

Genetic Factors in the Aetiology of Mouth Ulcers

Robert I.E. Lake,^{1*} Steve J. Thomas,^{1,2} and Nicholas G. Martin¹

¹Queensland Institute of Medical Research, Brisbane, Australia

²College of Medicine, University of Wales, Cardiff, Wales

The aetiology of mouth ulcers is uncertain, and prior research has indicated both environmental and genetic factors. In this study, information on mouth ulcer incidence was collected for 290 twin pairs—127 monozygous (MZ) and 163 dizygous (DZ)—and their parents, a total of 1,160 people. Self-reported data on mouth ulcer incidence were available for the twins, and in each family the mother also reported on the mouth ulcer incidence in the twins, the twins' father, and herself. A structural equation model—combining a measurement model, a rater bias model, and a model including genetic and environmental influences—was used to explain variation in mouth ulcers. The fitted model explained the variation in a latent phenotype of mouth ulcer incidence for the twins in terms of an additive genetic factor (64%), a common environment factor (26%), and a specific environment factor (10%). The mothers' ratings showed a significant positive bias. *Genet. Epidemiol.* 14:17–33, 1997. © 1997 Wiley-Liss, Inc.

Key words: aphthous ulcers; twin studies; genetic influences

INTRODUCTION

Recurrent aphthous stomatitis (RAS), commonly known as mouth ulcers, is a painful condition that affects over 20% of the population [Embil et al., 1975] but little is known about its aetiology [Porter and Scully, 1991]. RAS ulcers may be single or multiple, usually occurring initially in the second decade and frequently recurring more than once per year [Axell and Henricsson, 1985]. The role of environmental risk factors in the aetiology of RAS is poorly understood. Deficiencies in

Contract grant sponsor: Queensland Cancer Fund; Contract grant sponsor: The National Health and Medical Research Council.

*Correspondence to: Robert Lake, Epidemiology and Population Health Unit, Queensland Institute of Medical Research, Post Office, Royal Brisbane Hospital, Queensland 4029, Australia.

Received 21 March 1996; Accepted 1 August 1996

© 1997 Wiley-Liss, Inc.

iron, folic acid, and vitamin B12 are more common in RAS patients [Wray et al., 1975]. An association has been noted between psychiatric factors and RAS but not in all studies [Pedersen, 1989]. Viral or bacterial mechanisms may operate, and RAS displays some features consistent with a viral aetiology [Pedersen, 1991]. RAS is also a feature of HIV/AIDS [Liang et al., 1993]. Rarely, RAS can be a manifestation of an underlying systemic disorder such as Behçet's syndrome. However, in Behçet's the HLA type associated with mouth ulcers differs from that associated with other aspects of the disease [Ozbakir et al., 1987].

In general HLA studies of patients with RAS have been inconclusive, although links with HLA A, B, C, and DR have been suggested [Porter and Scully, 1991].

From the foregoing it is apparent that RAS may be due to a range of environmental and genetic factors. The search for putative causal factors will be enhanced by the knowledge of whether these factors are likely to be genetic or environmental. The twin method—utilising identical and fraternal twins—is a powerful tool to decompose the variance of a phenotype into genetic and environmental components.

Miller et al. [1977] examined the incidence of RAS in 19 young twin pairs (12 MZ and seven DZ, mean age 13.1 years) and in six pedigrees comprising 318 individuals. They concluded that a genetic factor was involved on the basis of high concordance rates in MZ twins versus DZ twins (92% vs. 57%) and a stable occurrence rate of about 40% over generations in the six pedigrees but did not find a specific mode of inheritance. Their study was limited because they did not conduct any statistical test of significance and because the small size of the study limited the power to quantify or detect genetic effects. Structural equation models can be used to partition the variance of a phenotype, and sophisticated models have been developed in the area of behavioral genetics using twin and family studies [Boomsma et al., 1989; Neale and Cardon, 1992]. In this study we use structural equation models to examine data on mouth ulcers in 290 young twin pairs and their parents. The aim of the study is to clarify the contributions of genetic and environmental factors in the aetiology of RAS.

MATERIALS AND METHODS

Study Protocol

The subjects were 290 family groups, defined here as a twin pair and their biological parents. The subjects initially were recruited by writing to the principals of all primary schools in the Brisbane Statistical Division, seeking twins aged 10 to 12 years for participation in a longitudinal study of melanocytic naevus development. Subjects were also ascertained by word of mouth, and as a result of media appearances by investigators from the study. Informed consent was received from all subjects. The twins and usually the mother but not the father were examined at the Queensland Institute of Medical Research over the period June 1992 to August 1995.

Data were collected for 306 family groups; however, data were missing for the father in 16 groups, leaving a final sample size of 290 groups. The breakdown by zygosity for the twin pairs was monozygotic (MZ) female pairs, 65; MZ male pairs, 62; dizygotic (DZ) female pairs, 42; DZ male pairs, 50; DZ opposite sex pairs, 71. Zygosity was determined by similarity of appearance or, in doubtful cases, by concordance of eight highly polymorphic DNA microsatellite markers (MZ) or discordance of two markers (DZ).

Measures

The response variable was the incidence of mouth ulcers. During a clinical visit for a count of moles, the twins and their mother individually completed questionnaires on a range of life style, health, and sun exposure matters. The questionnaire for the twins contained the question: "Have you ever had mouth ulcers?" The mother's questionnaire had the questions: "Has Twin A [or B] ever had mouth ulcers?"; "Have you ever had mouth ulcers?"; and "Has the Twin's father ever had mouth ulcers?" The response scale was 1, frequently; 2, sometimes; 3, rarely; and 4, never.

If either twin had mouth braces, both twins were told about the distinction between trauma or abrasions associated with braces and mouth ulcers. If the twins or their mother asked questions about mouth ulcers, for example "What is a mouth ulcer?", they were shown photographs of mouth ulcers, and had explained to them the distinction between mouth ulcers and cold sores (herpes simplex type I). This only occurred with about half of subjects.

There were six response variables per family group. Two values were available for mouth ulcer incidence for each twin: a self-report and the mother's report. The mother also reported on her own mouth ulcer incidence and on the mouth ulcer incidence in the father of the twins.

Test-Retest Sample

A subsample of 33 family groups was retested, early in the study, between 2 and 11 weeks (median 4 weeks) after their initial visit.

