



Duke Clinical Research Institute

Duke BioLINCC Data Harmonization Project

Ann Marie Navar, MD PhD

Fellow, Cardiovascular Disease

Our Team

- **Michael Pencina PhD**, DCRI Director of Biostatistics
- **Eric Peterson MD MPH**, DCRI Executive Director
- **Tony Shibler, MS**, Programmer
- **Brian Tinga, MS**, Programmer
- **Daniel Wojdyla, MS**, Study statistician



DCRI's Utilization of NHLBI Cohorts

- **Initial study: Temporal changes in risk factor burden**
- **Our approach to harmonization**
- **Challenges and Opportunities**

First Use of Pooled NHLBI Datasets

- **Study question:** How has the attributable risk of modifiable risk factors changed over time?

First Use of Pooled NHLBI Datasets

- **Study question:** How has the attributable risk of modifiable risk factors changed over time?
- **Cohorts:** Framingham Offspring, Framingham Original, MESA, ARIC, CHS

First Use of Pooled NHLBI Datasets

- **Study question:** How has the attributable risk of modifiable risk factors changed over time?
- **Cohorts:** Framingham Offspring, Framingham Original, MESA, ARIC, CHS
- **Analytic Approach:**
 - Two matched cohorts free of CVD created before/after 2000
 - PAF of CHD due to smoking, lipids, SBP, and diabetes calculated in each era

DCRI Experience in Data Harmonization

Clinical Trials Data Pooling

- >130,000 patients in ACS trials

Sex differences in outcomes after ACS

Adverse events in ACS trials

Approach to Harmonization

1. Requisite variables specified
2. Documentation reviewed to identify possible candidates
3. Candidate variables reviewed and classified
4. Six datasets created

