ISSN: 2690- 912X

Research Article

Journal of Genetic Engineering and Biotechnology Research

Combining Genetic Similarities among Known Relatives that Connect to a Pair of Unknown Relatives

Stephen P Smith*, Cambrian Lopez and Nicole Lam

Kaiser Permanente Labor & Delivery, KSDHCPA (UNAC/ UHCP) Hospital President, USA

***Corresponding Author**
Bright Karim-Abdallah, University of Energy and Natural Resources, Quality Assurance Directorate and Department of Statistics

Submitted: 2023, Dec 04 ; **Accepted**: 2024, Jan 18; **Published**: 2024, Feb 08

Citation: Smith, S. P., Lopez, C., Lam, N. (2024). Combining Genetic Similarities among Known Relatives that Connect to a Pair of Unknown Relatives. *J Gene Engg Bio Res, 6*(1), 01-06.

Abstract

Smith, Lopez and Lam described how to combine genetic similarities, measured in centimorgans (cM), among declared relatives in an outside pedigree, and to concentrate those cM values into a single cM measurement for an envoy that is a representative of the outside pedigree. An unknown relative is presumed to be a descendant of the envoy, but has the cM values with relatives in the outside pedigree. That prior effort was a univariate analysis, where there is only one unknown relative with matches with others in the outside pedigree. The present paper presents a bivariate analysis, where there are two sisters that have matches with others in the outside pedigree. The cM values are now paired, where any DNA tested member of the pedigree has two cM values that match to both sisters. The bivariate analysis offers more efficient use of information, compared to two univariate analyses done for each sister in turn. This advantage comes with an increase in model complexity, in that a model is developed for treating three mutually exclusive categories representing genes found in the sisters: for genes in the first sister but not in the second sister; genes common to both sisters; or genes in the second sister but not in the first. The model is applied to the inheritance of the cM values in the pedigree. Even though the number of random effects is increased by a factor of three, the number of fixed effects that actually spend two degrees of freedom is unchanged from the univariate analysis. This is on top of the doubling of the number of observations for the bivariate analysis compared to one univariate analysis.

1. Introduction

DNA testing companies, Ancestry, My Family Tree DNA, My Heritage, 23andMe, offer their customers a collection of DNA matches to help locate family members. The matches are reported in centimorgans (cM), where the larger the cM measurement the closer the relationship to a customer, denoted by R. Unless the relationship to R is close, such as grandchild-grandparent or among 1st cousins, a single cM value is not that useful if the objective is to locate family that are more distant.

To get beyond the above limitation, Smith, Lopez and Lam [1], described a statistical method that combined a collection cM values from a cluster of relatives that were known among themselves (see Display 1), but collectively having an unknown relationship to R. Precisely, an envoy is attached to the cluster, where R is a descendant of the envoy, and the collection of cM values are statistically combined to provide a single cM estimate between R and the envoy. The envoy's estimated cM is computed with a statistical error to judge significance. Unlike a single cM value on a typical unknown relative, the envoy's cM (with R) can be notably large and indicating a real genetic path through the envoy to R that has been previously undiscovered. Smith, Lopez and Lam described their method for 14 cM values collected for each of two sisters (dark circles of Display 1). The details and data of that prior paper[1] will now be investigated again in the present paper, where adjustments are recommended for treating the fact that the two sisters contribute paired data (Table 1), rather than a single collection coming from one customer R. In short, Smith, Lopez and Lam applied the same set of calculations for each sister in turn, but a preferred analysis does a single joint analysis by incorporating the appropriate covariances for paired data.

[1]The prior paper serves as a prerequisite, and should be read first before the present paper is read.

Individual	ET	YP	KH	RM	JM	JK	SS	JS	LM	SK	PJ	KO	DP	TM
Sister 1	36	11	31	28	21	33.2	72	182	108	26	<6	15.5	60	60
Sister 2	15.5	11.4	29.4	26	<6	36	47	135	24.3	<6	<6	<6	39	13.3

Table 1. Ancestry's cM values between two sisters and 14 individuals that belong to the Rosa and Paula families; from Smith, Lopez and Lam.