Statistical Methods

Correlations. The correlations between response variables are estimated with the assumption that liability to mouth ulcers is a normally distributed continuous variable. The four categories of the ordinal measure of mouth ulcer incidence represent four contiguous intervals of the continuous liability distribution, and these intervals can be defined by three thresholds [Olsson, 1979]. The thresholds can be estimated using maximum likelihood by the analysis package PRELIS [Jöreskog and Sörbom, 1988], and from these thresholds standardised scores for each category can be determined to give a normally distributed variable with zero mean and unit variance. This approach has the advantage, for multivariate analysis, of stabilising the variance of the response variables. Multivariate methods are sensitive to heteroscedasticity and departures from normality, and both are optimised by this method.

Genetic analysis. The variance of a variable can be decomposed into hypothesised genetic and environmental sources. Structural equation modelling, as operationalised by software packages such as LISREL [Jöreskog and Sörbom, 1989] or Mx [Neale, 1995], allows a range of genotype-environmental models to be fitted to twin data. Hypotheses about various combinations of environmental and genetic factors can then be tested [Neale and Cardon, 1992].

The contributions of genetics and environment to variation in the phenotype are modelled in the twin method by exploiting the fact that monozygotic (MZ) twins share 100% of their genes and dizygotic (DZ) twins share, on average, 50% of their genes. It can be shown that the correlation for additive genetic influence is 1 for MZ twins and 0.5 for DZ same-sex twins, and that the correlation for dominant (nonadditive) genetic effects is 1 for MZ twins and 0.25 for DZ same-sex twins.

The possibility exists that the phenotypic expression of the genotype is different between males and females (scalar sex limitation) or that different genes influence the phenotype in males and females (nonscalar sex limitation). Two indications of sex limitation are that the moments of the distributions of the phenotype between males and females are different and that the phenotypic correlation for DZ opposite-sex twins is lower than the harmonic mean of the phenotypic correlations for DZ same-sex twins. In the case of nonscalar sex limitation, the correlation for additive genetic influence between DZ opposite-sex twins may be less than 0.5 and should be estimated as a parameter in the model.

In this study we do not model dominant factors, as Martin et al. [1978] showed that the power of a twin study to detect dominant factors is weak with small sample sizes.

The analysis of genetic influences is enriched when data on the parents are available as, under Mendelian inheritance, the correlation between a parent's genotype and a child's genotype will be 0.5.

The environmental component of the variation in the phenotype can be separated into aspects common to both twins (shared environment) and aspects unique to each family member (specific environment). By definition, the correlation of shared environment between twins—either MZ or DZ—is 1, and the specific environment correlation between family members is zero.

Measurement model. A measurement model is defined with two or more observed variables and the assumption that these variables are correlated because they have a single factor (a latent or unobserved variable) in common. The latent variable is the “true” value of the phenotype, and each observed variable is a measure, with error, of the latent variable [Jöreskog and Sörbom, 1989]. The measurement model has the advantage in the context of a genetic analysis in that it allows the variance of the “true” variable to be partitioned without the distraction of error variance contributing to the imperfect indicator variables. It also allows quantification of the extent to which the indicators measure the “true” variable.

A measurement model may be congeneric (the most general form), tau-equivalent, or parallel, depending on the relationship between scale values for the variables and the error terms [Bollen, 1989; Jöreskog and Sörbom, 1989]. In this study the latent variable is liability to mouth ulcers, and there are two observed indicators for each twin (self-report and mother's report). Although the observed variables are responses to similar questions using the same response scale, a conservative approach is to use a congeneric measurement model since this makes no assumptions about equality of scale scores or error terms for each observed variable. The congeneric model allows the observed variables to vary in both accuracy and precision. For instance, the cumulative temporal aspect to the measures means that the mother may be more reliable, owing to her age and experience, than the twins. Additionally, the phenotype may be more salient to the mother in her role as the carer and nurturer within the family unit, leading to a more accurate rating by the mother.

A distinction can be made between a “pure” measurement model—in which the observed indicator represents only the “true” value of a latent variable plus measurement error—and an “impure” measurement model—in which one or more observed variables may be manifestations both of the latent phenotype and of other latent variables. In this study the measurement model is used in a genetically informative context, and thus there are genetic and environmental components of variance for

the two observed variables for each twin as well as for the latent phenotype. If each indicator measures only the latent phenotype (with error) then these genetic and environmental components will not be significant.

Rater bias. Rater bias can occur when there are measurements by the one rater on more than one person. Rater bias in twin or family analysis can have two components. The first component is the bias of location where one rater will rate consistently higher than another. The second component is the bias of association, or “halo” bias, where the rater gives ratings based on other characteristics, in this case members of a group are rated similarly because of their group membership.

In a rater bias model these two components of bias can be specified if ratings are available on group members and non-group members—with the group being the twin pair. In a structural equation model, bias is specified as a latent variable which contributes to the variance of the observed variables [Hewitt et al., 1992]. The variance of a phenotype is modelled as the sum of the effects of the “true” variance, the error variance, and the variance due to the rater’s mean bias and the rater’s association bias.

Structural equation models. We develop two structural equation models: a measurement model involving only the two measurements of mouth ulcer incidence for each twin; and a combined model which includes a rater bias component and the measurements for mouth ulcer incidence in the parents. The principal reason for developing the simpler measurement model, which does not use all the information available, is to validate the combined model.

The structural equation model for the measurement model is represented by the path diagram in Figure 1. The observed variables are the mother’s report on each twin ($m-r_1$, $m-r_2$), and the twins’ self-report ($s-r_1$, $s-r_2$). The latent variables (P_{T1} and P_{T2}) represent the phenotype of liability to mouth ulcers in the twins, manifested by two observed variables—self-report and mother’s report of ulcer incidence. The path coefficients between the latent phenotype and the observed variables $s-r$ and $m-r$ are β_M and β_S respectively.

In the measurement model shown in Figure 1 the components of variance in the latent phenotypes (P_{T1} and P_{T2}) are contributed by the latent genotype A , latent common environment variable C , and the latent unique environment variable E . The components of variance for the observed variables ($m-r_1$, $m-r_2$, $s-r_1$, $s-r_2$) are the contribution of the latent phenotypes, the unique environment latent variables E_M and E_S . The latent variables X_M and X_S can be either the genetic components A_M and A_S or the common environment components C_M and C_S . The correlation between twin genotypes is specified as 1 for MZ twins and 0.5 for DZ twins. In addition, under conditions of no sex limitation, the paths can be equated across sex. The path coefficients are a , a_M , and a_S .

