

# Harnessing Qatar Biobank to Understand Type 2 Diabetes and Obesity in Adult Qataris from the First Qatar Biobank Project

Ehsan Ullah, Raghvendra Mall, Reda Rawi, Naima M Moustaid,  
Adeel A Butt, Halima Bensmail

## Gender Stratified Analysis

### Baseline Statistics

Table 1: Baseline characteristics for males in diabetes study. (Rows are sorted by significance)

	Diabetic ( $n = 160$ )	Non-Diabetic ( $n = 454$ )	P-value
Age	$48.92 \pm 10.69$	$38.51 \pm 11.51$	$5.54 \times 10^{-22}$
Magnesium	$0.8 \pm 0.07$	$0.85 \pm 0.06$	$4.92 \times 10^{-18}$
Chloride	$99.17 \pm 2.5$	$101.02 \pm 1.93$	$2.72 \times 10^{-15}$
Uric Acid	$298.29 \pm 78.49$	$351.46 \pm 67.79$	$5.25 \times 10^{-13}$
Sodium	$139.28 \pm 2.37$	$140.74 \pm 2.13$	$1.33 \times 10^{-12}$
Albumin	$45.12 \pm 2.6$	$46.46 \pm 2.56$	$2.24 \times 10^{-08}$
Free triiodothyronine	$4.4 \pm 0.58$	$4.68 \pm 0.59$	$1.81 \times 10^{-07}$
Folate Serum	$25.84 \pm 7.5$	$23.3 \pm 7.81$	$3.79 \times 10^{-04}$
High Density Lipoprotein (HDL-C)	$1.1 \pm 0.29$	$1.19 \pm 0.29$	$5.70 \times 10^{-04}$
Triglycerides	$1.93 \pm 1.09$	$1.58 \pm 1.19$	$9.18 \times 10^{-04}$
Total Bilirubin	$7.18 \pm 3.41$	$8.31 \pm 5.06$	$1.74 \times 10^{-03}$
Homocysteine	$9.04 \pm 2.66$	$9.85 \pm 3.42$	$2.27 \times 10^{-03}$
Insulin	$26.11 \pm 32.82$	$17.36 \pm 24.6$	$2.33 \times 10^{-03}$
Serum Creatinine	$74.53 \pm 13.72$	$79.87 \pm 33.35$	$5.17 \times 10^{-03}$
Vit. D	$19.9 \pm 9.75$	$17.61 \pm 9.02$	$7.13 \times 10^{-03}$
BMI	$29.7 \pm 5.21$	$28.5 \pm 5.09$	$1.11 \times 10^{-02}$
C-Reactive Protein	$7.68 \pm 6.18$	$6.33 \pm 5.06$	$1.34 \times 10^{-02}$
Low Density Lipoprotein (LDL-C)	$2.92 \pm 0.96$	$3.11 \pm 0.91$	$2.33 \times 10^{-02}$
GGT	$42.12 \pm 38.7$	$33.9 \pm 42.62$	$3.20 \times 10^{-02}$
AST (GOT)	$20.68 \pm 8.42$	$22.41 \pm 13.93$	$6.38 \times 10^{-02}$
Cholesterol	$4.85 \pm 1.09$	$5.01 \pm 0.98$	$8.48 \times 10^{-02}$
C-Peptide of insulin	$3.43 \pm 2.48$	$3.07 \pm 2.04$	$1.00 \times 10^{-01}$
Creatine Kinase	$133.5 \pm 115.99$	$156.66 \pm 229.71$	$1.02 \times 10^{-01}$
Phosphorus	$1.12 \pm 0.18$	$1.13 \pm 0.19$	$3.08 \times 10^{-01}$
Hemoglobin	$14.7 \pm 1.23$	$14.81 \pm 1.07$	$3.33 \times 10^{-01}$
Thyroid Stimulating Hormone (TSH)	$1.63 \pm 1.19$	$1.87 \pm 4.91$	$3.47 \times 10^{-01}$
Vitamin B12	$334.45 \pm 179.1$	$322.09 \pm 161.78$	$4.20 \times 10^{-01}$
Alkaline Phosphatase (ALP)	$71.99 \pm 20.14$	$70.63 \pm 18.74$	$4.41 \times 10^{-01}$
Free Thyroxine	$13.21 \pm 1.38$	$13.31 \pm 1.66$	$4.54 \times 10^{-01}$
Calcium	$2.4 \pm 0.09$	$2.4 \pm 0.09$	$5.69 \times 10^{-01}$
ALT (GPT)	$30.11 \pm 16.56$	$29.73 \pm 22.39$	$8.22 \times 10^{-01}$
Potassium	$4.39 \pm 0.3$	$4.39 \pm 0.32$	$9.49 \times 10^{-01}$

