HUMAN POPULATION STRUCTURE

Bruce S. Weir http://statgen.ncsu.edu

How is kinship measured?

We each have two copies, or alleles, of a gene – one from each of our parents. We pass on one of these two to each of our children.

The **coefficient of kinship**, θ , is the chance that an allele chosen from one person is identical to an allele chosen from the other. It is also the chance that a child of these two people receives two identical copies of the same gene, i.e. the **inbreeding coefficient**, F, of the child.

202

203

Consanguineous marriages

Marriages between relatives increases the chance that a child receives two identical copies of a gene. If these copies are deleterious, neither is protected by the other, and the child may have a medical handicap. From a collection of 38 studies of first-cousin marriages, a 4.4% depression in survival from six months gestation to age 10 years was found.

(Bittles and Neel, Nature Genetics 8:117, 1997.)

Kinship coefficients

For people who are not inbred, meaning that their parents were not related:

Relatives	θ
Identical twins	1/2
Father-daughter	1/4
Brother-sister	1/4
Grandmother-grandson	1/8
Brother-half sister	1/8
Aunt-nephew	1/8
Double first cousins	1/8
First cousins	1/16
First cousins once removed	1/32
Second cousins	1/64
Unrelated	0

205

Evolutionary relationship

A person cannot have all ancestors different for the last 30 generations (1200 AD) as 2^{30} is 1 billion and there were only half a billion people in 1200.

The fact that the number of distinct ancestors we have is limited by historical population sizes means that any two people have some degree of relationship.

If a population is kept at the same size N, then relative to some reference point t generations ago, $\theta \approx t/2N$.

Historically the human population has had an average effective size of about 100,000. There have been about 5,000-10,000 generations since modern humans are thought to have moved out of Africa. Might expect θ between 0.025 and 0.05.

Estimates of θ

Cavalli-Sforza et al. published "The History and Geography of Human Genes" in 1994. Their average estimates of θ were:

Region	$\widehat{m{ heta}}$
Sub-Saharan Africa	0.035
East Asia	0.025
Southeast Asia	0.035
Asia (India)	0.028
Europe	0.016
South America	0.059
New Guinea	0.039

207

206

Kinship coefficients

For people in a population where people in different families have an evolutionary relationship of $\theta = 0.025$:

Relatives	"θ"
Identical twins	$(1+\theta)/2 = 0.51$
Father-daughter	$(1+3\theta)/4=0.27$
Brother-sister	$(1+3\theta)/4=0.27$
Grandmother-grandson	$(1+7\theta)/8=0.147$
Brother-half sister	$(1+7\theta)/8=0.147$
Aunt-Nephew	$(1+7\theta)/8=0.147$
Double first cousins	$(1+7\theta)/8=0.147$
First cousins	$(1+15\theta)/16=0.086$
First cousins once removed	$(1+31\theta)/32=0.055$
Second cousins	$(1+63\theta)/64=0.032$
" Unrelated"	$\theta = 0.025$

Compare "fixed" populations with means.

Compare "random" populations with variances.

Assuming Hardy-Weinberg, compare two sets of allele counts by exact test with permutation procedure.

208

Two samples

Observed	Sample 1	Sample 2	Total
$\overline{A_1}$	6	6	12
A_2	5	1	6
A_3	1	1	2
Total	12	8	20

If both samples from the same population:

Expected	Sample 1	Sample 2	Total	
A_1	$12 \times \frac{12}{20} = 7.2$	$8 \times \frac{12}{20} = 4.8$	12	
A_2	$12 \times \frac{6}{20} = 3.6$	$8 \times \frac{6}{20} = 2.4$	6	
A_3	$12 \times \frac{2}{20} = 1.2$	$8 \times \frac{2}{20} = 0.8$	2	
Total	12	8	20	
$X^2 = \sum_{E} \frac{(O - E)^2}{E} = 1.94$				

210

Exact test for comparing populations

$$\begin{array}{ll} \Pr(\mathsf{sample}\ I) \ = \ \frac{n_I!}{\prod_i (n_{I_i})!} \prod_i (p_{I_i})^{n_{I_i}} \\ &= \ \frac{12!}{6!5!1!} p_1^6 p_2^5 p_3^1 \\ \\ \Pr(\mathsf{sample}\ II) \ = \ \frac{n_{II}!}{\prod_i (n_{II_i})!} \prod_i (p_{II_i})^{n_{II_i}} \\ &= \ \frac{8!}{6!1!1!} p_1^6 p_2^1 p_3^1 \end{array}$$