The structural equation model for the combined model is represented by the path diagram shown in Figure 2. In addition to the model represented in Figure 1 there are the observed variables of mother’s report on herself (moth) and on the father (fath), and a latent variable representing the mother’s rater bias.

In the combined model the genetic components of variance are contributed by latent genotypes (A_{FA} for the father, A_{MO} for mother, and A , A_M , A_S for offspring). The correlation between parental genotype and offspring phenotype is specified as 0.5, representing the genetic contribution each parent makes to the child. The corre-

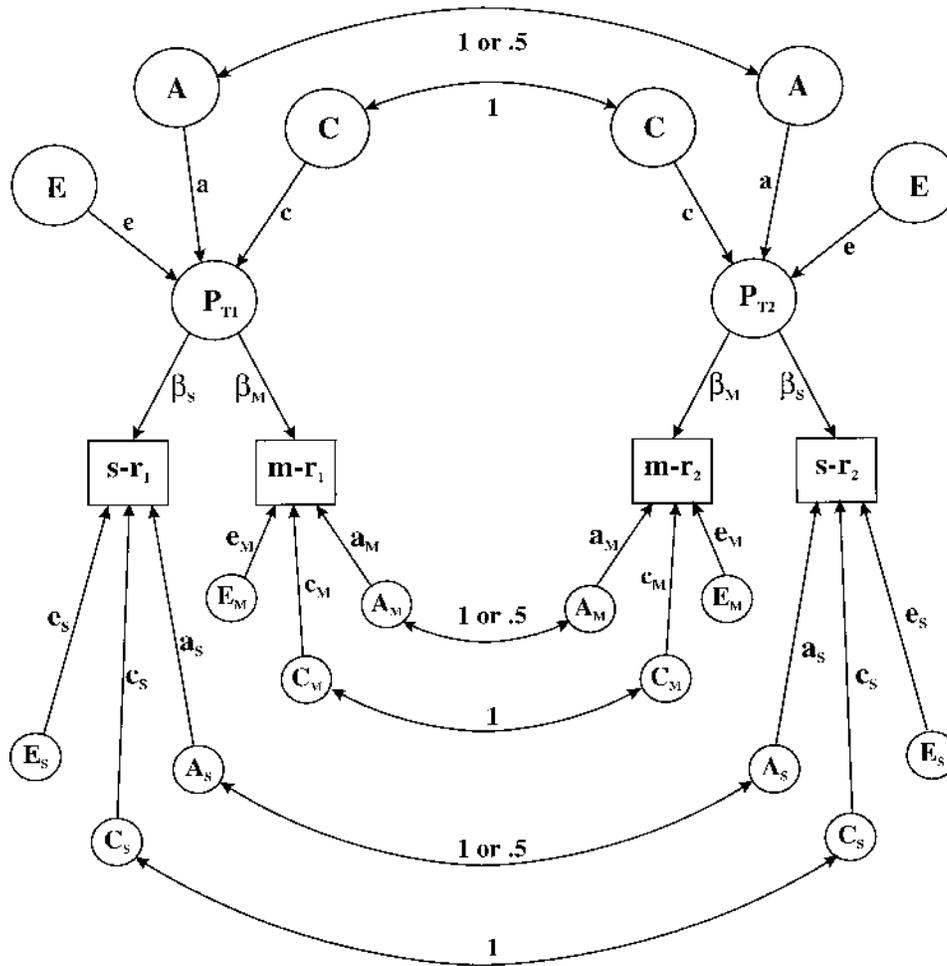


Fig. 1. The measurement model: path diagram showing the relationships between observed variables [twin self report (s-r) and mother's report on twin (m-r)] and hypothesised latent variables [additive genetic factors (A , A_M , A_S); common environment factors (C , C_M , C_S); unique environment factors (E , E_M , E_S); and the latent phenotype (P) for twins].

lation between twin genotypes is specified as 1 for MZ twins and 0.5 for DZ twins. There is no path between parental genotypes—an assumption of this model is that there is no assortative mating based on liability to mouth ulcers and, consequently, no direct relationship between parental genotypes. The path coefficients for the latent genetic variables for parents and offspring can be equated since it has been shown that, in the absence of cultural transmission or assortative mating on the phenotype, the loading of the parental genotype on the parental phenotype is equal to the loading of the offspring genotype on the offspring phenotype [Martin et al., 1986; Eaves et al., 1990]. In addition, under conditions of no sex limitation, the paths can be equated across sex. The path coefficients are a , a_M , and a_S .