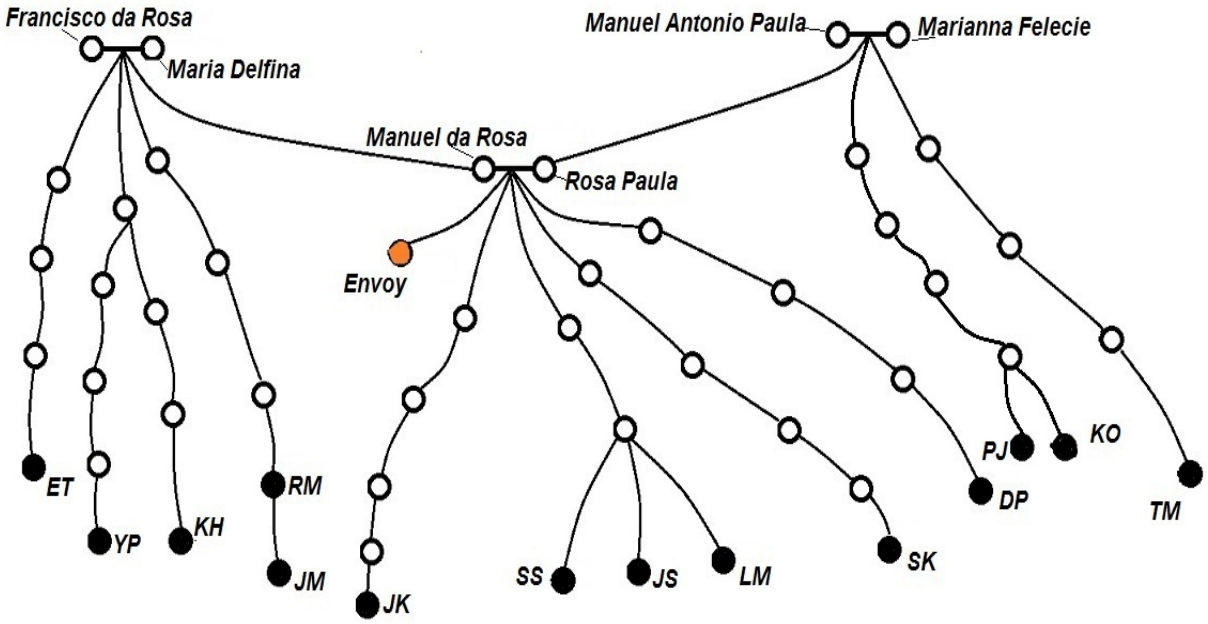

Figure 1: Family tree with central ancestors Manuel da Rose and Rosa Paula, with paternal common ancestors Francisco da Rosa and Maria Delfina, and with maternal common ancestors Manuel Antonio Paula and Marianna Felecie. Circles indicate individuals involved with gene flows that are common by decent between the envoy (red circle) and living descendants that took DNA tests (dark circles); from Smith, Lopez and Lam.

2. Statistical Model

A bivariate extension of the prior univariate analysis presented by [1] is straightforward, but only in principle. The fact is that the residual effects in the prior linear models that depict the inheritance pattern of the paired cM values, do not correspond to a superficially available variance-covariance matrix. In the univariate analysis the variances were found proportional to Pr(P=R) or Pr(O=R), where Pr(X=R) represents the probability of identity by descent that a random gene from X (i.e., from parent P or offspring O) is identical to a random gene from relative R. Or in the case of re-weighted iteration, the variances were found proportional to $\hat{u}_P$ or $\hat{u}_O$, where $\hat{u}_P$ or $\hat{u}_O$ are the imputed cM values for relative R matched with parent P and offspring O, respectively. No such proportionality adjustment was found for the bivariate case because the Markov chain properties of inheritance was found disrupting simple patterns, as will become apparent when the actual remedy is described.