Key Info

Events

Medical history

Labs

Medications

Vitals



Approach to Harmonization: Documentation

medhist_documentation - Microsoft Excel

File Home Insert Page Layout Formulas Data Review View

Clipboard Font Alignment Number Styles Cells Editing

G95 1059

	A	B	C	D	E	F	G	H	I
1	registry	mhterm	mhterm_o	mhtime_c	srcdom	srcvar	count	passed_qc	
2	ATHEROSCLEROSIS RISK IN COMMUN	Diabetes (ADA/MEDS)	DIABETES W/ FASTING GLUCOSE CUTPT.	BASELINE - VIS	/DATA/V1/DERIVE13.SAS7BDAT	DIABTS03	15040	YES	
3	ATHEROSCLEROSIS RISK IN COMMUN	Diabetes (ADA/MEDS)	DIABTS22 USING CHMB07=126	VISIT 2	/DATA/V1/DERIVE2_10.SAS7BDAT	DIABTS23	13540	YES	
4	ATHEROSCLEROSIS RISK IN COMMUN	Diabetes (ADA/MEDS)	DIABETES USING LOWER CUTPOINT 126 M	VISIT 3	/DATA/V1/DERIVE37.SAS7BDAT	DIABTS34	12118	YES	
5	ATHEROSCLEROSIS RISK IN COMMUN	Diabetes (ADA/MEDS)	DIABETES USING LOWER CUTPOINT 126 M	VISIT 4	/DATA/V1/DERIVE47.SAS7BDAT	DIABTS42	10912	YES	
6	CARDIOVASCULAR HEALTH STUDY (C	Diabetes (ADA/MEDS)	DIABETIC STATUS (ADA)	BASELINE	/DATA/BASELINE/BASEBOTHFINAL.SAS7BDAT	DIABADA	5888	YES	
7	CARDIOVASCULAR HEALTH STUDY (C	Diabetes (ADA/MEDS)	ADA DIABETIC STATUS	VISIT 3	/DATA/YEAR5/YR5OLDFINAL.SAS7BDAT	DIABADA	5201	YES	
8	CARDIOVASCULAR HEALTH STUDY (C	Diabetes (ADA/MEDS)	ADA DIABETIC STATUS	VISIT 7	/DATA/YEAR9/YR9FINAL.SAS7BDAT	DIABADA	5888	YES	
9	FRAMINGHAM-OFFSPRING	Diabetes (ADA/MEDS)	DIABETES MELLITUS STATUS	BASELINE - EXA	/DATA/LDIA1_7.SAS7BDAT	DIAB1	5013	YES	
10	FRAMINGHAM-OFFSPRING	Diabetes (ADA/MEDS)	DIABETES MELLITUS STATUS	EXAM 2	/DATA/LDIA1_7.SAS7BDAT	DIAB2	5013	YES	
11	FRAMINGHAM-OFFSPRING	Diabetes (ADA/MEDS)	DIABETES MELLITUS STATUS	EXAM 3	/DATA/LDIA1_7.SAS7BDAT	DIAB3	5013	YES	
12	FRAMINGHAM-OFFSPRING	Diabetes (ADA/MEDS)	DIABETES MELLITUS STATUS	EXAM 4	/DATA/LDIA1_7.SAS7BDAT	DIAB4	5013	YES	
13	FRAMINGHAM-OFFSPRING	Diabetes (ADA/MEDS)	DIABETES MELLITUS STATUS	EXAM 5	/DATA/LDIA1_7.SAS7BDAT	DIAB5	5013	YES	
14	FRAMINGHAM-OFFSPRING	Diabetes (ADA/MEDS)	DIABETES MELLITUS STATUS	EXAM 6	/DATA/LDIA1_7.SAS7BDAT	DIAB6	5013	YES	