Under hypothesis that both samples from same population, p's are equal and joint probability is

$$\left(\frac{12!}{6!5!1!}\right)\left(\frac{8!}{6!1!1!}\right)p_1^{12}p_2^6p_3^2$$

211

Exact test for comparing populations

If the two samples were from the same population, or from two populations with the same allele frequencies, the probability of the combined sample is

$$Pr(combined) = \frac{20!}{12!6!2!} p_1^{12} p_2^6 p_3^2$$

The probability of the two samples, conditional on the counts in the combined data set is, therefore,

probability =
$$\frac{\frac{12!}{6!5!1!} \times \frac{8!}{6!1!1!}}{\frac{20!}{12!6!2!}}$$

Is this small? Use permutation to find the proportion of all pairs of samples with the same total numbers of alleles that are as probable or less probable than the observed samples.

Common Definition of F_{ST}

$$F_{ST} = \frac{s^2}{\bar{p}(1-\bar{p})}$$

$$\propto X^2$$

Never a good idea to define a quantity of interest by a statistic. It cannot then be a parameter, and its values depend on (for example) the dimensions of the data.

213

ANOVA layout for fixed populations

Assuming Hardy-Weinberg, so that allelic rather than genotypic data analyzed. Indicator variable x_{ij} is for the jth allele from the ith population. It is 1 when the allele is type A, and is 0 otherwise. Each sample frequency \tilde{p}_{Ai} unbiased for population frequency p_{Ai} .

$$\sum_{i} \frac{x_i^2}{n_i} - \frac{x_{\cdot \cdot}^2}{n_{\cdot \cdot}} = \sum_{i} n_i (\tilde{p}_{Ai} - \tilde{p}_A)^2$$

Expected MS
$$\frac{1}{r-1} [\sum_i k_1 p_{Ai} (1-p_{Ai}) + \sum_i n_i (p_{Ai} - \bar{p}_A)^2]$$

SS within populations
$$\sum_{i=1}^{r} \sum_{j=1}^{n_i} x_{ij}^2 - \sum_{i} \frac{x_{i}^2}{n_i} = \sum_{i} n_i \tilde{p}_{Ai} (1 - \tilde{p}_{Ai})$$

Expected MS
$$\frac{1}{\sum_{i}(n_{i}-1)}\sum_{i}k_{2}p_{Ai}(1-p_{Ai})$$

214

Within individuals within populations

For a particular population i, the chance that an individual has two copies of allele A is

$$P_{AAi} = p_{Ai}^2 + f_i p_{Ai} (1 - p_{Ai})$$

The quantity f_i is the inbreeding coefficient within that population. Hardy-Weinberg implies that $f_i = 0$. The allele frequency p_{A_i} is for the particular population.

215

Within individuals within populations

If Hardy-Weinberg is not assumed, need to keep track of individuals. The indicator variable is now x_{ijk} for allele k = 1, 2 in individual $j(1,2,\ldots,n_i)$ in population $i(1,2,\ldots,r)$.

SS between alleles within individuals

$$\sum_{i,j,k} x_{ijk}^2 - \sum_{i,j} \frac{x_{ij}^2}{2}$$
$$= \sum_i n_i (\tilde{p}_{Ai} - \tilde{P}_{AAi})$$

Within individuals within populations

Taking expectations over samples from this specific set of populations:

Expected MS between populations: $\frac{1}{r-1} \left[\sum_{i} k_1 p_{Ai} (1-i) \right]$ p_{Ai}) + $\sum_{i} n_{i} (p_{Ai} - \bar{p}_{Ai})^{2}$

Expected MS between individuals within populations: $\frac{1}{n_i-r}\sum_i(n_i-1)(1+f_i)p_{Ai}(1-p_{Ai})$

Expected MS between alleles within individuals: $\frac{1}{n}\sum_{i}n_{i}(1-f_{i})p_{Ai}(1-p_{Ai})$

It might be assumed that the f_i were all the same. Still a fixed population approach.

Within individuals over all populations

Averaging over populations, the chance that an individual has two copies of allele A is

$$P_{AA} = p_A^2 + Fp_A(1-p_A)$$

The quantity F is the inbreeding coefficient for the whole collection of populations. Hardy-Weinberg in the entire population implies that F=0. The allele frequency p_A is an average over all the populations. Now a random population approach.