An approach that worked, if not the only approach, was to partition the cM values, that are contained as entries in a vector **u** in Smith, Lopez and Lam, into three classes:

Class k=1 is where genes are common to Sister 1 and excluded from Sister 2; Class k=2 is where genes are common to both Sister 1 and Sister 2; and Class k=3 is where genes are common to Sister 2 and excluded from Sister 1. If the cM value for individual X in **u** (matched with one of the sisters in turn) is denoted by u_X, then $u_X=u_{X1}+u_{X2}$ for Sister 1 and $u_X=u_{X2}+u_{X3}$ for Sister 2, where u_{Xk} is the partition of u_X into Class k. For the univariate analysis there are two vectors **u** for each sister in turn, but in the bivariate analysis there is only one vector **u** with three times as many entries that contain u_{Xk} for all X and k. Originally there were 52 relatives where the cM where predicted (not counting the envoy and the envoy's descendants), and therefore, in the bivariate analysis **u** has length 156[2].

With the above partition of the cM values into three classes, the inheritance pattern can be treated as independent happenings, where the three categories (A, B and C) presented in Smith, Lopez and Lam can be applied independently for each class in turn, and this permits the construction of the appropriate linear model involving the vector **u** and a residual ϵ, as **Pu** = ϵ. Even this remedy is approximate, however, because its quite possible

[2]Actually, it will have length 158 because two fixed effects will be added to the end.

for DNA segments to overlap the classes. But this approximation is preferable to the grosser approximation of trying the map out the cM values unpartitioned by assigning a possible variance-covariance matrix for residuals.

The only modification that is needed before applying the rules presented in Smith, Lopez and Lam, is to recognize that the probabilistic sizes of u_{Xk}, relative to u_X in the univariate analysis, are cut in half. For the first round of iteration, the adjustment is made by using the multiplicative factor of 1700 when multiplying the probability of identity by descent, rather than 3400, but for re-weighted iteration (if there are any additional rounds) the adjustment is already implicit with the partition that implies that $u_X=u_{X1}+u_{X2}$ or $u_X=u_{X2}+u_{X3}$. The rules for the three categories of cM values, appropriately adjusted, and for the three classes in turn (k = 1, 2, 3), follow.

A. From One Parent to an Offspring, With the Parent Not A Common Ancestor or Central Ancestor.

If u_{Pk} is the partitioned cM measurement between any parent, identified as P, and the unknown relative R, then let u_{Ok} be the partitioned cM measurement between that parent's offspring, identified as O, and relative R. Moreover, define Pr(P = R) as the probability that a random gene taken from P (at a given loci) is identical by decent to a random gene taken from the relative R (at the same loci) that is now assumed to pass through the envoy by following a stipulated path.

It is apparent with meiosis and crossovers (i.e., genetic recombination) that half of the parents genes will be passed on to the offspring, implying that

$$u_{Ok} = \tfrac{1}{2}u_{Pk} + \epsilon_{Ok} \quad (1)$$

where ϵ_{Ok} is a random residual, with a variance that is approximated as:

Variance $(\epsilon_{Ok}) \approx 1700\times Pr(P = R)$ for the first round, or

Variance $(\epsilon_{Ok}) \approx \tfrac{1}{2}\times \hat{u}_{Pk}$ during re-weighted iteration, where a fresh prediction of u_{Pk} is available as $\hat{u}_{Pk}$.

