Paper Part 1: Table 2

Part A: Unadjusted analysis

**> summary(glm(bmi~log(pot.den.ffq)+sex+age,subset=!is.na(sod.ffq)))**

**Call: glm(formula = bmi ~ log(pot.den.ffq) + sex + age, subset = !is.na(sod.ffq))**

**Deviance Residuals:**

**Min 1Q Median 3Q Max**

**-10.23257761 -3.611599501 -0.9421481251 2.613448039 23.6675035**

**Coefficients:**

**Value Std. Error t value**

**(Intercept) 27.27970304926 1.76370401199 15.4672795797**

**log(pot.den.ffq) -1.68861963058 0.93369142715 -1.8085414319**

**sex -0.37743182645 0.49133847595 -0.7681707111**

**age 0.03884958636 0.02940883789 1.3210173927**

**(Dispersion Parameter for Gaussian family taken to be 27.69892335 )**

**Null Deviance: 13429.12809 on 482 degrees of freedom**

**Residual Deviance: 13267.78429 on 479 degrees of freedom**

**Number of Fisher Scoring Iterations: 1**

**Correlation of Coefficients:**

**(Intercept) log(pot.den.ffq) sex**

**log(pot.den.ffq) -0.0628721082**

**sex -0.4397412696 -0.1967040840**

**age -0.8814587091 -0.1696406259 0.1021679250**

Part B: Prediction model used by RC

**summary(glm(logpot.den.urav[1:250]~log(pot.den.ffq[1:250])+sex[1:250]+age[1:250],subset=!is.na(logpot.den.urav[1:250])))**

**Call: glm(formula = logpot.den.urav[1:250] ~ log(pot.den.ffq[1:250]) + sex[1:250] + age[1:250], subset = !is.na**

**(**

**logpot.den.urav[1:250]))**

**Deviance Residuals:**

**Min 1Q Median 3Q Max**

**-2.189545227 -0.1627275503 0.03625760434 0.1838953913 0.7046653572**

**Coefficients:**

**Value Std. Error t value**

**(Intercept) -0.47889683240 0.161926014886 -2.957503973**

**log(pot.den.ffq[1:250]) 0.44920301576 0.089424195265 5.023282730**

**sex[1:250] 0.04590277636 0.045081815374 1.018210469**

**age[1:250] 0.00591434811 0.002651430005 2.230625775**

**(Dispersion Parameter for Gaussian family taken to be 0.1052102843 )**

**Null Deviance: 25.87866777 on 211 degrees of freedom**

**Residual Deviance: 21.88373913 on 208 degrees of freedom**

**Number of Fisher Scoring Iterations: 1**

**Correlation of Coefficients:**

**(Intercept) log(pot.den.ffq[1:250]) sex[1:250]**

**log(pot.den.ffq[1:250]) -0.0878465874**

**sex[1:250] -0.4689452119 -0.1290678365**

**age[1:250] -0.8601394553 -0.1942118020 0.0998851244**

Part C: Adjusted analysis by RC – parameter estimates

**> open$predlogpotden2<- -0.4789 + 0.4492\*log(pot.den.ffq) + 0.04590\*sex + 0.005914\*age**

**> attach(open)**

**> summary(glm(bmi~predlogpotden2+sex+age,subset=!is.na(predlogpotden2)))**

**Call: glm(formula = bmi ~ predlogpotden2 + sex + age, subset = !is.na(predlogpotden2))**

**Deviance Residuals:**

**Min 1Q Median 3Q Max**

**-10.23257761 -3.611599501 -0.9421481251 2.613448039 23.6675035**

**Coefficients:**

**Value Std. Error t value**

**(Intercept) 25.47943603884 1.96996484231 12.9339547040**

**predlogpotden2 -3.75917103868 2.07856506489 -1.8085414319**

**sex -0.20488587577 0.51861113345 -0.3950664815**

**age 0.06108132388 0.03374378255 1.8101504709**

**(Dispersion Parameter for Gaussian family taken to be 27.69892335 )**

**Null Deviance: 13429.12809 on 482 degrees of freedom**

**Residual Deviance: 13267.78429 on 479 degrees of freedom**

**Number of Fisher Scoring Iterations: 1**

**Correlation of Coefficients:**

**(Intercept) predlogpotden2 sex**

**predlogpotden2 0.4490115768**

**sex -0.5497653766 -0.3703245244**

**age -0.9260653147 -0.5121407902 0.2464656623**

**>**

Part C: Adjusted analysis by RC – bootstrap SEs

> strat.boot2mod

function(nboot, y, xobs, z1, z2, xref)