Between individuals within populations

Averaging over populations, the chance that two individuals in the same population both have a copy of allele A is

$$P_{A,A} = p_A^2 + \theta p_A (1 - p_A)$$

The quantity θ is the coancestry coefficient for the whole collection of populations (measures relatedness of genes within populations relative to that between populations). A lack of population structure implies that $\theta = 0$. The allele frequency p_A is an average over all the populations.

218

219

Within vs. between Populations

There are differences among populations as a consequence of the relatedness of individuals within populations. The expression

$$f = \frac{F - \theta}{1 - \theta}$$

shows that f = 0 is possible even if F and θ are not zero. Populations are different even if each population is in Hardy-Weinberg equilibrium.

ANOVA layout for random populations

For an analysis of allele frequencies, when each population has same expected allele frequency:

SS between populations $\sum_{i=1}^{r} \frac{x_{i\cdot}^2}{n_i} - \frac{x_{\cdot\cdot}^2}{n_{\cdot}}$

$$\sum_{i=1}^{r} \frac{x_{i}^{2}}{n_{i}} - \frac{x_{i}^{2}}{n_{i}}$$

Expected MS between populations

$$p_A(1-p_A)[(1-\theta)+n_c\theta]$$

SS within populations

$$\sum_{i=1}^{r} \sum_{j=1}^{n_i} x_{ij}^2 - \sum_{i} \frac{x_{i}^2}{n_i}$$

Expected MS within populations

$$p_A(1-p_A)(1-\theta)$$

220

Estimation of θ

Therefore

$$\mathcal{E}[(MSB - MSW)/n] = p_A(1 - p_A)\theta$$

$$\mathcal{E}[MSW + (MSB - MSW)/n] = p_A(1 - p_A)$$

so that

$$\hat{\theta} = \frac{\mathsf{MSB} - \mathsf{MSW}}{\mathsf{MSB} + (n-1)\mathsf{MSW}}$$

Unequal Sample Sizes

If the samples are of unequal size n_i :

MSB =
$$\frac{1}{r-1} \sum_{i=1}^{r} n_i (\tilde{p}_{Ai} - \tilde{p}_A)^2$$

MSW = $\frac{1}{\sum_{i} (n_i - 1)} \sum_{i=1}^{r} n_i \tilde{p}_{Ai} (1 - \tilde{p}_{Ai})$

Then

$$\hat{\theta} = \frac{MSB - MSW}{MSB + (n_c - 1)MSW}$$

where

$$n_c = \frac{1}{r-1} \left(\sum_i n_i - \frac{\sum_i n_i^2}{\sum_i n_i} \right)$$

222

223

Estimation of θ

Each allele gives an estimate of the form

$$\hat{\theta} = \frac{\text{Numerator}}{\text{Denominator}}$$

An overall estimate is obtained as the ratio of the sums of these numerators and denominators.

$$\hat{\theta} = \frac{\sum_{\text{alleles}} \text{Numerators}}{\sum_{\text{alleles}} \text{Denominators}}$$

ANOVA structure

$$\begin{aligned} \mathsf{MSB} \; &= \; \frac{n}{r-1} \sum_{i=1}^r (\tilde{p}_{Ai} - \bar{p}_A)^2 \\ \mathsf{MSW} \; &= \; \frac{1}{r(n-1)} \sum_{i=1}^r \tilde{p}_{Ai} (1 - \tilde{p}_{Ai}) \\ \mathcal{E}(\mathsf{MSB}) \; &= \; p_A (1 - p_A) [(1 - \theta) + n \theta] \\ \mathcal{E}(\mathsf{MSW}) \; &= \; p_A (1 - p_A) (1 - \theta) \end{aligned}$$

The two mean squares will differ only when $\theta \neq 0$.