B. From Common Ancestors, Or Central Ancestors, To an Offspring, When the Offspring Is Not One of the Central Ancestors.

As long as the flow of genes that are common to the unknown relative flow down from the two parents (P1 and P2) to an offspring (O), we can apply model (1) to represent the uniting gametes from the two parents. Any allele that is common to the unknown relative can only occupy one loci across both parents, and hence the common genes are passed on independently with (1) applied twice to give model (2).

$$u_{Ok}=\tfrac{1}{2}u_{P1k}+\tfrac{1}{2}u_{P2k}+\epsilon_{Ok} \quad (2)$$

The term ϵ_{Ok} is again the residual, but now with a variance that can be approximated by:

Variance $(\epsilon_{Ok}) \approx 1700\times[Pr(P1 = R)+Pr(P2 = R)]$ for the first round,

or

Variance $(\epsilon_{Ok}) \approx \tfrac{1}{2}\times[\hat{u}_{P1k} + \hat{u}_{P2k}]$ during re-weighted iteration.

C. From a Paternal (or maternal) Central Ancestor Back against the Flow of Genes to the Paternal (or Maternal) Common Ancestors.

The paternal common ancestors are the *parents* for the male central ancestor, and the maternal common ancestors are the *parents* for the female central ancestor. Here gene flow is reversed to place the common genes, i.e., found identical in the unknown relative, that are also in the central ancestors, but now finding them in the noted parents. This makes two equations given by (3) for the parents rather than one for the offspring, and done for both the paternal and maternal sides of the central ancestors.

$$\begin{aligned} u_{P1k}&=\tfrac{1}{2}u_{Ok} + \epsilon_{P1k} \\ u_{P2k}&=\tfrac{1}{2}u_{Ok} + \epsilon_{P2k} \end{aligned} \quad (3)$$

Every common allele (at a particular loci) found in P1 is an allele missing in P2, and vice versa. Therefore ϵ_{P1k} and ϵ_{P2k} have a perfect negative correlation, and the associated 2×2 variance-covariance matrix is a rank-1 matrix, approximated by the following.

$$Var\begin{bmatrix}\varepsilon_{P1k}\\ \varepsilon_{P2k}\end{bmatrix}=\begin{bmatrix}v & -v\\ -v & v\end{bmatrix}$$

Where

v = 1700×Pr(O=R) for the first round, or

v = $\tfrac{1}{2}\hat{u}_{Ok}$ during re-weighted iteration.

The residuals in equation (3) are meant to have the *conditional* variance-covariance matrix given above, i.e., conditional on u_{Ok} as opposed to being unconditional, and its easy to conflate an unconditional variance or covariance with a better approximation that it is not. The conditional distribution is better approximated as a binomial distribution, rather than a Poisson distribution, and this leads to the above singular variance matrix once v, or the mean of the distribution, is substituted and allowances are made for extra-Poisson variation that acts as a proportionality constant over the variances and covariances for all residuals. This pattern of using a multinomial distribution to derive a variance matrix (the binomial distribution in the above example), followed by the substitution of the mean of the distribution, is again employed below where specifications were found required of the central ancestors but now using a trinomial distribution.

D. Equations for Central Ancestors.

There are now equations and residuals coming form (1), (2) and (3) for all individuals in the pedigree that is to be analyzed, excluding the central ancestors and the envoy, and the envoy's descendants leading to the unknown relative. But new equations for the central ancestors are needed too, because absent any other adjustment the

bivariate analysis will end up spending six degrees of freedom in estimating all six fixed effects that are otherwise associated with the two central ancestors. In the univariate analysis there were only two fixed effects associated with the central ancestors, but in the present situation the partitioning inflates that number by a factor of three. It had been advantageous to let those prior fixed effects float freely to absorb all the 14 cM values that had been measured (for each sister in turn). But floating all six parameters is overkill and creates instabilities as well as using up too much information, even if the total information grows to 28 cM values for both sisters. The bivariate analysis should result in a more efficient use in information, not less, and so a new adjustment is required.