The shared environment components of variance are contributed by the latent

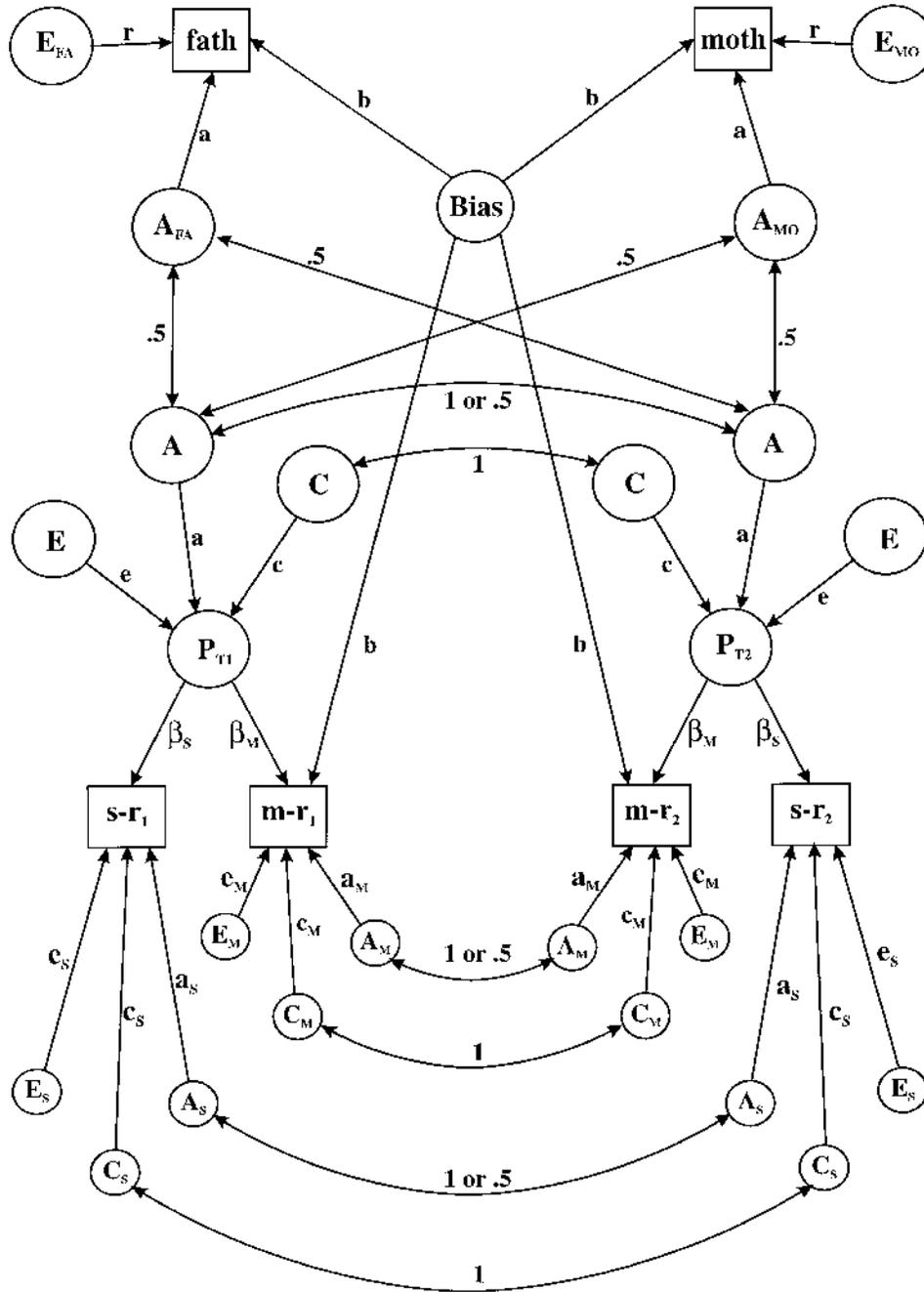


Fig. 2. The combined measurement and rater bias model: path diagram showing the relationships between observed variables [mother's report on father (fath); mother's self-report (moth); twin self-report (s-r); and mother's report on twin (m-r)] and hypothesised latent variables [additive genetic factors (A_{FA}, A_{MO}, A, A_M, A_S); common environment factors (C, C_M, C_S); unique environment factors (E_{FA}, E_{MO}, E, E_M, E_S); rater bias (Bias); and the latent phenotype (P) for twins].

variables C , C_M , and C_S with path coefficients c , c_M , and c_S respectively. The unique environment components of variance are contributed by the latent variables E , E_M , and E_S . The path coefficients are r for parents and e , e_M , e_S for offspring.

The mean rater bias is represented by the latent variable Bias, which contributes to the variance of fath, moth, m-r₁, and m-r₂ with path coefficient b . Any association bias caused by the mother's rating the twins more similarly is confounded by the latent variable C_M . An assumption in this model is that the components of variance of the variables fath and moth are equivalent, which may not hold since one (moth) is a self-report and the other is not. An indicator of the equivalence of these measures is the first- and second-order statistics.

Table I shows the algebraic expectations for each statistic for the combined model.

Model fitting. In the context of a measurement model the genetic and common environment variance components of s-r and m-r will be zero if s-r and m-r are manifestations only of the latent phenotype. However, given the rater bias model, the variance component of m-r due to C_M may be non-zero if an association rater bias exists. The path coefficient b will be non-zero if a mean rater bias exists.

We use an hierarchical approach to model fitting, which reflects the theoretical basis of the model. The first stage is to develop the most parsimonious measurement and rater bias model. The likelihood ratio chi-square test is used to assess the effect of the components relating to the measurement and rater bias aspects of the model (A_M , A_S , C_M , C_S , and Bias). The specific environment components (E_M , E_S) are not tested because they represent measurement error.

TABLE I. Algebraic Expectations for Variances and Covariances From the Combined Model*

Variance(m-r)	=	$\beta_M^2(a^2 + c^2 + e^2) + b^2 + a_M^2 + c_M^2 + e_M^2$
Variance(s-r)	=	$\beta_S^2(a^2 + c^2 + e^2) + a_S^2 + c_S^2 + e_S^2$
Variance(fath)	}	$r^2 + a^2 + b^2$
Variance(moth)		
Covariance(s-r ₁ ,m-r ₁)	}	$\beta_S\beta_M(a^2 + c^2 + e^2)$
Covariance(s-r ₂ ,m-r ₂)		
Covariance(m-r ₁ ,m-r ₂)	=	$\beta_M(H.a^2 + c^2) + b^2 + H.a_M^2 + c_M^2$
Covariance(s-r ₁ ,s-r ₂)	=	$\beta_S(H.a^2 + c^2) + H.a_S^2 + c_S^2$
Covariance(s-r ₁ ,m-r ₂)	}	$\beta_S\beta_M(H.a^2 + c^2)$
Covariance(s-r ₂ ,m-r ₁)		
Covariance(fath,moth)	=	b^2
Covariance(fath,s-r ₁)	}	$.5\beta_S a^2$
Covariance(fath,s-r ₂)		
Covariance(moth,s-r ₁)		
Covariance(moth,s-r ₂)		
Covariance(fath,m-r ₁)	}	$.5\beta_M a^2 + b^2$
Covariance(fath,m-r ₂)		
Covariance(moth,m-r ₁)		
Covariance(moth,m-r ₂)	=	$.5\beta_M a^2 + b^2$

*Note: a , c , e , a_M , a_S , c_M , c_S , b_M , β_S , r , and b are the path coefficients shown in Figure 2, and H is the additive genetic correlation: 1 for MZ and .5 for DZ twins.

The second stage is to examine the hypotheses of the study by comparing the fit of the various nested submodels (AE, CE, and E) with the full model. We use Akaike's Information Criterion (AIC) to compare the submodels and to choose the model which best explains the data.

RESULTS

Test-Retest Sample

The test-retest reliability of the response variables is assessed using the subsample of 33 family groups who were retested between 2 and 11 weeks later. The twin pairs for the retest sample had a median sex split. Data were not recorded for three fathers. The test-retest correlations are 0.601 (Twin 1) and 0.696 (Twin 2) for twin self-report, 1.00 for both Twin 1 and Twin 2 for the mother's report on the twin, 1.00 for the mother's self report; and 0.960 for mother's report on the father.