15	FRAMINGHAM-OFFSPRING	Diabetes (ADA/MEDS)	DIABETES MELLITUS STATUS	EXAM 7	/DATA/LDIA1_7.SAS7BDAT	DIAB7	5013	YES	
16	FRAMINGHAM-ORIGINAL	Diabetes (ADA/MEDS)	DIABETES MELLITUS STATUS	EXAM 11	/DATA/VR_DIAB_EX28_0_0601D.SAS7BDAT	BG200_CU	5079	YES	
17	FRAMINGHAM-ORIGINAL	Diabetes (ADA/MEDS)	DIABETES MELLITUS STATUS	EXAM 12	/DATA/VR_DIAB_EX28_0_0601D.SAS7BDAT	BG200_CU	5079	YES	
18	FRAMINGHAM-ORIGINAL	Diabetes (ADA/MEDS)	DIABETES MELLITUS STATUS	EXAM 13	/DATA/VR_DIAB_EX28_0_0601D.SAS7BDAT	BG200_CU	5079	YES	
19	FRAMINGHAM-ORIGINAL	Diabetes (ADA/MEDS)	DIABETES MELLITUS STATUS	EXAM 14	/DATA/VR_DIAB_EX28_0_0601D.SAS7BDAT	BG200_CU	5079	YES	
20	FRAMINGHAM-ORIGINAL	Diabetes (ADA/MEDS)	DIABETES MELLITUS STATUS	EXAM 15	/DATA/VR_DIAB_EX28_0_0601D.SAS7BDAT	BG200_CU	5079	YES	
21	FRAMINGHAM-ORIGINAL	Diabetes (ADA/MEDS)	DIABETES MELLITUS STATUS	EXAM 16	/DATA/VR_DIAB_EX28_0_0601D.SAS7BDAT	BG200_CU	5079	YES	
22	FRAMINGHAM-ORIGINAL	Diabetes (ADA/MEDS)	DIABETES MELLITUS STATUS	EXAM 17	/DATA/VR_DIAB_EX28_0_0601D.SAS7BDAT	BG200_CU	5079	YES	
23	FRAMINGHAM-ORIGINAL	Diabetes (ADA/MEDS)	DIABETES MELLITUS STATUS	EXAM 18	/DATA/VR_DIAB_EX28_0_0601D.SAS7BDAT	BG200_CU	5079	YES	
24	FRAMINGHAM-ORIGINAL	Diabetes (ADA/MEDS)	DIABETES MELLITUS STATUS	EXAM 19	/DATA/VR_DIAB_EX28_0_0601D.SAS7BDAT	BG200_CU	5079	YES	
25	FRAMINGHAM-ORIGINAL	Diabetes (ADA/MEDS)	DIABETES MELLITUS STATUS	EXAM 20	/DATA/VR_DIAB_EX28_0_0601D.SAS7BDAT	BG200_CU	5079	YES	
26	FRAMINGHAM-ORIGINAL	Diabetes (ADA/MEDS)	DIABETES MELLITUS STATUS	EXAM 21	/DATA/VR_DIAB_EX28_0_0601D.SAS7BDAT	BG200_CU	5079	YES	
27	FRAMINGHAM-ORIGINAL	Diabetes (ADA/MEDS)	DIABETES MELLITUS STATUS	EXAM 22	/DATA/VR_DIAB_EX28_0_0601D.SAS7BDAT	BG200_CU	5079	YES	
28	FRAMINGHAM-ORIGINAL	Diabetes (ADA/MEDS)	DIABETES MELLITUS STATUS	EXAM 23	/DATA/VR_DIAB_EX28_0_0601D.SAS7BDAT	BG200_CU	5079	YES	
29	FRAMINGHAM-ORIGINAL	Diabetes (ADA/MEDS)	DIABETES MELLITUS STATUS	EXAM 24	/DATA/VR_DIAB_EX28_0_0601D.SAS7BDAT	BG200_CU	5079	YES	