{

b <- rep(0, nboot)

bz1 <- rep(0, nboot)

bz2 <- rep(0, nboot)

xref1 <- xref[!is.na(xref)]

xobs1 <- xobs[!is.na(xref)]

z11 <- z1[!is.na(xref)]

z21 <- z2[!is.na(xref)]

y1 <- y[!is.na(xref)]

y2 <- y[is.na(xref)]

xobs2 <- xobs[is.na(xref)]

z12 <- z1[is.na(xref)]

z22 <- z2[is.na(xref)]

for(i in 1:nboot) {

sampk <- sample(seq(length(xref1)), replace = T)

runpred <- glm(xref1[sampk] ~ xobs1[sampk] + z11[sampk] + z21[sampk])

sampj <- sample(seq(length(y2)), replace = T)

ynew <- c(y1[sampk], y2[sampj])

xobsnew <- c(xobs1[sampk], xobs2[sampj])

z1new <- c(z11[sampk], z12[sampj])

z2new <- c(z21[sampk], z22[sampj])

xpred <- runpred$coef[1] + runpred$coef[2] \* xobsnew + runpred$coef[3] \* z1new + runpred$coef[4] \* z2new

run <- glm(ynew ~ xpred + z1new + z2new)

b[i] <- run$coef[2]

bz1[i] <- run$coef[3]

bz2[i] <- run$coef[4]

}

c(mean(b), stdev(b), median(b), quantile(b), mean(bz1), stdev(bz1), mean(bz2), stdev(bz2))

}

> strat.boot2mod(5000,bmi[!is.na(pot.den.ffq)],log(pot.den.ffq[!is.na(pot.den.ffq)]),age[!is.na(pot.den.ffq)],sex[!is.na(pot.den.ffq)],logpot.den.urav.mod[!is.na(pot.den.ffq)])

0% 25% 50% 75% 100%

-3.847527074 **2.428333046** -3.701628903 -27.56516548 -5.285744176 -3.701628903 -2.254706107 6.034609342 0.06202949411

**0.03369161495** -0.190893721 **0.5761791425**

Table 3: Part A: Unadjusted analysis

**> summary(glm(bmi~log(pot.den.ffq)+log(sod.den.ffq)+sex+age,subset=!is.na(pot.den.ffq)))**

**Call: glm(formula = bmi ~ log(pot.den.ffq) + log(sod.den.ffq) + sex + age, subset = !is.na(pot.den.ffq))**

**Deviance Residuals:**

**Min 1Q Median 3Q Max**

**-10.27140643 -3.564291142 -0.9620607239 2.626061781 23.31995155**

**Coefficients:**

**Value Std. Error t value**

**(Intercept) 26.74070172003 1.81387189307 14.7423320369**

**log(pot.den.ffq) -1.93487001284 0.95339917880 -2.0294437586**

**log(sod.den.ffq) 1.50849436585 1.19813392221 1.2590365216**

**sex -0.32813747133 0.49259704368 -0.6661377195**

**age 0.03816717705 0.02939588788 1.2983849035**

**(Dispersion Parameter for Gaussian family taken to be 27.6651261 )**

**Null Deviance: 13429.12809 on 482 degrees of freedom**

**Residual Deviance: 13223.93028 on 478 degrees of freedom**

**Number of Fisher Scoring Iterations: 1**

**Correlation of Coefficients:**

**(Intercept) log(pot.den.ffq) log(sod.den.ffq) sex**

**log(pot.den.ffq) -0.0113782529**

**log(sod.den.ffq) -0.2360178755 -0.2051463582**

**sex -0.4447252166 -0.2082167515 0.0794816840**

**age -0.8520589928 -0.1622218234 -0.0184382641 0.1003618813**

**>**

Table 3: Part B1: Prediction model for log potassium density

**> summary(glm(logpot.den.urav[1:250]~log(pot.den.ffq[1:250])+log(sod.den.ffq[1:250])+sex[1:250]+age[1:250],subset=!is.na(logpot.den.urav[1:250])&!is.na(pot.den.ffq[1:250])&!is.na(sod.den.ffq[1:250])))**