Univariate Analysis

Table II shows the frequency distributions of the six observed variables pooled across zygosity groups. The equality of distributions across zygosity groups can be tested using the PRELIS2.12e package by computing the differences in chi-square for all correlations between a model with a pooled estimate of thresholds versus a model with separate estimates of thresholds for each zygosity group [Jöreskog and Sörbom, 1993]. There are no significant differences for any zygosity group, which satisfies an assumption of the twin method that there is no difference in distributions between MZ and DZ twins and also indicates that there is no scalar sex limitation.

The modal response for five of the six variables is "Rarely." Very few in the sample responded "Frequently." The distributions for mother's report on father and mother's self-report are not markedly different, suggesting that the two are equivalent measures.

Correlations

Table III shows the cross-twin, within-trait correlations calculated by PRELIS2.12e for the six observed variables by zygosity. The DZ opposite-sex twins are ordered female-male. Three features are evident. First, the correlations for MZ twins are higher than for DZ twins—indicating a genetic component of variance in mouth ulcer liability. Second, the correlations for DZ opposite sex pairs are roughly the harmonic mean of the correlations for DZ same-sex pairs, which suggests that there is no

TABLE II. Distribution of the Six Observed Variables Pooled Over Zygosity (Sample size = 290 Family Groups)*

Category	Mother's report Twin 1 (%)	Self-report Twin 1 (%)	Mother's report Twin 2 (%)	Self-report Twin 2 (%)	Mother's report Father (%)	Self-report Mother (%)
Frequently	3	5	4	6	5	4
Sometimes	28	36	22	26	19	28
Rarely	37	34	42	42	52	51
Never	32	25	32	27	24	17

*Percentages may not add to 100 due to rounding.

TABLE III. Cross-Twin, Within-Trait Polychoric Correlations of the Six Observed Variables by Twin Zygosity (Standard Errors in Parentheses)

	Self-report	Mother's report
MZ female	0.626 (0.090)	0.797 (0.084)
MZ male	0.509 (0.134)	0.987 (0.009)
DZ female	0.212 (0.159)	0.316 (0.265)
DZ male	0.356 (0.160)	0.703 (0.130)
DZ opposite sex	0.288 (0.137)	0.638 (0.114)

nonscalar sex limitation. Third, the correlation for mother's report on the twins is much higher than the twins' self-report correlation, indicating rater bias.

Table IV shows the polychoric correlations of the six response variables by zygosity. The data are pooled over sex because there is no evidence for sex limitation. This has the advantage for the analysis of increasing the power by increasing the size of the analytic groups. The parent-offspring correlations are generally equivalent for both zygosity groups and for mother and father, which accords with the expectation that there should be no zygosity effect on the relationship between an individual and parent and the assumption that the measures for father and mother are equivalent.

Structural Equation Models

The polychoric correlation matrices and their associated asymptotic covariance matrices, generated using PRELIS2.12e, were then analysed using Mx 1.34 [Neale, 1995].

Model specification. The structural equation models represented in Figures 1 and 2 are not identified, that is, the number of parameters to be estimated exceeds the number of statistics with unique expectations [Jöreskog and Sörbom, 1989; Neale and Cardon, 1992]. In the case of the measurement model shown in Figure 1, there are 11 parameters to be estimated. The two 4×4 input matrices contain 20 statistics, but a number of the statistics have identical expectations and are not unique. The eight diagonal elements have the same expectation because they are correlations rather than covariances. Across zygosity groups, the two within-twin, cross-trait correlations have the same expectation, and the two between-twin, cross-trait correlations have the same expectation. There are only nine statistics in the data set that have

TABLE IV. Polychoric Correlation Matrices of Six Mouth Ulcer Ratings for MZ Twins (Lower Triangle) and DZ Twins (Upper Triangle)*

	Mother's report Twin 1	Self-report Twin 1	Mother's report Twin 2	Self-report Twin 2	Mother's report Father	Self-report Mother
DZ twins (163 pairs)						
Mother's report, Twin 1	—	0.427	0.582	0.314	0.369	0.562
Self-report, Twin 1	0.598	—	0.230	0.299	0.105	0.192
Mother's report, Twin 2	0.880	0.519	—	0.602	0.470	0.321
Self-report, Twin 2	0.422	0.563	0.493	—	0.291	0.296
Mother's report, Father	0.399	0.385	0.424	0.325	—	0.245
Self-report, Mother	0.387	0.265	0.347	0.021	0.193	—
MZ twins (127 pairs)						

*The twin intra-pair correlations for the two observed variables are shown in bold.

unique expectations. Similarly, with the combined model shown in Figure 2 there are 13 parameters but only 12 unique statistics. The 12 diagonal elements have the same expectation because they are correlations rather than covariances. Across zygosity groups, the two within-twin, cross-trait correlations have the same expectation; the two between-twin, cross-trait correlations have the same expectation; and the parental-offspring and between-parent correlations have the same expectation.

The identification problem can be solved by either fixing a parameter or equating parameters. In the measurement model the path coefficient e can be set at an arbitrary value since the latent phenotype has no natural scale. The model is still not identified, and the approach we choose is to estimate two models, including either the specific common environment components C_M and C_S or the specific genetic components A_M and A_S . There are now only eight parameters to be estimated and the model is identified. With the combined model, the path coefficient r can be determined by subtraction since the variance for the parental phenotypes is constrained to 1, i.e.: $a^2 + b^2 + r^2 = 1$. With this constraint there are now 12 parameters to be estimated and the model is identified.

The degrees of freedom for the two models reflect the fact that the diagonal elements—being correlations—are all unity. This is equivalent to constraining the variance of each of the four variables (twin self-report, mother's report on twin, self, and father) to unity, so the diagonal elements represent only two statistics for the measurement model and four statistics for the combined model. The number of statistics is thus 14 for the measurement model and 34 for the combined model.

Model fitting. Table V summarises the parameter estimates and model fit statistics for the two measurement models and the combined model. Model 1 is the measurement model including C_M and C_S , and model 5 is the measurement model including A_M and A_S . Models 2 to 4 and 6 to 8 are the nested submodels AE, CE, and E. Model 9 is the full combined model from Figure 2. Models 10 to 14 are the submodels relating the observed variables and reflect the development of the most parsimonious measurement and rater bias model to determine the twin's phenotype. Models 15, 16, and 17 test the hypotheses about the contribution of genetic and environmental factors to the variance of the latent phenotype of liability to mouth ulcers.