Table 2: Baseline characteristics for females in diabetes study. (Rows are sorted by significance)

	Diabetic ( $n = 152$ )	Non-Diabetic ( $n = 444$ )	P-value
Age	$53.16 \pm 8.8$	$39.53 \pm 12.72$	$1.60 \times 10^{-38}$
Triglycerides	$1.71 \pm 0.81$	$1.19 \pm 0.71$	$1.25 \times 10^{-11}$
Alkaline Phosphatase (ALP)	$81.35 \pm 21.03$	$68.23 \pm 20.68$	$4.16 \times 10^{-11}$
Magnesium	$0.78 \pm 0.09$	$0.83 \pm 0.06$	$1.07 \times 10^{-10}$
Chloride	$99.72 \pm 2.7$	$101.34 \pm 2.04$	$1.13 \times 10^{-10}$
BMI	$33.23 \pm 6.13$	$29.61 \pm 6.66$	$6.25 \times 10^{-09}$
Vit. D	$23.95 \pm 9.32$	$18.71 \pm 9.73$	$1.17 \times 10^{-08}$
High Density Lipoprotein (HDL-C)	$1.33 \pm 0.33$	$1.49 \pm 0.35$	$1.90 \times 10^{-06}$
GGT	$33.32 \pm 32.71$	$20.51 \pm 20.51$	$1.01 \times 10^{-05}$
Albumin	$43.33 \pm 2.82$	$44.46 \pm 2.81$	$2.08 \times 10^{-05}$
Vitamin B12	$369.02 \pm 247.06$	$285.97 \pm 149.7$	$1.30 \times 10^{-04}$
Calcium	$2.4 \pm 0.1$	$2.37 \pm 0.1$	$3.50 \times 10^{-04}$
Free triiodothyronine	$4.2 \pm 0.82$	$4.46 \pm 0.65$	$3.60 \times 10^{-04}$
C-Peptide of insulin	$3.17 \pm 1.78$	$2.6 \pm 1.46$	$4.07 \times 10^{-04}$
Insulin	$27.91 \pm 56.3$	$12.2 \pm 10.28$	$8.00 \times 10^{-04}$
ALT (GPT)	$22.19 \pm 12.94$	$17.63 \pm 17.9$	$8.10 \times 10^{-04}$
C-Reactive Protein	$8.7 \pm 5.83$	$7.16 \pm 5.13$	$3.93 \times 10^{-03}$
Free Thyroxine	$14.14 \pm 2.44$	$13.72 \pm 2.01$	$5.76 \times 10^{-02}$
Uric Acid	$268.82 \pm 75.94$	$256.39 \pm 64.1$	$7.16 \times 10^{-02}$
Total Bilirubin	$5.63 \pm 2.45$	$6.05 \pm 3.15$	$8.77 \times 10^{-02}$
Sodium	$139.48 \pm 2.71$	$139.85 \pm 2.29$	$1.34 \times 10^{-01}$
Phosphorus	$1.19 \pm 0.15$	$1.17 \pm 0.15$	$1.39 \times 10^{-01}$
Homocysteine	$7.86 \pm 2.59$	$8.2 \pm 2.46$	$1.41 \times 10^{-01}$
AST (GOT)	$18.89 \pm 9.42$	$17.53 \pm 10.32$	$1.50 \times 10^{-01}$
Thyroid Stimulating Hormone (TSH)	$1.83 \pm 1.36$	$2.21 \pm 5.23$	$1.57 \times 10^{-01}$
Creatine Kinase	$82.32 \pm 78.87$	$73.44 \pm 41.57$	$1.86 \times 10^{-01}$
Folate Serum	$27.31 \pm 8.07$	$26.37 \pm 7.79$	$2.06 \times 10^{-01}$
Hemoglobin	$12.51 \pm 1.25$	$12.38 \pm 1.25$	$2.77 \times 10^{-01}$
Cholesterol	$5 \pm 1.11$	$4.92 \pm 0.88$	$4.55 \times 10^{-01}$
Potassium	$4.39 \pm 0.31$	$4.37 \pm 0.37$	$5.09 \times 10^{-01}$
Low Density Lipoprotein (LDL-C)	$2.87 \pm 0.96$	$2.91 \pm 0.79$	$6.54 \times 10^{-01}$
Serum Creatinine	$56.97 \pm 14.09$	$56.54 \pm 8.91$	$7.25 \times 10^{-01}$