**Call: glm(formula = logpot.den.urav[1:250] ~ log(pot.den.ffq[1:250]) + log(sod.den.ffq[1:250]) + sex[1:250] + a**

**ge[1:**

**250], subset = !is.na(logpot.den.urav[1:250]) & !is.na(pot.den.ffq[1:250]) & !is.na(sod.den.ffq[1:250]**

**))**

**Deviance Residuals:**

**Min 1Q Median 3Q Max**

**-2.191501292 -0.1627006302 0.03722671942 0.1849110677 0.7050642608**

**Coefficients:**

**Value Std. Error t value**

**(Intercept) -0.476396757504 0.167836661289 -2.83845468472**

**log(pot.den.ffq[1:250]) 0.450134847608 0.091040770791 4.94432157919**

**log(sod.den.ffq[1:250]) -0.006288077257 0.107385070953 -0.05855634495**

**sex[1:250] 0.045601527840 0.045482098593 1.00262585171**

**age[1:250] 0.005915839443 0.002657926718 2.22573459316**

**(Dispersion Parameter for Gaussian family taken to be 0.1057167954 )**

**Null Deviance: 25.87866777 on 211 degrees of freedom**

**Residual Deviance: 21.88337665 on 207 degrees of freedom**

**Number of Fisher Scoring Iterations: 1**

**Correlation of Coefficients:**

**(Intercept) log(pot.den.ffq[1:250]) log(sod.den.ffq[1:250]) sex[1:250]**

**log(pot.den.ffq[1:250]) -0.0391836151**

**log(sod.den.ffq[1:250]) -0.2543854262 -0.1747944592**

**sex[1:250] -0.4793818517 -0.1460366815 0.1131124640**

**age[1:250] -0.8293676591 -0.1895382355 -0.0095820331 0.0981556773**

**>**

Table 3: Part B2: Prediction model for log sodium density

**> summary(glm(logsod.den.urav[1:250]~log(pot.den.ffq[1:250])+log(sod.den.ffq[1:250])+sex[1:250]+age[1:250],subset=!is.na(logpot.den.urav[1:250])&!is.na(pot.den.ffq[1:250])&!is.na(sod.den.ffq[1:250])))**