Approach to Harmonization: Documentation

1	A	B	C	D	E	F	G	H
	REGISTRY	VSTEST	UNIT	VISIT	VAL MISS NONMISS	VAL MEAN MEDIAN Q1 Q3	VAL MIN MAX	DAYS MEAN MED
202	FRAMINGHAM-ORIGINAL	Systolic Blood Pressure	MM HG	EXAM 24	(27, 686)	140.05, 138.00 (123.00 - 154.00)	(91.00 - 225.00)	16947.79, 16954.00 (16779.00 - 17118.00)
203	FRAMINGHAM-ORIGINAL	Systolic Blood Pressure	MM HG	EXAM 25	(6, 665)	139.58, 138.00 (125.00 - 153.00)	(87.00 - 212.00)	17649.28, 17644.00 (17452.00 - 17838.00)
204	FRAMINGHAM-ORIGINAL	Systolic Blood Pressure	MM HG	EXAM 26	(4, 538)	138.85, 138.00 (125.00 - 150.00)	(87.00 - 235.00)	18402.96, 18416.00 (18175.00 - 18648.00)
205	FRAMINGHAM-ORIGINAL	Systolic Blood Pressure	MM HG	EXAM 27	(4, 407)	139.11, 137.00 (124.00 - 152.00)	(86.00 - 227.00)	19055.00, 19028.00 (18812.00 - 19244.00)
206	FRAMINGHAM-ORIGINAL	Systolic Blood Pressure	MM HG	EXAM 28	(5, 297)	135.56, 133.00 (120.00 - 149.00)	(87.00 - 215.00)	19754.83, 19734.00 (19490.00 - 20000.00)
207	MULTI-ETHNIC STUDY OF ATHEROSCLEROSIS (MESA)	Systolic Blood Pressure	MM HG	BASELINE	(3, 6811)	126.59, 123.50 (111.00 - 140.00)	(67.00 - 230.50)	0.00, 0.00 (0.00 - 0.00)
208	MULTI-ETHNIC STUDY OF ATHEROSCLEROSIS (MESA)	Systolic Blood Pressure	MM HG	EXAM 2	(584, 6230)	124.42, 121.00 (109.50 - 137.50)	(60.00 - 230.00)	597.78, 590.00 (504.00 - 672.00)
209	MULTI-ETHNIC STUDY OF ATHEROSCLEROSIS (MESA)	Systolic Blood Pressure	MM HG	EXAM 3	(878, 5936)	123.30, 119.50 (108.50 - 136.50)	(61.50 - 210.50)	1166.93, 1144.00 (1089.00 - 1213.00)
210	MULTI-ETHNIC STUDY OF ATHEROSCLEROSIS (MESA)	Systolic Blood Pressure	MM HG	EXAM 4	(1116, 5698)	123.68, 120.50 (109.00 - 136.00)	(66.50 - 245.00)	1763.73, 1753.00 (1680.00 - 1837.00)
211	ATHEROSCLEROSIS RISK IN COMMUNITIES (ARIC)	Weight, Truncation	KG	BASELINE - VISIT 1	(16, 14925)	78.71, 77.57 (66.68 - 88.91)	(36.29 - 181.89)	0.00, 0.00 (0.00 - 0.00)
212	ATHEROSCLEROSIS RISK IN COMMUNITIES (ARIC)	Weight, Truncation	KG	VISIT 2	(18, 13519)	79.45, 78.02 (67.13 - 89.36)	(32.66 - 180.53)	1069.12, 1063.00 (1022.00 - 1110.00)
213	ATHEROSCLEROSIS RISK IN COMMUNITIES (ARIC)	Weight, Truncation	KG	VISIT 3	(0, 12100)	80.44, 78.93 (68.04 - 90.72)	(33.57 - 188.70)	2181.72, 2180.00 (2122.00 - 2239.00)
214	ATHEROSCLEROSIS RISK IN COMMUNITIES (ARIC)	Weight, Truncation	KG	VISIT 4	(0, 10880)	81.00, 79.38 (68.49 - 91.63)	(35.83 - 184.16)	3266.04, 3248.00 (3219.00 - 3297.00)
215	(CHS)	Weight, Truncation	KG	BASELINE	(111, 5777)	72.46, 71.67 (62.14 - 81.74)	(50.80 - 97.98)	0.00, 0.00 (0.00 - 0.00)
216	(CHS)	Weight, Truncation	KG	VISIT 1	(1103, 4785)	72.30, 71.67 (61.69 - 81.65)	(50.80 - 97.98)	353.99, 350.00 (342.00 - 359.00)
217	(CHS)	Weight, Truncation	KG	VISIT 2	(1398, 4490)	72.15, 71.21 (61.69 - 81.65)	(50.35 - 97.98)	722.92, 718.50 (710.00 - 729.00)
218	(CHS)	Weight, Truncation	KG	VISIT 3	(1007, 4194)	71.66, 71.21 (61.46 - 80.74)	(49.90 - 97.07)	1070.58, 1069.00 (1052.00 - 1087.00)
219	(CHS)	Weight, Truncation	KG	VISIT 4	(1492, 4396)	72.62, 72.12 (62.14 - 81.65)	(50.35 - 98.88)	1299.11, 1425.00 (1384.00 - 1450.00)