Table 3: Baseline characteristics for males in obesity study. (Rows are sorted by significance)

	Obese ( $n = 203$ )	Normal ( $n = 122$ )	P-value
C-Peptide of insulin	$3.83 \pm 2.49$	$2.4 \pm 1.64$	$1.50 \times 10^{-09}$
High Density Lipoprotein (HDL-C)	$1.11 \pm 0.27$	$1.29 \pm 0.29$	$8.65 \times 10^{-09}$
Triglycerides	$1.88 \pm 1.46$	$1.26 \pm 0.67$	$3.29 \times 10^{-07}$
Insulin	$26.35 \pm 35.89$	$12.26 \pm 13.11$	$7.58 \times 10^{-07}$
Albumin	$45.43 \pm 2.54$	$46.88 \pm 2.67$	$1.78 \times 10^{-06}$
Uric Acid	$347.47 \pm 77.57$	$311.66 \pm 55.68$	$8.47 \times 10^{-06}$
ALT (GPT)	$32.43 \pm 24.96$	$24.99 \pm 14.4$	$7.38 \times 10^{-04}$
Total Bilirubin	$7.31 \pm 4.78$	$9.31 \pm 5.64$	$1.25 \times 10^{-03}$
Free Thyroxine	$13.06 \pm 1.39$	$13.61 \pm 1.47$	$1.31 \times 10^{-03}$
Age	$43.12 \pm 11.37$	$38.84 \pm 13.68$	$4.04 \times 10^{-03}$
GGT	$41.95 \pm 55.7$	$28.55 \pm 27.75$	$4.21 \times 10^{-03}$
HBA1C %	$6.44 \pm 1.73$	$5.94 \pm 1.43$	$4.83 \times 10^{-03}$
Cholesterol	$5.03 \pm 1.02$	$4.72 \pm 0.93$	$6.71 \times 10^{-03}$
Vit. D	$16.77 \pm 7.27$	$19.17 \pm 10.5$	$2.70 \times 10^{-02}$
AST (GOT)	$23 \pm 14.64$	$20.44 \pm 6.22$	$3.00 \times 10^{-02}$
Calcium	$2.39 \pm 0.09$	$2.42 \pm 0.09$	$3.04 \times 10^{-02}$
Vitamin B12	$306.36 \pm 120.77$	$350.68 \pm 206.4$	$3.22 \times 10^{-02}$
Free triiodothyronine	$4.54 \pm 0.49$	$4.67 \pm 0.65$	$4.28 \times 10^{-02}$
Magnesium	$0.83 \pm 0.07$	$0.85 \pm 0.07$	$4.50 \times 10^{-02}$
Low Density Lipoprotein (LDL-C)	$3.08 \pm 0.89$	$2.88 \pm 0.84$	$4.53 \times 10^{-02}$
Glucose	$6.94 \pm 3.39$	$6.24 \pm 2.94$	$5.91 \times 10^{-02}$
Chloride	$100.56 \pm 2.31$	$100.25 \pm 2.21$	$2.41 \times 10^{-01}$
Potassium	$4.38 \pm 0.29$	$4.34 \pm 0.34$	$2.68 \times 10^{-01}$
Sodium	$140.18 \pm 2.1$	$140.43 \pm 2.45$	$3.16 \times 10^{-01}$
C-Reactive Protein	$7.22 \pm 6.09$	$6.57 \pm 7.1$	$3.77 \times 10^{-01}$
Phosphorus	$1.11 \pm 0.17$	$1.13 \pm 0.19$	$6.25 \times 10^{-01}$
Thyroid Stimulating Hormone (TSH)	$1.63 \pm 1.26$	$1.7 \pm 2.66$	$7.57 \times 10^{-01}$
Serum Creatinine	$77.58 \pm 14.56$	$77.19 \pm 10.62$	$7.79 \times 10^{-01}$
Alkaline Phosphatase (ALP)	$71.64 \pm 20.61$	$72.16 \pm 19.41$	$8.21 \times 10^{-01}$
Homocysteine	$9.76 \pm 3.21$	$9.8 \pm 4.04$	$9.09 \times 10^{-01}$
Creatine Kinase	$150.7 \pm 134.29$	$149.91 \pm 192.61$	$9.67 \times 10^{-01}$
Folate Serum	$23.63 \pm 7.29$	$23.61 \pm 7.77$	$9.83 \times 10^{-01}$
Hemoglobin	$14.75 \pm 1.1$	$14.76 \pm 1.08$	$9.83 \times 10^{-01}$