Both measurement models provide a good fit to the data (model 1 $\chi^2_6 = 6.632$, $P = 0.356$; model 5 $\chi^2_6 = 4.985$, $P = .546$). In the full ACE model additive genetic factors account for two-thirds of the phenotypic variance. The preferred model, using AIC, is an AE model (models 2 and 6), in which about 87% of the variance in the phenotype is attributable to additive genetic factors.

The full combined model (model 9) provides a reasonable fit to the data ($\chi^2_{22} = 32.63$, $P = .067$). The model comparisons for the combined model are summarised in Table VI. The first four lines represent the development of the measurement model. The additive genetic components of variance for twin self-report and mother's report on twin and the shared environment component for twin self-report can be set to zero without significant loss of fit. The shared environment component for mother's report on twin cannot be removed from the model without significantly reducing the fit. The final measurement model thus incorporates this component. There is significant rater bias, indicated by the significant worsening of fit when the bias parameter is removed. Model 12 is the most parsimonious measurement and rater bias model. The parameter estimates for model 12 are quite similar to those for models 1 and 5,

TABLE V. Summary of Parameter Estimates and Model Fit Statistics for the Measurement Model and the Combined Measurement and Rater Bias Model*

	a^a	c^a	e^a	a_M	a_S	c_M	c_S	e_M	e_S	b	r	β_M	β_S	Parameters	χ^2	df	P	AIC
Model 1	0.86	0.27	0.37	—	—	0.42	0.43	0.11	0.68	—	—	0.33	0.22	8	6.63	6	.356	-5.37
Model 2	0.94	—	0.35	—	—	0.47	0.41	0.16	0.67	—	—	0.30	0.22	7	7.01	7	.427	-6.99
Model 3	—	0.88	0.48	—	—	0.64	0.17	0.27	0.65	—	—	0.34	0.35	7	17.86	7	.013	3.86
Model 4	—	—	1.00	—	—	0.80	0.46	0.43	0.65	—	—	0.41	0.61	6	110.70	8	.000	94.70
Model 5	0.80	0.48	0.37	0.20	0.56	—	—	0.00	0.64	—	—	0.37	0.20	8	4.99	6	.546	-7.02
Model 6	0.93	—	0.36	0.49	0.48	—	—	0.10	0.63	—	—	0.31	0.22	7	6.49	7	.484	-7.51
Model 7	—	0.87	0.48	0.55	0.47	—	—	0.00	0.64	—	—	0.40	0.29	7	12.22	7	.094	-1.78
Model 8	—	—	1.00	0.85	0.55	—	—	0.31	0.63	—	—	0.42	0.55	6	94.58	8	.000	78.58
Model 9	0.80	0.51	0.32	0.26	0.00	0.42	0.00	0.24	0.59	0.46	0.39	0.70	0.81	12	32.63	22	.067	-11.37
Model 10	0.80	0.51	0.32	0.26	—	0.42	0.00	0.24	0.59	0.46	0.39	0.70	0.81	11	32.63	23	.088	-13.37
Model 11	0.80	0.51	0.32	0.26	—	0.42	—	0.24	0.59	0.46	0.39	0.70	0.81	10	32.63	24	.112	-15.37
Model 12	0.80	0.51	0.32	—	—	0.47	—	0.25	0.60	0.46	0.39	0.71	0.81	9	32.81	25	.136	-17.19
Model 13	0.76	0.57	0.30	—	—	—	—	0.25	0.67	0.48	0.36	0.80	0.71	8	49.01	26	.004	-2.99
Model 14	0.76	0.58	0.30	—	—	0.51	—	0.21	0.69	—	0.31	0.61	0.52	8	54.55	26	.001	2.55
Model 15	0.96	—	0.28	—	—	0.48	—	0.26	0.59	0.47	0.60	1.04	1.20	8	35.24	26	.107	-16.76
Model 16	—	0.82	0.57	—	—	0.48	—	0.31	0.67	0.56	0.83	0.61	0.76	8	111.07	26	.000	59.07
Model 17	—	—	1.00	—	—	0.60	—	0.37	0.83	0.58	0.82	0.56	0.75	7	20.34	27	.000	146.34

*Two measurement models are developed: models 1 to 4 include a_M and a_S ; models 5 to 8 include c_M and c_S . Models 9 to 14 develop the combined measurement and rater bias model; models 15 to 17 are the nested submodels (AE, CE, and E) of model 12. Preferred model is shown in bold.

^aThe estimates for a , c , and e are standardised so that $a^2 + c^2 + e^2 = 1$.

although these cannot be formally tested since models 1 and 5 are not nested submodels of model 12. Nevertheless, this similarity suggests that the inclusion of parental data and a rater bias variable have not altered the components of variance of the phenotype.

The final three lines of Table VI are the comparisons of the nested submodels AE, CE, and E against the ACE model (model 12). Using AIC, none of the submodels is preferred over the full model. Both the CE and E models represent a significant worsening of fit from the ACE model, and neither model fits the data as shown in Table V. The AE model is not a significantly worse fit than the ACE model and gives an adequate fit to the data, but has a greater AIC than the ACE model. Figure 3 shows the path diagram and parameter estimates for model 12.

In the ACE model, 64% of the variance in liability to mouth ulcers is accounted for by additive genetic factors, 26% by an environmental factor common to the twins, and the 10% is accounted for by environmental factors specific to the individual. Sixty-four percent of the variance in twin self-report and 49% of the variance in mother's report on the twin is due to the latent phenotype. An additional 22% of the variance in mother's report on the twin is contributed by a common environment latent variable. Twenty-one percent of the variance in the mother's ratings of herself or other family members is due to the latent bias variable.

DISCUSSION

We have found in this study that there is a sizeable genetic component of variation in the incidence of recurrent aphthous stomatitis, with a heritability of about 0.64, and this is consistent with previous twin research.