**Call: glm(formula = logsod.den.urav[1:250] ~ log(pot.den.ffq[1:250]) + log(sod.den.ffq[1:250]) + sex[1:250] + a**

**ge[1:**

**250], subset = !is.na(logpot.den.urav[1:250]) & !is.na(pot.den.ffq[1:250]) & !is.na(sod.den.ffq[1:250]**

**))**

**Deviance Residuals:**

**Min 1Q Median 3Q Max**

**-1.296534341 -0.1706700093 0.01347986835 0.1965336188 1.002107819**

**Coefficients:**

**Value Std. Error t value**

**(Intercept) 0.201714948168 0.159070566966 1.268084675**

**log(pot.den.ffq[1:250]) -0.143329836727 0.086285719196 -1.661107285**

**log(sod.den.ffq[1:250]) 0.418660664014 0.101776357972 4.113535524**

**sex[1:250] -0.025408227850 0.043106572512 -0.589428163**

**age[1:250] 0.003749758628 0.002519103435 1.488529044**

**(Dispersion Parameter for Gaussian family taken to be 0.0949620288 )**

**Null Deviance: 21.69022875 on 211 degrees of freedom**

**Residual Deviance: 19.65713996 on 207 degrees of freedom**

**Number of Fisher Scoring Iterations: 1**

**Correlation of Coefficients:**

**(Intercept) log(pot.den.ffq[1:250]) log(sod.den.ffq[1:250]) sex[1:250]**

**log(pot.den.ffq[1:250]) -0.0391836151**

**log(sod.den.ffq[1:250]) -0.2543854262 -0.1747944592**

**sex[1:250] -0.4793818517 -0.1460366815 0.1131124640**

**age[1:250] -0.8293676591 -0.1895382355 -0.0095820331 0.0981556773**

**>**

Table 3: Part C: Adjusted analysis using regression calibration

**> open$predlogpotden3<- -0.4764+0.4501\*log(pot.den.ffq)-0.0063\*log(sod.den.ffq)+0.04560\*sex+0.005916\*age**

**> open$predlogsodden3<- 0.2017-0.1433\*log(pot.den.ffq)+0.4187\*log(sod.den.ffq)-0.02541\*sex+0.003750\*age**

**> attach(open)**

**> summary(glm(bmi~predlogpotden3+predlogsodden3+sex+age,subset=!is.na(predlogpotden3)&!is.na(predlogsodden3)))**

**Call: glm(formula = bmi ~ predlogpotden3 + predlogsodden3 + sex + age, subset = !is.na(predlogpotden3) & !is.na**

**(**

**predlogsodden3))**

**Deviance Residuals:**

**Min 1Q Median 3Q Max**

**-10.27140643 -3.564291142 -0.9620607239 2.626061781 23.31995155**

**Coefficients:**

**Value Std. Error t value**

**(Intercept) 24.51492155135 2.11350988429 11.5991515978**

**predlogpotden3 -3.16688835429 2.13743887932 -1.4816275613**

**predlogsodden3 3.55515397473 2.86893251382 1.2391905204**

**sex -0.09339089987 0.52566750506 -0.1776615427**

**age 0.04357066115 0.03663874466 1.1891963427**

**(Dispersion Parameter for Gaussian family taken to be 27.6651261 )**

**Null Deviance: 13429.12809 on 482 degrees of freedom**

**Residual Deviance: 13223.93028 on 478 degrees of freedom**

**Number of Fisher Scoring Iterations: 1**

**Correlation of Coefficients:**

**(Intercept) predlogpotden3 predlogsodden3 sex**

**predlogpotden3 0.3210754660**

**predlogsodden3 -0.3638624749 0.2346398795**

**sex -0.5657378645 -0.3152686032 0.1676434656**

**age -0.6517084248 -0.5499659185 -0.3908973420 0.1579857432**

**>**

> strat.boot3mod

function(nboot, y, x1obs, x2obs, z1, z2, x1ref, x2ref)

{

b1 <- rep(0, nboot)

b2 <- rep(0, nboot)

bz1 <- rep(0, nboot)

bz2 <- rep(0, nboot)

x1ref1 <- x1ref[!is.na(x1ref) & !is.na(x2ref)]

x2ref1 <- x2ref[!is.na(x1ref) & !is.na(x2ref)]

x1obs1 <- x1obs[!is.na(x1ref) & !is.na(x2ref)]

x2obs1 <- x2obs[!is.na(x1ref) & !is.na(x2ref)]

z11 <- z1[!is.na(x1ref) & !is.na(x2ref)]

z21 <- z2[!is.na(x1ref) & !is.na(x2ref)]

y1 <- y[!is.na(x1ref) & !is.na(x2ref)]

y2 <- y[is.na(x1ref) | is.na(x2ref)]

x1obs2 <- x1obs[is.na(x1ref) | is.na(x2ref)]

x2obs2 <- x2obs[is.na(x1ref) | is.na(x2ref)]

z12 <- z1[is.na(x1ref) | is.na(x2ref)]

z22 <- z2[is.na(x1ref) | is.na(x2ref)]

for(i in 1:nboot) {

sampk <- sample(seq(length(x1ref1)), replace = T)

runpred1 <- glm(x1ref1[sampk] ~ x1obs1[sampk] + x2obs1[sampk] + z11[sampk] + z21[sampk])