Table 4: Baseline characteristics for females in obesity study. (Rows are sorted by significance)

	Obese ( $n = 305$ )	Normal ( $n = 102$ )	P-value
Age	$46.85 \pm 11.81$	$30.46 \pm 9.6$	$4.28 \times 10^{-32}$
Albumin	$43.16 \pm 2.52$	$46.24 \pm 2.51$	$1.39 \times 10^{-23}$
C-Peptide of insulin	$3.15 \pm 1.65$	$1.88 \pm 0.96$	$8.64 \times 10^{-19}$
HBA1C %	$6.58 \pm 1.61$	$5.41 \pm 0.96$	$7.66 \times 10^{-17}$
Triglycerides	$1.43 \pm 0.71$	$0.9 \pm 0.51$	$2.08 \times 10^{-14}$
Uric Acid	$275.72 \pm 68.94$	$224.59 \pm 50.39$	$4.40 \times 10^{-14}$
Alkaline Phosphatase (ALP)	$76.99 \pm 21.43$	$60.51 \pm 16.62$	$5.79 \times 10^{-14}$
C-Reactive Protein	$8.92 \pm 6.06$	$5.62 \pm 2.82$	$8.75 \times 10^{-13}$
High Density Lipoprotein (HDL-C)	$1.38 \pm 0.32$	$1.63 \pm 0.35$	$1.86 \times 10^{-10}$
Glucose	$6.88 \pm 3.45$	$5.02 \pm 2.13$	$5.41 \times 10^{-10}$
Magnesium	$0.8 \pm 0.07$	$0.84 \pm 0.06$	$1.30 \times 10^{-06}$
Insulin	$20.38 \pm 40.59$	$8.64 \pm 7.18$	$2.02 \times 10^{-06}$
GGT	$26.47 \pm 25.31$	$15.49 \pm 17.69$	$2.41 \times 10^{-06}$
Chloride	$100.62 \pm 2.32$	$101.6 \pm 1.96$	$5.03 \times 10^{-05}$
Creatine Kinase	$79.54 \pm 54.82$	$63.9 \pm 27.59$	$2.00 \times 10^{-04}$
Total Bilirubin	$5.44 \pm 2.65$	$6.87 \pm 3.53$	$2.39 \times 10^{-04}$
Serum Creatinine	$57.77 \pm 10.82$	$54.71 \pm 7.02$	$1.13 \times 10^{-03}$
Calcium	$2.37 \pm 0.11$	$2.39 \pm 0.08$	$1.57 \times 10^{-02}$
Phosphorus	$1.17 \pm 0.15$	$1.2 \pm 0.13$	$2.16 \times 10^{-02}$
Thyroid Stimulating Hormone (TSH)	$2.15 \pm 2.78$	$1.7 \pm 1.26$	$2.81 \times 10^{-02}$
Vit. D	$20.23 \pm 8.74$	$17.65 \pm 11.53$	$4.00 \times 10^{-02}$
Cholesterol	$4.96 \pm 0.89$	$4.75 \pm 0.82$	$4.75 \times 10^{-02}$
Low Density Lipoprotein (LDL-C)	$2.93 \pm 0.82$	$2.75 \pm 0.74$	$5.50 \times 10^{-02}$
Potassium	$4.37 \pm 0.34$	$4.45 \pm 0.4$	$7.00 \times 10^{-02}$
Hemoglobin	$12.26 \pm 1.33$	$12.5 \pm 1.11$	$7.04 \times 10^{-02}$
Free triiodothyronine	$4.35 \pm 0.66$	$4.49 \pm 0.58$	$7.20 \times 10^{-02}$
Free Thyroxine	$13.79 \pm 2.13$	$13.98 \pm 1.7$	$3.56 \times 10^{-01}$
ALT (GPT)	$19.09 \pm 10.87$	$16.7 \pm 32.8$	$4.71 \times 10^{-01}$
Homocysteine	$8.34 \pm 2.63$	$8.14 \pm 2.53$	$4.87 \times 10^{-01}$
Folate Serum	$26.47 \pm 7.96$	$25.65 \pm 7.2$	$4.92 \times 10^{-01}$
Sodium	$139.71 \pm 2.57$	$139.88 \pm 2.11$	$5.12 \times 10^{-01}$
AST (GOT)	$17.57 \pm 6.77$	$18.11 \pm 19.31$	$7.84 \times 10^{-01}$
Vitamin B12	$302.49 \pm 176.64$	$299.54 \pm 184.57$	$8.84 \times 10^{-01}$

