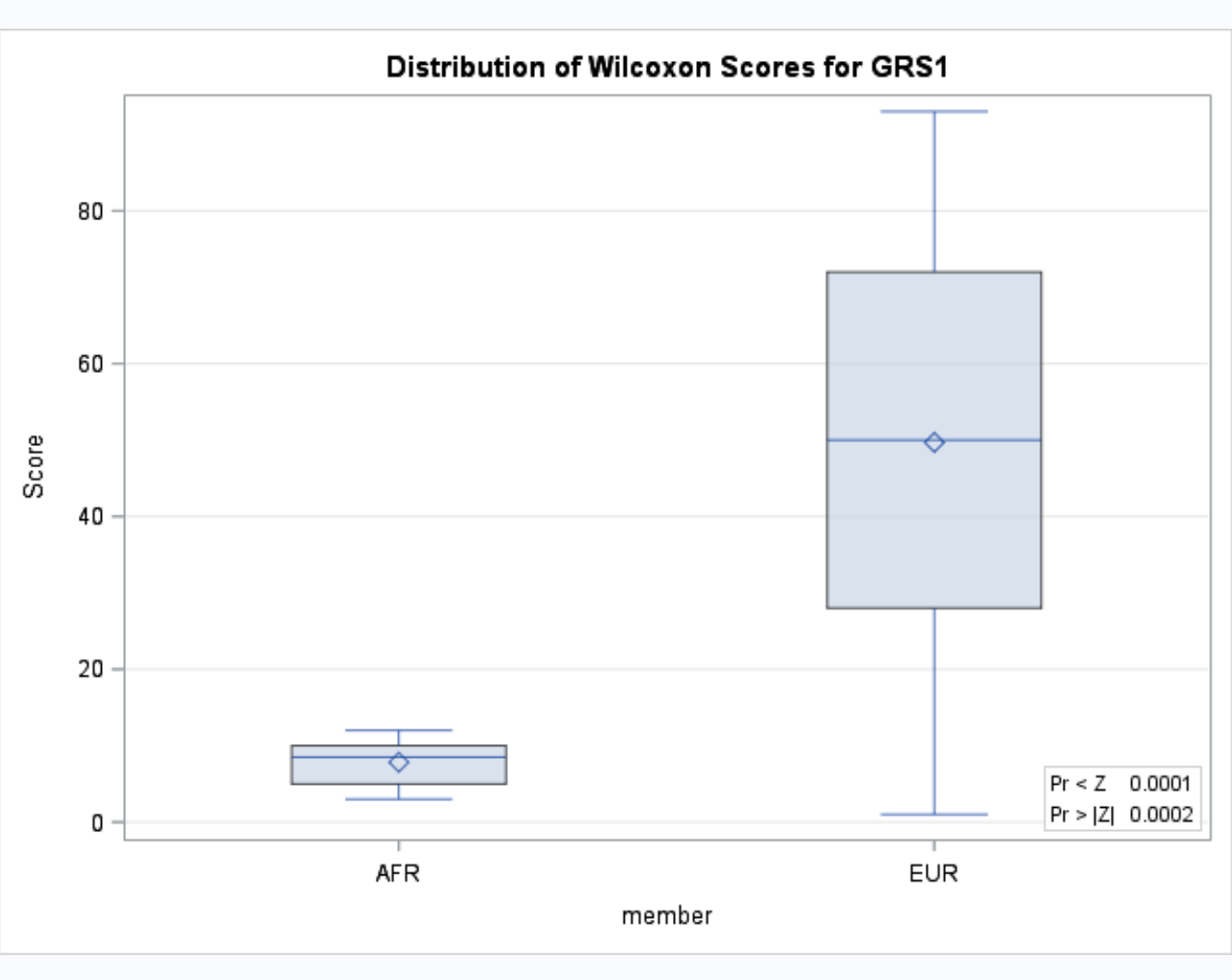
Wilcoxon Two-Sample Test

Statistic	47.0000
Normal Approximation	
Z	-3.6672
One-Sided Pr < Z	0.0001
Two-Sided Pr > |Z|	0.0002
t Approximation	
One-Sided Pr < Z	0.0002
Two-Sided Pr > |Z|	0.0004

Z includes a continuity correction of 0.5.

Kruskal-Wallis Test

Chi-Square	13.5057
DF	1
Pr > Chi-Square	0.0002



Hodges-Lehmann Estimation

Location Shift (AFR - EUR) -0.0463

95% Confidence Limits		Interval Midpoint	Asymptotic Standard Error
-0.0634	-0.0273	-0.0453	0.0092

The SAS System

The NPAR1WAY Procedure

Wilcoxon Scores (Rank Sums) for Variable GRS
Classified by Variable member

member	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
AFR	6	170.0	282.0	63.860796	28.333333
EUR	87	4201.0	4089.0	63.860796	48.287356

Average scores were used for ties.