xpred1 <- runpred1$coef[1] + runpred1$coef[2] \* x1obs + runpred1$coef[3] \* x2obs + runpred1$coef[4] \* z1 +

runpred1$coef[5] \* z2

runpred2 <- glm(x2ref1[sampk] ~ x1obs1[sampk] + x2obs1[sampk] + z11[sampk] + z21[sampk])

xpred2 <- runpred2$coef[1] + runpred2$coef[2] \* x1obs + runpred2$coef[3] \* x2obs + runpred2$coef[4] \* z1 +

runpred2$coef[5] \* z2

sampj <- sample(seq(length(y2)), replace = T)

ynew <- c(y1[sampk], y2[sampj])

x1obsnew <- c(x1obs1[sampk], x1obs2[sampj])

x2obsnew <- c(x2obs1[sampk], x2obs2[sampj])

z1new <- c(z11[sampk], z12[sampj])

z2new <- c(z21[sampk], z22[sampj])

xpred1 <- runpred1$coef[1] + runpred1$coef[2] \* x1obsnew + runpred1$coef[3] \* x2obsnew + runpred1$coef[4]

\*

z1new + runpred1$coef[5] \* z2new

xpred2 <- runpred2$coef[1] + runpred2$coef[2] \* x1obsnew + runpred2$coef[3] \* x2obsnew + runpred2$coef[4]

\*

z1new + runpred2$coef[5] \* z2new

run <- glm(ynew ~ xpred1 + xpred2 + z1new + z2new)

b1[i] <- run$coef[2]

b2[i] <- run$coef[3]

bz1[i] <- run$coef[4]

bz2[i] <- run$coef[5]

}

c(mean(b1), stdev(b1), quantile(b1, probs = seq(0, 1, 0.1)), mean(b2), stdev(b2), quantile(b2, probs = seq(0, 1,

0.1)), mean(bz1), stdev(bz1), quantile(bz1, probs = seq(0, 1, 0.1)), mean(bz2), stdev(bz2), quantile(bz2,

probs = seq(0, 1, 0.1)))

}

>

strat.boot3mod(5000,bmi[!is.na(pot.den.ffq)&!is.na(sod.den.ffq)],log(pot.den.ffq[!is.na(pot.den.ffq)&!is.na(sod.den.ffq)]),log(sod.den.ffq[!is.na(pot.den.ffq)&!is.na(sod.den.ffq)]),age[!is.na(pot.den.ffq)&!is.na(sod.den.ffq)],sex[!is.na(pot.den.ffq)&!is.na(sod.den.ffq)],logpot.den.urav.mod[!is.na(pot.den.ffq)&!is.na(sod.den.ffq)],logsod.den.urav.mod[!is.na(pot.den.ffq)&!is.na(sod.den.ffq)])

0% 10% 20% 30% 40% 50% 60%

-3.096007905 **2.343909357** -13.50357359 -5.957270295 -4.966114044 -4.277516504 -3.672524077 -3.089150982 -2.538405443

70% 80% 90% 100% 0% 10% 20%

-1.91351659 -1.244790409 -0.2314993294 24.83159995 3.684305052 **3.820581422** -14.31513798 -0.2243576686 1.150939219

30% 40% 50% 60% 70% 80% 90% 100%

2.015877774 2.778576446 3.511777648 4.234173642 5.077118866 6.085243544 7.680309306 86.3221475 0.04254842176

0% 10% 20% 30% 40% 50% 60%

**0.03815590667** -0.1923375415 -0.004336184183 0.01244864936 0.0242481552 0.03406533545 0.04363979445 0.05289355563

70% 80% 90% 100% 0% 10% 20%

0.06221359363 0.07330957105 0.09008448526 0.23969918 -0.1059413867 **0.5988896467** -2.261993676 -0.848702353 -0.5901675493

30% 40% 50% 60% 70% 80% 90% 100%

-0.4056957298 -0.2525383692 -0.1111230762 0.03231706856 0.1767637103 0.3598294348 0.6195122983 6.68339857