# Regularization Methods

Table 5: Significant results of elastic net, glinternet, lasso proj and hdi. (Rows are sorted by the absolute value of elastic net coefficients)

	elastic net coefficient ( $\beta$ )	glinetnet coefficient ( $\beta$ )	lasso proj P-value	hdi P-value
<b>Diabetes Study (Males)</b>				
Magnesium	$-9.23 \times 10^{-1}$	$-1.50 \times 10^{+1}$	$8.10 \times 10^{-5}$	$1.08 \times 10^{-3}$
High Density Lipoprotein	$-8.36 \times 10^{-2}$	$-8.17 \times 10^{-2}$	$5.78 \times 10^{-1}$	
Free triiodothyronine	$-6.37 \times 10^{-2}$	$-1.19 \times 10^{-1}$	$6.55 \times 10^{-4}$	$1.67 \times 10^{-1}$
Chloride	$-4.70 \times 10^{-2}$	$-1.70 \times 10^{-1}$	$9.99 \times 10^{-13}$	$1.96 \times 10^{-11}$
Albumin	$-2.08 \times 10^{-2}$	$-2.12 \times 10^{-2}$	$2.36 \times 10^{-6}$	$7.54 \times 10^{-4}$
Folate Serum	$4.50 \times 10^{-3}$	$1.07 \times 10^{-2}$	$8.67 \times 10^{-4}$	$1.10 \times 10^{-1}$
Uric Acid	$-1.16 \times 10^{-3}$	$-8.13 \times 10^{-3}$	$5.10 \times 10^{-14}$	$1.17 \times 10^{-7}$
Insulin	$4.94 \times 10^{-4}$	$6.76 \times 10^{-6}$	$3.94 \times 10^{-1}$	
<b>Diabetes Study (Females)</b>				
Magnesium	$-8.66 \times 10^{-1}$	$-2.91 \times 10^{+0}$	$1.44 \times 10^{-4}$	$2.09 \times 10^{-4}$
Calcium	$8.47 \times 10^{-2}$	$5.67 \times 10^{-1}$	$2.72 \times 10^{-1}$	
Triglycerides	$6.83 \times 10^{-2}$	$2.20 \times 10^{-1}$	$4.46 \times 10^{-2}$	$7.47 \times 10^{-4}$
Free triiodothyronine	$-2.29 \times 10^{-2}$	$-1.05 \times 10^{+0}$	$9.26 \times 10^{-4}$	
Chloride	$-1.82 \times 10^{-2}$	$1.93 \times 10^{-2}$	$6.98 \times 10^{-4}$	$2.05 \times 10^{-1}$
High Density Lipoprotein	$-1.36 \times 10^{-2}$	$-4.62 \times 10^{-1}$	$6.17 \times 10^{-1}$	
BMI	$2.92 \times 10^{-3}$	$-1.47 \times 10^{-3}$	$7.12 \times 10^{-1}$	
Vitamin D	$2.89 \times 10^{-3}$	$3.62 \times 10^{-3}$	$1.33 \times 10^{-1}$	$9.41 \times 10^{-1}$