225

ANOVA test for $\theta = 0$

Li (Ph.D. thesis, 1996) showed that, if $\theta = 0$

$$\frac{\mathsf{MSB}}{p_A(1-p_A)} \overset{d}{\to} \frac{\chi^2_{(r-1)}}{r-1}, \text{ as } n \to \infty$$

$$\frac{\mathsf{MSW}}{p_A(1-p_A)} \overset{p}{\to} 1, \text{ as } n \to \infty$$

$$\frac{\mathsf{MSB}}{\mathsf{MSW}} \overset{d}{\to} \frac{\chi^2_{(r-1)}}{r-1}$$

For non-normal data, Boos and Brownie (Statistics and Probability Letters 23:183–191, 1995) showed that

$$F^* = \frac{\mathsf{MSB}}{\mathsf{MSW}} \overset{d}{\to} F_{r(n-1)}^{(r-1)}$$

This could provide a test for $H_0: \theta = 0$.

226

Moment estimator of θ

$$\hat{\theta} = \frac{MSB - MSW}{MSB + (n-1)MSW}$$

When $\theta \neq 0$, Li (Ph.D. thesis, 1996) showed that

$$\frac{\mathsf{MSB}}{p_A(1-p_A)[1+(n-1)\theta]} \overset{d}{\to} \frac{\chi^2_{(r-1)}}{r-1}, \text{ as } n \to \infty$$

$$\frac{\mathsf{MSW}}{p_A(1-p_A)(1-\theta)} \overset{p}{\to} 1, \text{ as } n \to \infty$$

$$F^* = \frac{\mathsf{MSB}}{\mathsf{MSW}} \overset{d}{\to} \frac{1+(n-1)\theta}{(1-\theta)(r-1)} \chi^2_{(r-1)}$$

227

Mean and Variance of Moment Estimator of θ

The mean and variance of a chi-square variable are the df and twice the df. So

$$\mathcal{E}(F^*) = \frac{1 + (n-1)\theta}{1 - \theta}$$

$$Var(F^*) = \frac{2}{r-1} \left(\frac{1+(n-1)\theta}{1-\theta} \right)^2$$

Now note that

$$\widehat{\theta} = \frac{F^* - 1}{F^* + n - 1}$$

so that

$$\mathcal{E}(\hat{\theta}) \approx \theta - \frac{1-\theta}{r-1} \left[\frac{(n-1)\theta+1}{n} \right]^2$$

$$\operatorname{Var}(\hat{\theta}) \approx \frac{2}{r-1} \left[\frac{(1-\theta)[(n-1)\theta+1]}{n} \right]^2$$

228

Confidence Interval for θ

$$\frac{(1-\theta)(r-1)}{1+(n-1)\theta}F^* \sim \chi^2_{(r-1)}$$

Look for an $100(1-\alpha)\%$ confidence interval, based on the middle $(1-\alpha)$ of the chi-square distribution. Write this interval as (L,U).

$$0.95 = \Pr(L \le \frac{(1-\theta)(r-1)}{1+(n-1)\theta} F^* \le U)$$
$$= \Pr(\theta_L \le \theta \le \theta_U)$$

Confidence Interval for θ

$$\begin{array}{l} \theta_L \; = \; \frac{(r-1)F^* - U}{(r-1)F^* + (n-1)U} \\ \\ \; = \; 1 - \frac{1}{\frac{n-1}{n} + \frac{r-1}{U} \left(\frac{\hat{\theta}}{1 - \hat{\theta}} + \frac{1}{n}\right)} \\ \theta_U \; = \; \frac{(r-1)F^* - L}{(r-1)F^* + (n-1)L} \\ \\ \; = \; 1 - \frac{1}{\frac{n-1}{n} + \frac{r-1}{L} \left(\frac{\hat{\theta}}{1 - \hat{\theta}} + \frac{1}{n}\right)} \end{array}$$

For example, suppose $r=5, n=100, \hat{\theta}=0.01$. With 4 df, (L=0.48, U=11.1) for a 95% interval. The 95% CI for θ is (-0.003, 0.173).

Confidence Interval for θ

If, instead,

$$\frac{(1-\theta)}{1+(n-1)\theta}F^* \sim F_{r(n-1)}^{r-1}$$

the CI becomes

$$\theta_L = 1 - \frac{1}{\frac{n-1}{n} + \frac{1}{U} \left(\frac{\hat{\theta}}{1-\hat{\theta}} + \frac{1}{n}\right)}$$

$$\theta_U = 1 - \frac{1}{\frac{n-1}{n} + \frac{1}{L} \left(\frac{\hat{\theta}}{1-\hat{\theta}} + \frac{1}{n}\right)}$$

For example, suppose $r=5, n=100, \hat{\theta}=0.01$. With 4 and 396 df, (L=0.36, U=2.79) for a 95% interval. The 95% CI for θ is (-0.003, 0.044). Not as wide as the interval based on the chi-square distribution.