Item Reliability

The test-retest correlations for the mother's reports were effectively unity for all four measures which indicates either a high level of stability in how the mother rates the phenotype, or a high level of precision in the mother's ratings combined with a stable phenotype. The presence of a significant rater bias supports the former explanation. The

TABLE VI. Summary of Model Comparison and Hypothesis Testing Statistics for the Combined Model

Hypothesis	Model comparison	χ^2	df	<i>P</i>
a_S equals zero?	2 vs. 1	0.00	1	1.00
a_M equals zero?	3 vs. 2	0.00	1	1.00
c_S equals zero?	4 vs. 3	0.18	1	.67
c_M equals zero?	5 vs. 4	16.19	1	<.001
No rater bias	6 vs. 4	21.74	1	<.001
No shared environment (C) component—AE model	7 vs. 4	2.43	1	.119
No additive genetic (A) component—CE model	8 vs. 4	78.26	1	<.001
No shared environment or additive genetic component—E model	9 vs. 4	167.5	1	<.001

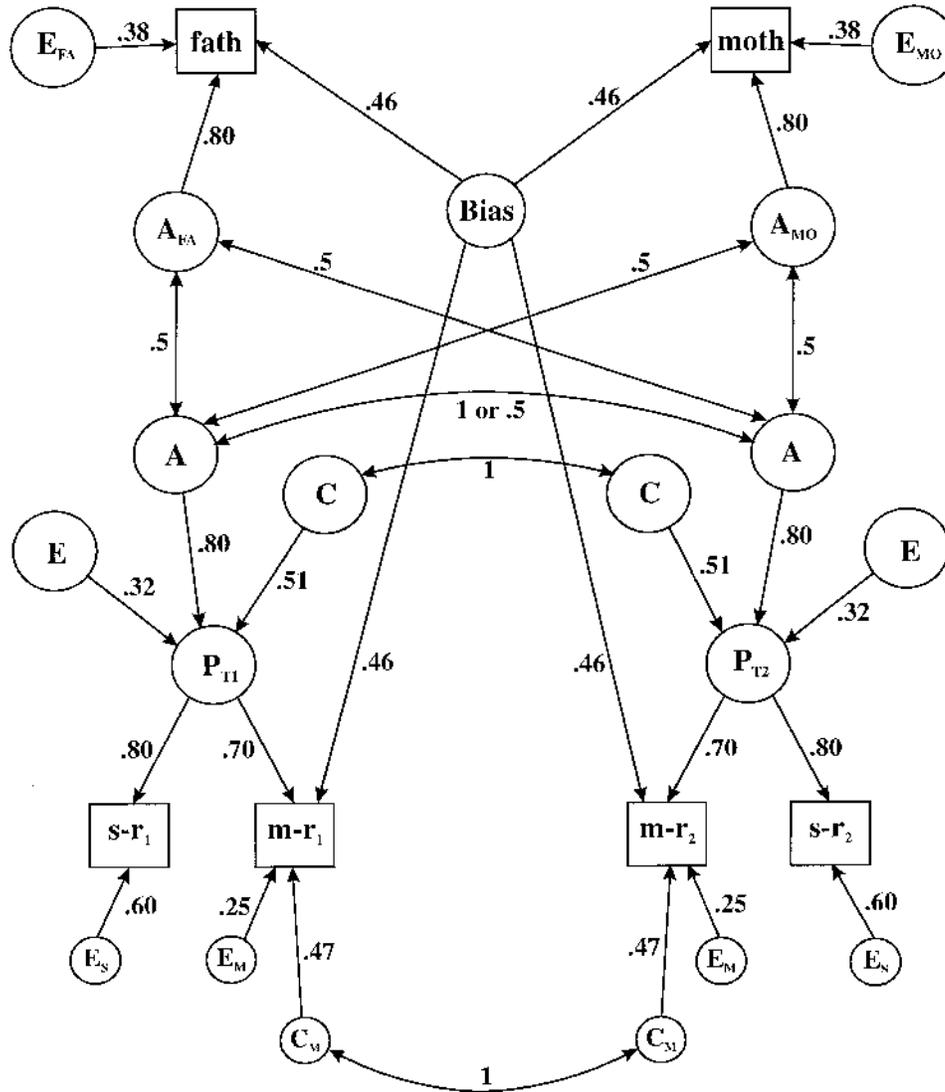


Fig. 3. Path diagram showing the parameter estimates for the combined measurement and rater bias model in which the variance in the latent phenotype P is a combination of additive genetic, common environment, and unique environment components.

phenotype may be more salient to the mother so that her rating of ulcer incidence is based on a long-term awareness of the phenotype. By contrast, the twin's self-reports have a much lower repeatability; this may be due to the phenotype being less salient to them, and hence their assessment of its cumulative incidence is less stable.

Ulcer Prevalence

Lifetime prevalence of mouth ulcers is high with 74% of the sample of 1,160 people reporting that they had had an ulcer. This is higher than the lifetime prevalence of 39%

for males and 50% for females reported in Embil et al. [1975]. That study used a criterion of two or more occurrences and was based on 10,532 cases. The disparity in lifetime prevalence between this study and Embil et al. may reflect both the differences in sample size and the relative imprecision of the measure used here. In a more selected population of young adults 66.2% reported a history of RAS [Ship et al., 1967].

Structural Equation Model

The best fitting model as judged by AIC explains the variance in the phenotype of liability to mouth ulcers by additive genetic factors, common environment factors, and by specific environmental factors. The model supports a significant genetic aetiology for mouth ulcers with 64% of the variance in the liability to mouth ulcers attributed to genetic factors. This is in accord with previous research on recurrent mouth ulcers. The choice of the ACE model as the best fitting model is a marginal decision: the AIC for the AE model (model 16) is only slightly greater (by 0.43), and the AE model is more parsimonious. In the AE model, 92% of the variance is due to additive genetic factors and 8% is due to specific environmental factors. It is possible that, with a larger sample size, the AE model may have been chosen, giving strong support to a primarily genetic aetiology for RAS.

The value of using a measurement model to explain covariation in phenotypes is highlighted in this study. A large proportion of the variance in the observed variables was not accounted for by the latent phenotype. The measurement model partitioned measurement error in the observed variables from specific environment effects in the latent phenotype. A genetic-environmental analysis on only the observed variables would have given an inflated estimate of the effect of specific environment. This would have led to underestimates of the effect of shared environment and of the heritability of mouth ulcers. The utility of a rater bias model is also shown. In the presence of significant rater bias, a failure to account for this bias in the model will inflate the estimate of the contribution of shared environment because assessment by the one rater is structurally equivalent to shared environment. With the data in this study 21% of the variance in the mother's ratings was due to her bias. The significant shared environment effect for the mother's report on each twin is likely to be due to her association bias in rating them similarly due to their twin or sibling status. Thus an additional 22% of the variance in mother's report may represent rater bias, giving a sum of 43% of the variance possibly due to bias. Rater bias can be controlled to an extent by using a quantitative scale rather than, as in this study, a qualitative scale, and also by external validation of the ratings.