<b>Obesity Study (Males)</b>				
High Density Lipoprotein	$-2.18 \times 10^{-1}$	$4.05 \times 10^{-1}$	$7.94 \times 10^{-4}$	$1.09 \times 10^{-1}$
C-Peptide of insulin	$2.31 \times 10^{-2}$	$3.60 \times 10^{-1}$	$6.92 \times 10^{-2}$	$2.99 \times 10^{-1}$
Albumin	$-1.69 \times 10^{-2}$	$4.52 \times 10^{-3}$	$4.38 \times 10^{-1}$	$6.30 \times 10^{-1}$
Free Thyroxine	$-1.61 \times 10^{-2}$	$-1.40 \times 10^{-1}$	$4.50 \times 10^{-1}$	
HBA1C %	$1.22 \times 10^{-2}$	$9.83 \times 10^{-2}$	$2.06 \times 10^{-1}$	
Cholesterol	$5.70 \times 10^{-3}$	$-7.58 \times 10^{-2}$	$3.71 \times 10^{-1}$	
Triglycerides	$2.62 \times 10^{-3}$	$4.12 \times 10^{-3}$		
Vitamin D	$-1.22 \times 10^{-3}$	$-5.27 \times 10^{-2}$	$5.18 \times 10^{-1}$	
Uric Acid	$8.29 \times 10^{-4}$	$1.21 \times 10^{-3}$	$3.23 \times 10^{-5}$	$7.39 \times 10^{-2}$
ALT (GPT)	$7.78 \times 10^{-4}$	$-2.11 \times 10^{-2}$		
Total Bilirubin	$-2.65 \times 10^{-4}$	$3.93 \times 10^{-3}$		
<b>Obesity Study (Females)</b>				
High Density Lipoprotein	$-5.24 \times 10^{-2}$	$1.90 \times 10^{-2}$	$1.84 \times 10^{-1}$	
Albumin	$-4.12 \times 10^{-2}$	$-1.44 \times 10^{-1}$	$1.90 \times 10^{-9}$	$3.87 \times 10^{-10}$
Triglycerides	$3.78 \times 10^{-2}$	$-3.07 \times 10^{-1}$		$8.12 \times 10^{-1}$
C-Peptide of insulin	$2.03 \times 10^{-2}$	$-2.32 \times 10^{-1}$	$2.57 \times 10^{-1}$	
HBA1C %	$1.93 \times 10^{-2}$	$-3.57 \times 10^{-1}$	$1.55 \times 10^{-2}$	
Potassium	$-1.46 \times 10^{-2}$	$-2.41 \times 10^{-2}$		
Alkaline Phosphatase	$1.84 \times 10^{-3}$	$4.31 \times 10^{-3}$	$5.65 \times 10^{-2}$	$3.86 \times 10^{-1}$
Chloride	$-1.49 \times 10^{-3}$	$-6.24 \times 10^{-3}$	$1.03 \times 10^{-1}$	
Phosphorus	$-1.27 \times 10^{-3}$	$-1.02 \times 10^{-1}$	$3.62 \times 10^{-1}$	
Uric Acid	$8.94 \times 10^{-4}$	$2.14 \times 10^{-3}$	$3.39 \times 10^{-5}$	$6.79 \times 10^{-4}$