220	(CHS)	Weight, Truncation	KG	VISIT 5	(1717, 4171)	72.09, 71.67 (61.92 - 81.19)	(49.90 - 97.98)	1663.30, 1782.00 (1750.00 - 1814.00)
221	(CHS)	Weight, Truncation	KG	VISIT 6	(2046, 3842)	71.96, 71.21 (61.69 - 80.74)	(49.90 - 97.98)	2029.28, 2154.00 (2123.00 - 2189.00)
222	(CHS)	Weight, Truncation	KG	VISIT 7	(2379, 3509)	71.83, 71.21 (61.69 - 80.74)	(49.90 - 97.98)	2394.35, 2525.00 (2485.00 - 2634.00)
223	(CHS)	Weight, Truncation	KG	VISIT 8	(2696, 3192)	71.66, 70.99 (61.69 - 80.29)	(49.90 - 97.98)	2748.54, 2878.00 (2843.00 - 2913.00)
224	(CHS)	Weight, Truncation	KG	VISIT 9	(2859, 3029)	70.98, 70.31 (61.24 - 79.83)	(49.40 - 97.07)	3127.30, 3252.00 (3214.00 - 3290.00)
225	(CHS)	Weight, Truncation	KG	VISIT 16	(4819, 1069)	69.73, 68.95 (59.65 - 78.47)	(34.93 - 131.32)	5642.79, 5777.00 (5739.00 - 5815.00)
226	FRAMINGHAM-OFFSPRING	Weight, Truncation	KG	BASELINE - EXAM 1	(0, 5013)	72.91, 72.58 (61.24 - 83.92)	(45.36 - 120.20)	0.00, 0.00 (0.00 - 0.00)
227	FRAMINGHAM-OFFSPRING	Weight, Truncation	KG	EXAM 2	(0, 3773)	74.46, 72.58 (61.24 - 83.92)	(45.36 - 120.20)	2882.12, 2912.00 (2747.00 - 2969.00)
228	FRAMINGHAM-OFFSPRING	Weight, Truncation	KG	EXAM 3	(36, 3748)	76.27, 74.84 (63.50 - 86.18)	(45.36 - 120.20)	4494.04, 4516.00 (4341.00 - 4669.00)
229	FRAMINGHAM-OFFSPRING	Weight, Truncation	KG	EXAM 4	(10, 3915)	77.61, 77.11 (65.77 - 88.45)	(45.36 - 120.20)	5751.40, 5721.00 (5656.00 - 5846.00)
230	FRAMINGHAM-OFFSPRING	Weight, Truncation	KG	EXAM 5	(17, 3694)	78.81, 77.11 (65.77 - 88.45)	(45.36 - 120.20)	7086.45, 7084.00 (6995.00 - 7175.00)
231	FRAMINGHAM-OFFSPRING	Weight, Truncation	KG	EXAM 6	(18, 3416)	79.95, 79.38 (68.04 - 90.72)	(45.36 - 120.20)	8564.39, 8595.00 (8479.00 - 8670.00)
232	FRAMINGHAM-OFFSPRING	Weight, Truncation	KG	EXAM 7	(31, 3461)	80.20, 79.38 (68.04 - 90.72)	(45.36 - 120.20)	9649.36, 9655.00 (9366.00 - 9932.00)
233	FRAMINGHAM-OFFSPRING	Weight, Truncation	KG	EXAM 8	(10, 2991)	78.96, 76.20 (64.86 - 89.81)	(45.36 - 131.54)	12052.50, 12071.00 (11718.00 - 12405.00)
234	FRAMINGHAM-ORIGINAL	Weight, Truncation	KG	EXAM 11	(22, 2830)	71.85, 70.31 (61.24 - 79.38)	(45.36 - 124.74)	7272.33, 7282.00 (7226.00 - 7328.00)
235	FRAMINGHAM-ORIGINAL	Weight, Truncation	KG	EXAM 12	(2, 2195)	71.70, 70.31 (61.24 - 79.38)	(45.36 - 124.74)	9899.27, 9924.00 (7960.00 - 11890.00)

Challenges

1. Variability in documentation style across cohorts
2. Differences in event classification
 - Composite endpoints vary
 - Differential disease classification: e.g. PAD, CHF
3. De-identification of CARDIA → too few events to include
4. Slight differences in n between NHLBI and cohorts

Opportunities

- New projects using pooled data underway
- Infrastructure allows rapid incorporation of new cohorts and variables
- Collaboration with CCC

DCRI Commitment to Open Science

- Available resources: 90+ statisticians/programmers and 18 PhD level statistical faculty
- Data coordinating center for multiple national registries: NCDR, GWTG, STS
- Improving access to clinical trials data
- Opening up clinical data to outside investigators



SOAR

SUPPORTING OPEN ACCESS FOR RESEARCHERS





Thank you