Any central ancestor (denoted by C), will also show an expected cM value with relative R given by 6800×Pr(C=R), but this falls to 3400×Pr(C=R) for the expected cM values representing each of the three classes or partitions, for u_{C1}, u_{C2} and u_{C3}. Every gene passed from ancestor C to at lease one of the sisters has an equal a chance of falling in any of the three partitions. If there are N_e effective DNA segments of length "cM, segregating independently, more or less, then the trinomial distribution is available to define the needed prior distribution[3] of which the mean vector and variance matrix are sought and are listed below.

$$\alpha N_e \begin{bmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{bmatrix} \qquad \alpha^2 N_e \begin{bmatrix} \frac{2}{9} & -\frac{1}{9} & -\frac{1}{9} \\ -\frac{1}{9} & \frac{2}{9} & -\frac{1}{9} \\ -\frac{1}{9} & -\frac{1}{9} & \frac{2}{9} \end{bmatrix}$$

The mean value for all three classes, symbolically u_C, had been approximated as 3400×Pr(C=R). The substitution $u_C = ⅓\times\alpha N_e$ is plugged into the variance matrix to produce the following.

$$\alpha u_C \begin{bmatrix} \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \end{bmatrix}$$

This matrix must still be calibrated with all the other residual variances used for categories A, B and C above, because it was implicit that ½α was a common factor that became part of the extra-Poisson variation. Following calibration[4], the equations for the both central ancestors are found, and are presented next.

$$(4) \qquad \begin{aligned} u_{C1} &= u_C + \epsilon_{C1} \\ u_{C2} &= u_C + \epsilon_{C2} \\ u_{C3} &= u_C + \epsilon_{C3} \end{aligned}$$

Where the variance matrix for ϵ_{C1}, ϵ_{C2} and ϵ_{C3} is given by the following.

$$6800 \times \Pr(C = R) \begin{bmatrix} \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \end{bmatrix}$$ for the first round,

or

$$2 \times \hat{u}_C \begin{bmatrix} \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \end{bmatrix}$$ during re-weighted iteration.

[3]A prior distribution is defined in the Bayesian context, where it is sampled once at the beginning of the Markov processes, and not repetitively downstream.

[4]That is, dividing by ½ α

The mean value u_C, and its estimate $\hat{u}_C$, are different for each central ancestor; i.e., there are two sets of them, and now estimation spends two degrees of freedom as originally sought. The two u_C values are treated as a fixed effects, and are floated to represent all the cM measurements.

This information in equations (1), (2), (3) and (4) can be collected and expressed in matrix notation as follows

$$\mathbf{Pu} = \epsilon$$

$$\text{Var}(\epsilon) = \mathbf{R}$$

Where **P** is a rectangular matrix with 156 rows and 158 columns, with each row representing an equation of the form (1) or (2) or where two rows are given by (3) or three rows given by (4), where most elements in any row are set to zero expect for the numbers 1, -1 or -½ that are found at the appropriate places. The column vector , represents the collection of 156 residuals. The variance matrix **R**[5] is almost completely diagonal, except for six 2×2 blocks that correspond to the paternal and maternal common ancestors and two 3×3 blocks that correspond to the central ancestors . The matrix **R** is rank deficient[6] and this property is treated correctly with the matrix tools described by Smith, Lopez and Lam.

The data are the paired cM values found on living relatives that also belong to the terminal ends of the pedigree presented in Display 1. The linear model for the observations is the following.

y=Zu

Where **y** a column vector of length 28 containing the 14 paired cM values observed on some of the living relatives, and **Z** is an incidence matrix of 28 rows and 158 columns containing mostly zeros except for two entries containing the number one in each row that picks out the appropriate two elements in **u** so that it is matched with the corresponding element in **y**. These are linear equations of the form: $u_X=u_{X1}+u_{X2}$ for Sister 1 and $u_X=u_{X2}+u_{X3}$ for Sister 2. The linear model for the observations contains no additional error terms, following the standard presented by Smith, Lopez and Lam.