# Machine Learning Methods

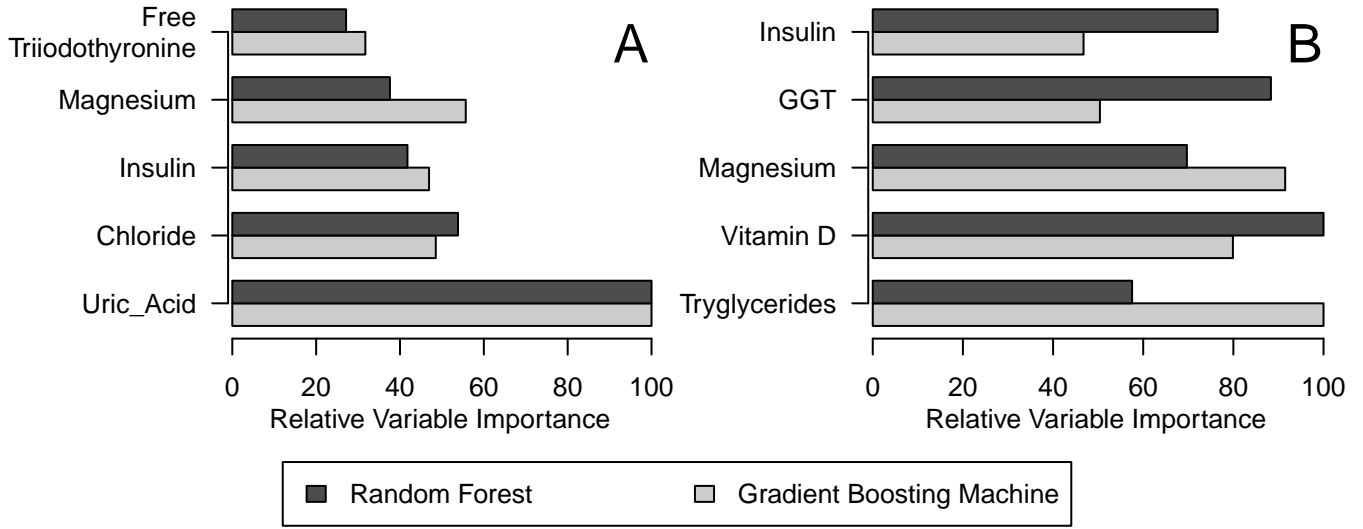


Figure 1: Relative variable importance of top variables of machine learning methods for (A) Males and (B) Females in Diabetes studies.

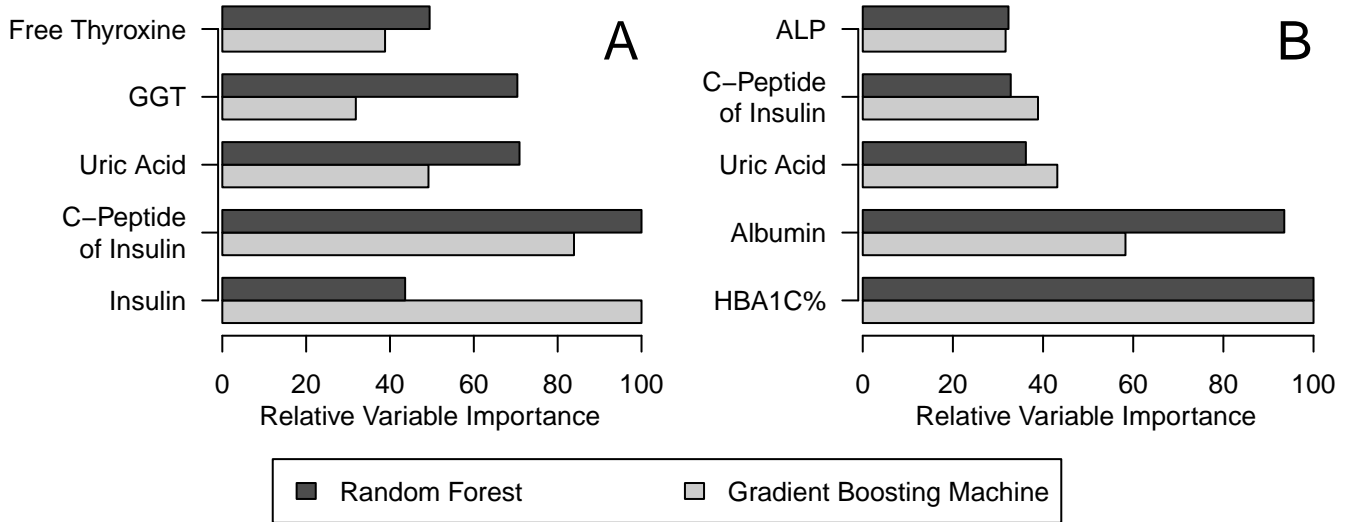


Figure 2: Relative variable importance of top variables of machine learning methods for (A) Males and (B) Females in Obesity studies.