Limitations

One limitation to this study is the data quality. The data are categorical with only four categories. Using categorical rather than continuous data for a quantitative phenotype limits the variance and thereby reduces the precision of any estimate. Similarly, the inconsistent application of the protocol identifying mouth ulcers may have led to misclassification problems with a corresponding loss of precision. Given the significant genetic nature of liability to mouth ulcers a logical next step in the description of the aetiology of mouth ulcers would be the search for quantitative trait loci which account for the genetic variance. This search would be better informed with reliable, unbiased continuous data on the phenotype.

Another limitation in the study is whether the results can be extended to the general population. The participants were volunteer twins aged about 12 years, and participation in the study may be related in some manner to the criterion variable. This is unlikely given that the primary purpose of the study was investigation of melanocytic naevi. Alternatively, young adolescent twins may be different from the remainder of the population in either the incidence or the aetiology of mouth ulcers. The only guide to this is the incidence in the sample versus other studies and, as shown earlier, the data are not of sufficient quality to be conclusive.

CONCLUSIONS

In this study we have added to our knowledge on the aetiology of RAS. We fitted a genetically informative structural equation model to data from family groups and have shown that there is a significant genetic component in the variation in the liability to mouth ulcers in twins aged 12 years. This indicates a strong genetic factor in the aetiology of RAS in a nonclinical population. However, the 36% of variance not accounted for by genetic factors suggests that nongenetic factors have a role to play in RAS.

We have extended the previous research on the incidence of mouth ulcers using twins [Miller et al., 1977] by using data on 1,160 individuals—representing 290 twin pairs and their parents—with a commensurate increase in power and the precision of the estimates of heritability.

The estimate of heritability of mouth ulcers, and the search for quantitative trait loci, will be enhanced with an objective, quantitative scale which is less subject to the rater bias evident in this study.

ACKNOWLEDGMENTS

We thank Ann Eldridge and Marlene Grace for the data collection, John Pearson and Theresa Pangan for the data storage and extraction, and John Hewitt for comments on the model. We also thank two anonymous reviewers for their suggestions. The data collection was funded by grants from the Queensland Cancer Fund and the National Health and Medical Research Council. The airfare to Australia of one author (S.J.T.) was paid for by the Cancer Research Campaign.

REFERENCES

- Axell T, Henricsson V (1985): The occurrence of recurrent aphthous ulcers in an adult Swedish population. *Acta Odontol Scand* 43:121–125.
- Bollen KA (1989): "Structural Equations With Latent Variables." New York: John Wiley & Sons.
- Boomsma DI, Martin NG, Neale MC (eds) (1989): Genetic analysis of twin and family data: structural modeling using LISREL. Special edition of *Behavior Genetics* ed.
- Eaves LJ, Martin NG, Heath AC (1990): Religious affiliation in twins and their parents: Testing a model of cultural inheritance. *Behav Genet* 20:1–22.
- Embil JA, Stephens RG, Manuel FR (1975): Prevalence of recurrent herpes labialis and aphthous ulcers among young adults on six continents. *Can Med Assoc J* 113:627–630.
- Hewitt JK, Silberg J, Neale MC, Eaves LJ, Erickson M (1992): The analysis of parental ratings of children's behavior using LISREL. *Behav Genet* 22:293–317.
- Jöreskog KG, Sörbom D (1988): "PRELIS A Program for Multivariate Data Screening and Data Summarization. A preprocessor for LISREL," 2nd ed. Mooresville, IN: Scientific Software Inc.

- Jöreskog KG, Sörbom D (1989): "LISREL 7: A Guide to the Program and Applications," 2nd ed. Chicago: SPSS Inc.
- Jöreskog KG, Sörbom D (1993): "New features in PRELIS2." Chicago, IL: Scientific Software.
- Liang GS, Daikos GL, Serfling U, Zhu WY, Pecoraro M, Leonardi CL, Fischl MA, Penneys NS (1993): An evaluation of oral ulcers in patients with AIDS and AIDS-related complex. *J Am Acad Dermatol* 29:563–568.
- Martin NG, Eaves LJ, Kearsley MJ, Davies P (1978): The power of the classical twin study. *Heredity* 40:97–116.
- Martin NG, Eaves LJ, Heath AC, Jardine R, Feingold LM, Eysenck HJ (1986): Transmission of social attitudes. *Proc Natl Acad Sci USA* 83:4364–4368.
- Miller MF, Garfunkel AA, Ram C, Ship II (1977): Inheritance patterns in recurrent aphthous ulcers: twin and pedigree data. *Oral Surg Oral Med Oral Pathol* 43:886–891.
- Neale MC (1995): "Mx Statistical Modelling," 3rd ed. Richmond, VA: Department of Psychiatry, Medical College of Virginia.
- Neale MC, Cardon LR (eds.) (1992): "Methodology for Genetic Studies of Twins and Families." Dordrecht: Kluwer Academic Publishers.
- Olsson U (1979): Maximum likelihood estimation of the polychoric correlation coefficient. *Psychometrika* 44:443–460.
- Ozbakir F, Yazici H, Mat C, Tuzun Y, Yurdakul S, Yilmazer S (1987): HLA antigens in recurrent oral ulceration: evidence against a common disease spectrum with Behcet's syndrome. *Clin Exp Rheumatol* 5:263–265.
- Pedersen A (1989): Psychological stress and recurrent aphthous ulceration. *J Oral Pathol Med* 18:119–122.
- Pedersen A (1991): Are recurrent oral aphthous ulcers of viral etiology? *Med Hypotheses* 36:206–210.
- Porter SR, Scully C (1991): Aphthous stomatitis—an overview of aetiopathogenesis and management. *Clin Exp Dermatol* 16:235–243.
- Ship II, Brightman VJ, Laster LL (1967): The patient with recurrent aphthous ulcerations and the patient with recurrent herpes labialis: a study of two population samples. *J Am Dent Assoc* 74:645–653.
- Wray D, Feruson MM, Mason DK, Hutcheon AW, Dagg JH (1975): Recurrent aphthae: treatment with vitamin B12, folic acid and iron. *Br Med J* 2:490–493.