The matrices (**P, Z** and **R**) and the vector (**y**) are now plugged into the matrix **M** below, following Smith, Lopez and Lam.

$$\mathbf{M} = \begin{bmatrix} \mathbf{R} & & \mathbf{P} & \\ & & \mathbf{Z} & \mathbf{y} \\ \mathbf{P}^T & \mathbf{Z}^T & & \\ & \mathbf{y}^T & & \end{bmatrix}$$

The empty space in **M** is understood to represent entries of zero. The only modifications necessary to the 6-step protocol of Smith, Lopez and Lab involve steps 3 and 6. For the sake of completeness and flow, all 6 steps are repeated again, with modifications.

• The matrix **M** is constructed as described above, using approximate variances derived for the Poisson distribution and the probabilities of identity by descent.

• A permutation matrix **Q** is found dynamically with the implementation of the LU factorization (see Smith 2001), to compute the unit lower triangular matrix **L** and an upper triangular matrix **U** such that $\mathbf{LU=QMQ}^T$, while restricting the permitted permutations to leave the last row and column of **M** fixed in the last position.

• The chi-square statistic (x^2) with 2×(N – 1) degrees of freedom[7] is retrieve in the last diagonal element of **U** which is present as $-x^2$. The expectation is that this statistics will show significance because the Poisson distribution comes with small relative variances, and it is therefore easy to generate a poor fit. Significance implies the presence of extra-Poisson variation, with variance term σ^2 noted below.

$$\sigma^2 = \frac{\chi^2}{2N-2}$$

• To calculate the predictions of shared cM for all the known relatives (i.e., to calculate the prediction of the vector **u**), retrieve the last column $\bar{U}$ but excluding the last element where the chi-square statistic was found, and put it in the work vector **r**. Remove the last row and column of **U**, making a smaller upper triangular matrix **U**. The column vector **r** has already been subjected to implicit forward substitution with the LU factorization. Complete the process now by solving **s** in $\bar{U}\hat{\mathbf{s}}=\mathbf{r}$ by backward substitution. The prediction of **u**, now defined as $\hat{\mathbf{u}}$, is found scattered in **s** depending on the permutations. However, because the permutations are done implicitly by software, $\hat{\mathbf{u}}$ is found in $\hat{\mathbf{s}}$ as if there had been no permutations.

• With the chi-square statistic significant, the matrix **M** can be rebuilt for re-weighted iteration by using the current value of $\hat{\mathbf{u}}$. The calculation then returns to Step 2 above, and this iteration repeated as many of times as necessary until the chi-square statistic stabilizes. Ideally, the chi-square statistic should fall initially, if only a little, otherwise reweighted iteration is not recommended. Once this is done, the last estimate of σ^2 found in Step 3 is taken as the extra-Poisson variation.

• To predict the shared cM for the envoy add the appropriate cM predictions for thecentral ancestors (C=A or B in turn) together; i.e., add four elements ($\hat{u}_{A1}+\hat{u}_{A2}+\hat{u}_{B1}+\hat{u}_{B2}$) of û together for Sister 1 and add four elements ($\hat{u}_{A2}+\hat{u}_{A3}+\hat{u}_{B2}+\hat{u}_{B3}$) for Sister 2. Initialize the work vector r used in Step 4 to zero everywhere except for the respective four entries that correspond to the central ancestors (i.e., corresponding to $\hat{u}_{A1}$, $\hat{u}_{A2}$, $\hat{u}_{B1}$, $\hat{u}_{B2}$ for Sister 1 or $\hat{u}_{A2}$, $\hat{u}_{A3}$, $\hat{u}_{B2}$, $\hat{u}_{B3}$ for Sister 2) that are set to the number one. With the permutations treated implicitly, use forward substitution to solve for the vector **s** in $\bar{U}^T\mathbf{s}=r$, for each sister in turn. Calculate the negative weighted sum of squares,

$$\delta^2 = -\sum_i \bar{U}_i s_i^2$$

[5]That is, a matrix R with its rows and columns that are suitably permuted.

[6]The 2×2 blocks have rank 1, and the 3×3 blocks have rank 2.