230

231

Genetic Distance

Distance between populations designed to measure time since they diverged. Must therefore be based on a model of the divergence process.

Suppose two populations have allele frequencies $\{p_i\}$ and $\{q_i\}$. Euclidean distances of the form

$$d = \sqrt{\sum_{i} (p_i - q_i)^2}$$

are not based on an evolutionary model.

Pure drift model

Now define coancestry θ as the probability of two alleles being identical by descent (having a common ancestral allele). If mating is completely at random, within each population, the coancestry coefficient θ behaves like:

$$\theta_{t+1} = \frac{1}{2N} + (1 - \frac{1}{2N})\theta_t$$

so, if $\theta_0 = 0$,

$$\theta_t = 1 - (1 - \frac{1}{2N})^t$$

$$\ln(1 - \theta_t) = t \ln(1 - \frac{1}{2N})$$

$$\theta_t \approx t/2N$$

For small values of θ , it is approximately proportional to time and so can serve as a measure of distance.

232

Trees of Populations

Suppose data are available from a series of populations i. The indicator variable x_{ij} applies to the jth allele from the ith sample:

$$x_{ij} = \left\{ egin{array}{ll} 1 & ext{allele is } A \ 0 & ext{otherwise} \end{array}
ight.$$

Taking expectations over samples and over replicates of the evolutionary process,

$$\begin{array}{rcl} \mathcal{E}(x_{ij}) & = & p_A \\ \mathcal{E}(x_{ij}x_{ij'}) & = & p_A^2 + \theta_i p_A (1-p_A), \ j \neq j' \\ \mathcal{E}(x_{ij}x_{i'j'}) & = & p_A^2 + \theta_{ii'} p_A (1-p_A), \ j \neq j' \end{array}$$

where θ_i refers to population i and $\theta_{ii'}$ refers to the population from which populations i and i' diverged.

Trees of Populations

The usual analysis of variance format for variation among and within populations is, when n is the sample size from the ith population of r populations,

Source	d.f.	Mean Squares
Among	r-1	MSA
Within	r(n-1)	MSW

235

Trees of Populations

Can show that

$$\begin{split} \mathcal{E}(\mathsf{M}SA) &= p_A(1-p_A)(1-\bar{\theta}_w) \\ &+ np_A(1-p_A)(\bar{\theta}_w-\bar{\theta}_a) \\ \mathcal{E}(\mathsf{M}SW) &= p_A(1-p_A)(1-\bar{\theta}_w) \end{split}$$

where

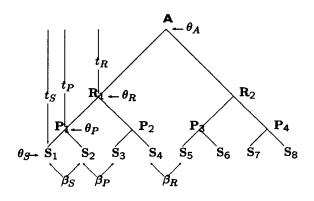
$$\bar{\theta}_w = \frac{1}{r} \sum_{i=1}^r \theta_i$$

$$\bar{\theta}_a = \frac{1}{r(r-1)} \sum_{i=1}^r \sum_{i' \neq i} \theta_{ii'}$$

This leads to an estimate, independent of allele frequencies, of

$$\beta = \frac{\bar{\theta}_w - \bar{\theta}_a}{1 - \bar{\theta}_a}$$

Trees of Populations



Idealized evolutionary history (A=ancestral, R=race, P=population, S=subpopulation.)