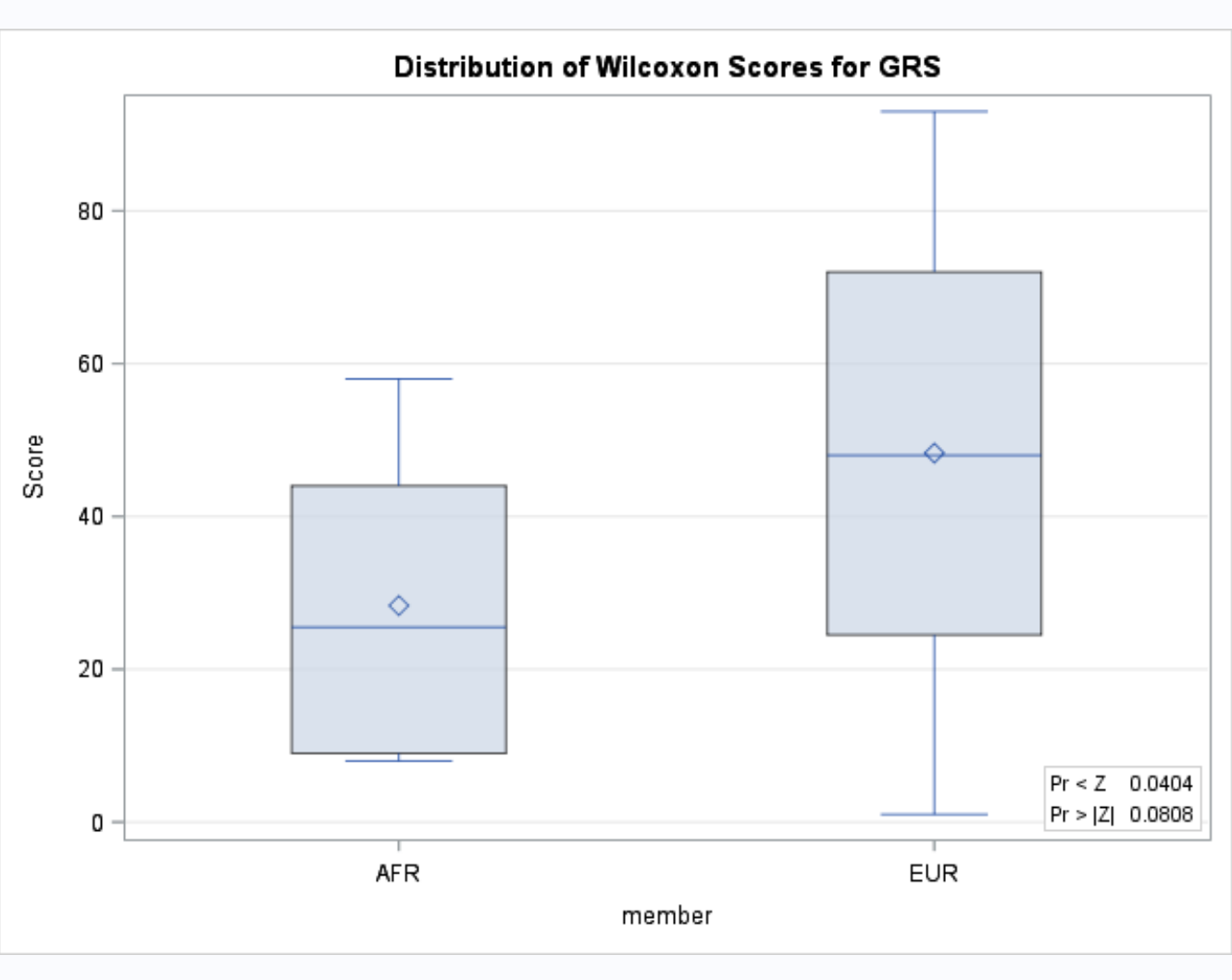
Wilcoxon Two-Sample Test

Statistic	170.0000
Normal Approximation	
Z	-1.7460
One-Sided Pr < Z	0.0404
Two-Sided Pr > |Z|	0.0808
t Approximation	
One-Sided Pr < Z	0.0421
Two-Sided Pr > |Z|	0.0842

Z includes a continuity correction of 0.5.

Kruskal-Wallis Test

Chi-Square	3.0759
DF	1
Pr > Chi-Square	0.0795



Hodges-Lehmann Estimation

Location Shift (AFR - EUR) -1.6880

95% Confidence Limits		Interval Midpoint	Asymptotic Standard Error
-3.5800	0.2470	-1.6665	0.9763

REFERENCES

1. Pugliese, A., et al. The Juvenile Diabetes Research Foundation Network for Pancreatic Organ Donors with Diabetes (nPOD) Program: goals, operational model and emerging findings. Pediatr Diabetes 15, 1-9 (2014).
2. Campbell-Thompson, M., et al. Network for Pancreatic Organ Donors with Diabetes (nPOD): developing a tissue biobank for type 1 diabetes. Diabetes Metab Res Rev 28, 608-617 (2012).
3. Carr, A.L.J., et al. Histological validation of a type 1 diabetes clinical diagnostic model for classification of diabetes. Diabetic Medicine 37, 2160-2168 (2020).
4. Cortes, A. & Brown, M.A. Promise and pitfalls of the Immunochip. Arthritis Research & Therapy 13, 101 (2011).
5. Onengut-Gumuscu, S., et al. Fine mapping of type 1 diabetes susceptibility loci and evidence for colocalization of causal variants with lymphoid gene enhancers. Nature Genetics 47, 381-386 (2015).
6. Type 1 Diabetes Genetic Risk Score: A Novel Tool to Discriminate Monogenic and Type 1 Diabetes. Diabetes 65, 2094-2099 (2016).
7. Oram, R.A., et al. A Type 1 Diabetes Genetic Risk Score Can Aid Discrimination Between Type 1 and Type 2 Diabetes in Young Adults. Diabetes Care 39, 337-344 (2016).
8. Onengut-Gumuscu, S., et al. Type 1 Diabetes Risk in African-Ancestry Participants and Utility of an Ancestry-Specific Genetic Risk Score. Diabetes Care 42, 406-415 (2019).
9. Alexander, D.H., Novembre, J. & Lange, K. Fast model-based estimation of ancestry in unrelated individuals. Genome Res 19, 1655-1664 (2009).
10. Auton, A., et al. A global reference for human genetic variation. Nature 526, 68-74 (2015).