[7]N=14, so there are 26 degrees of freedom for the bivariate analysis and the paired data set, a number confirmed by Fact 10 of Smith [2].

where $\bar{U}_i$ is the i-th diagonal of $\bar{U}$, and s_i is the i-th element of **s**. The standard error for the shared cM prediction for the particular sister matched with the envoy is $\sigma\delta=(\sigma^2\delta^2)^{½}$.

3. Numerical Example

3.1 Stability Issues

When the calculation were applied to the data in Table 1, for the bivariate analysis, it was discovered that the chi-square statistic (initially at 208.9 with 26 degrees of freedom) did not fall with re-weighted iteration and so re-weighted iteration was not performed with the calculations stopping after one round. In the univariate analysis, reweight iteration was performed for Sister 1, but not for Sister 2. The predicted cM values tended to be smaller for Sister 2 than Sister 1. Because of partitioning the predicted cM values in the bivariate analysis also tended to be smaller, as much as 50%.

It would be a stability issue that arises more generally, if small cM values tended to prohibit the initiation of re-weighted iteration by showing no further reduction in the chisquare statistic.

Six of the predicted cM values, out of a total of 156, were predicted to be negative. Because these negative numbers were for individuals located at the terminal ends of the pedigree, they never had a chance to enter as weights during re-weighted iteration. Only positive weights can be used.

In contrast, none of the 52 cM values were predicted negative for the univariate analysis. But with partitioning, and a more complicated linear model, the opportunity for negative cM predictions to arise increases, and this possible instability should be checked.

The occurrence of negative cM predictions has an easy remedy, however. A Bayesian trick[8] can be used to restrict the few troublesome predictions to zero, or near zero. The matrix **M** has a negative block, initially set to a sub-matrix containing only zeros and occupying the lower right corner. The diagonal elements of that sub-matrix have a oneto-one correspondence to the elements of **u**. The corresponding diagonal elements of **M**, that correspond to negative predictions in **u**, can be initialize to a large negative number, like the -1000 that was actually used. This will force the few predictions to follow a Bayesian prior that is concentrated sharply around zero, making the predictions come out close to zero. In theory, this may be an iterative process because forcing some predictions to zero might also force other predictions to become negative. However, iteration was not required with the bivariate analysis once restrictions were set in place for the 6 cM predictions that had been negative.

3.2 Statistical Results

Table 2 presents the side by side comparisons of the bivariate and univariate analyses, listing the cM predictions (for the sisters matched with the envoy), standard errors and 95% confidence regions. The two types of analyses have very similar results, with Sister 1 showing a stronger signal than Sister 2. The respective standard errors dropped by 6% and 3%, for Sister 1 and 2, respectively. A reduction was expected because the bivariate analysis is more efficient than the univariate analysis, but this was a small improvement.

Sister	cM of Envoy	Standard Error	95% Confidence Region
Univariate Analysis			
Sister 1	1003.7	±157.8	cM>744.9
Sister 2	566.8	±153.9	cM>314.4
Bivariate Analysis			
Sister 1	890.1	±148.7	cM>646.2
Sister 2	743.9	±148.7	cM>500.0

Table 2. Side by side, comparison of two univariate analysis compared to one bivariate analysis.

The big effect of the bivariate analysis was to bring the results for the two sisters closer together. As reports that the cM values on a greatgrandparent can vary between 547 to 1110, the bivariate analysis brings a closer agreement with the possibility that the envoy is a great-grandparent of both sisters [3].

References

1. Smith, S. P., Lopez, C., & Lam, N. (2017). *Combining Genetic Similarities Among Known Relatives that Connect to an Unknown Relative.* Memo, Achieved in Quantitative Biology of viXra.
2. Smith, S. P. (2001). The factorability of symmetric matrices and some implications for statistical linear models. *Linear Algebra and its Applications, 335*(1-3), 63-80.
3. Bettinger, B. T. (2019). *The Family Tree guide to DNA testing and genetic genealogy.* Penguin.

[8]See Section 3.3 of Smith [2] describing non-negativity constrains for interior-point methodology for linear programming.