237

236

Trees of Populations

Populations	$ ilde{ heta}_w$	$ar{ heta}_a$	$\beta = (\bar{\theta}_w - \bar{\theta}_a)/(1 - \bar{\theta}_a)$
S_1, S_2	θ_S	θ_P	$rac{ heta_S - heta_P}{1 - heta_P}$
S_1, S_3	$ heta_S$	$ heta_R$	$\frac{\theta_S-\theta_R}{1-\theta_R}$
S_1, S_5	$oldsymbol{ heta}_{\mathcal{S}}$	θ_A	$\frac{\theta_S - \theta_A}{1 - \theta_A}$
S_1,S_2,S_3	θ_S	$\frac{\theta_P+2\theta_R}{3}$	$\frac{3\theta_S - \theta_P - 2\theta_R}{3 - \theta_P - 2\theta_R}$
S_1,S_2,S_5	θ_S	$\frac{\theta_P + 2\theta_A}{3}$	$\frac{3\theta_S - \theta_P - 2\theta_A}{3 - \theta_P - 2\theta_A}$
S_1, S_2, S_3, S_4	$\boldsymbol{\theta}_{S}$	$\frac{2\theta_P+4\theta_R}{6}$	$\frac{3\theta_S - \theta_P - 2\theta_R}{3 - \theta_P - 2\theta_R}$
S_1, S_2, S_3, S_4, S_5	$ heta_S$	$\frac{2\theta_P + 4\theta_R + 4\theta_A}{10}$	$\frac{5\theta_S - \theta_P - 2\theta_R - 2\theta_A}{5 - \theta_P - 2\theta_R - 2\theta_A}$
$S_1, S_2, S_3, S_4, S_5, S_6, S_7, S_8$	$oldsymbol{ heta}_{\mathcal{S}}$	40 _P +80 _R +160 _A 28	$\frac{7\theta_S - \theta_P - 2\theta_R - 4\theta_A}{7 - \theta_P - 2\theta_R - 4\theta_A}$

Trees of Populations

Under a pure drift model, and if all populations are of the same size N,

$$\begin{aligned} \theta_R &= 1 - \left(1 - \frac{2N - 1}{2N}\right)^{t_R} \approx \frac{t_R}{2N} \\ \theta_P &= 1 - \left(1 - \frac{2N - 1}{2N}\right)^{t_P} \approx \frac{t_P}{2N} \\ \theta_S &= 1 - \left(1 - \frac{2N - 1}{2N}\right)^{t_S} \approx \frac{t_S}{2N} \end{aligned}$$

so that, if θ_A is taken to be zero,

$$eta_R pprox rac{t_S}{2N} \ eta_P pprox rac{t_S - t_R}{2N - t_R} pprox rac{t_S - t_R}{2N} \ eta_S pprox rac{t_S - t_P}{2N - t_P} pprox rac{t_S - t_P}{2N}$$

238

239

Trees of Populations

The β 's therefore furnish estimates of time since most recent common ancestral population of each pair of populations and so serve as the basis for reconstructing trees of populations.

Note that this development has used moment estimators. A more flexible approach is provided by modern methods of variance component estimation.

Human disease genes

Relatives who are affected by the same disease are likely to share alleles at the disease locus. If more pairs of relatives than expected also share alleles at a marker locus, then that marker is likely to be "linked" to the disease locus. If the chromosomal location of the marker locus is known, this may help to locate the disease gene.

The expected occurrence of marker allele sharing depends on the degree of relationship, as well as the background evolutionary relationship as measured by θ .

240

Affected sib pair test

When the parents are drawn randomly from a population characterized by coancestry θ , the probability that sibs have the same marker genotype is

$$\begin{split} &\left(6\theta^3 + \theta^2(1-\theta)(23+21\sum_{i}p_i^2)/4 \right. \\ &+ \theta(1-\theta)^2(4+7\sum_{i}p_i^2 + \sum_{i}p_i^3)/2 \\ &+ (1-\theta)^3[\frac{1}{4} + \frac{1}{2}\sum_{i}p_i^2 \right. \\ &\left. + \frac{1}{2}(\sum_{i}p_i^2)^2 - \frac{1}{4}\sum_{i}p_i^4]\right)/(1+\theta)(1+2\theta) \end{split}$$

```
AFPC NSWC QLDC SAC TASC VICC WAC FBIC WAAB FBIH FBIB QLDA VICA

NUCC .000 .001 .001

SAC .000 .001 .001

TASC .000 .001 .001 .000

VICC .000 .000 .001 .000 .000

WAC .000 .000 .001 .000 .000 .000

WAC .000 .000 .001 .000 .000 .000

FBIC .000 .000 .001 .000 .000 .000

FBIC .000 .000 .001 .000 .000 .000

FBIB .012 .011 .010 .011 .012 .011 .011 .012 .017

FBIB .014 .013 .016 .016 .018 .015 .015 .014 .026 .026

QLDA .023 .019 .020 .022 .022 .019 .021 .026 .015 .019 .041

VICA .018 .015 .016 .017 .016 .014 .016 .020 .010 .019 .037 .000

VICV .026 .022 .026 .027 .026 .022 .025 .030 .019 .031 .041 .003 .000
```

242