

Estimating Age-specific Prevalence of Testosterone Deficiency in Men Using Normal Mixture Models

Yungtai Lo

Mount Sinai School of Medicine

Abstract: Testosterone levels decline as men age. There is little consensus on what testosterone levels are normal for aging men. In this paper, we estimate age-specific prevalence of testosterone deficiency in men using normal mixture models when no generally agreed on cut-off value for defining testosterone deficiency is available. The Box-Cox power transformation is used to determine which transformation is most appropriate for correcting skewness in data and best suits normal mixture distributions. Parametric bootstrap tests are used to determine the number of components in a normal mixture.

Key words: Box-Cox power transformation, EM algorithm, normal mixture, parametric bootstrap, testosterone level.

1. Introduction

Testosterone is a male sex hormone that helps maintain bone mass, fat distribution, male hair patterns, muscle mass and strength, sex drive and sperm production throughout male adult life. Testosterone levels decline as men age. Harman *et al.*(2001) reported that most of change in male testosterone levels occurs before age 50. Some reported effects of testosterone deficiency in men include decreased energy, reduced muscle mass and strength, decreased cognitive function, less sexual interest or potency, increased breast size, and a depressed mood.

What are normal testosterone levels for healthy older men? There is no general agreement on what testosterone levels are normal for healthy aging men. A recent endocrine society annual andropause consensus meeting suggested that 300 nanograms per deciliter (ng/dl) is the lower limit for normal testosterone levels, and total testosterone levels less than 200 ng/dl clearly indicate hypogonadism in healthy young men. Harman *et al.*(2001) reported age-specific prevalences of hypogonadism of 12, 19, 28, and 49 percent for men in their 50s, 60s 70s, and 80s, respectively. They defined hypogonadism as total testosterone levels < 325

ng/dl. Mohr *et al.*(2005) used data from men aged 40 to 79 years to aid in the diagnosis of testosterone deficiency. They defined total testosterone levels less than the age-specific 2.5th percentile as abnormally low testosterone levels; thus about 2.5 % of men would have abnormally low testosterone in their male aging study. The corresponding age-specific cut-off values are 251, 216, 196, and 156 ng/dl for men in their 40s, 50s, 60s, and 70s, respectively.

The prevalence of testosterone deficiency is conventionally estimated using a pre-specified cut-off value to classify each serum total testosterone. When there is a gold standard for defining or determining testosterone deficiency, the conventional approach will in general give a good estimate of the testosterone deficiency prevalence. When there is no gold standard, prevalence estimates obtained using the conventional approach to define deficiency status will often vary with the definition used. Establishing cut-offs that are generally agreed is difficult or impossible when there is no clear separation between normal testosterone levels and abnormal testosterone levels. As testosterone levels decline gradually with age in men, a subgroup of men will have testosterone deficiency and have signs and symptoms related to hypogonadism whereas a subgroup will still have testosterone levels within the normal range throughout their lifetimes causing no significant problems at all. If we assume that serum total testosterone samples are taken from a mixture of men who have testosterone deficiency and those who do not have testosterone deficiency, the prevalence of testosterone deficiency can be estimated through fitting a two-component mixture model to the total testosterone levels. Unlike the conventional approach for estimating prevalence, mixture models do not need a pre-specified cut-off value to classify each serum total testosterone. Mixture models have been used to estimate the disease prevalence in other studies in the absence of a gold standard measure for a definitive diagnosis of disease status; see, for example, Gay (1996), Pfeiffer, Gail and Brown (2000), and Pfeiffer *et al.*(2008).

In this paper, we use mixture models to account for individual testosterone levels that can arise either from a subpopulation with testosterone deficiency or from a subpopulation without testosterone deficiency, and to estimate the age-specific rate of testosterone deficiency in men in a particular age group when no generally agreed on cut-off value for that age group is available. Mixture models are often used to model data that arise from heterogeneous populations where the subpopulation to which an individual observation belongs is unknown. The component distributions can arise either from the same or different parametric families. For examples of applications of mixture models, see Everitt and Hand (1981), Titterington, Smith and Makov (1985), McLachlan and Basford (1988), and McLachlan and Peel (2000). The Box-Cox family of power transformation is used to transform data to a mixture of normal distributions (Box

and Cox, 1964). The maximum likelihood estimators of the unknown parameters are obtained through application of the EM algorithm and a grid search over a range of possible power transformations (Dempster, Laird and Rubin, 1977). We illustrate the use of the Box-Cox power transformation in mixture models to estimate age-specific prevalence of testosterone deficiency with a sample of serum total testosterone levels obtained from an HIV at-risk aging men's prospective study. Data and model description are given in Section 2. Parametric bootstrap tests for determining the number of components in a normal mixture distribution are given in Section 3. In Section 4, we adjust effects of covariates on total testosterone levels. Discussions and concluding remarks are presented in Section 5.

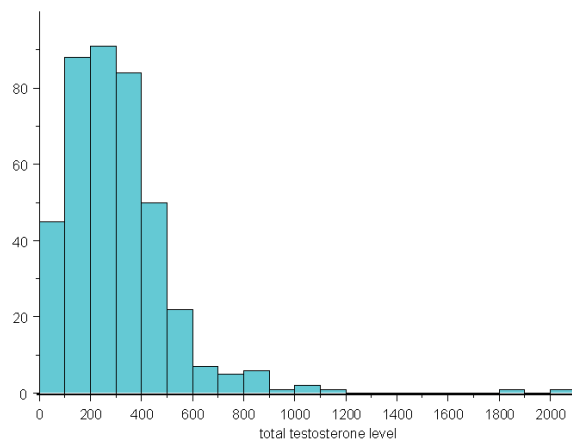


Figure 1: Histogram of total testosterone levels from 404 men in their 50s

2. Data and Model Description

Serum total testosterone levels were measured in duplicate using time-resolved immunofluorometric assays (DELFLIA; Pharmacia) in 404 men who were 50 to 59 years of age and currently not taking androgens from a cohort of HIV at-risk aging men's prospective study; details on the study design are given in Klein *et al.*(2005). We wish to estimate the age-specific rate of testosterone deficiency in men aged 50 to 59 with or at risk for HIV infection using a two-component normal mixture model when no generally agreed on cut-off for testosterone deficiency exists. The mean (\pm standard error) total testosterone level was 304.94 ± 10.81 ng/dl with a median (range) value of 273.5 (2 - 2070) ng/dl. Of these 404 men, 180

(44.6%) had total testosterone levels greater than 300 ng/dl, 91(22.5%) between 200 and 300 ng/dl, and 133 (32.9%) less than 200 ng/dl. The histogram of total testosterone levels is given in Figure 1.

From inspection of Figure 1, it appears that the distribution of total testosterone levels is skewed to the right and does not seem to suggest a bimodal distribution that may support using a mixture of two normal distributions to fit the data and estimate proportion of men with testosterone deficiency. It is noted that skewness is an inherent part of mixture distributions, in particular, when the component distributions are not well separated. If evidence for a mixture of two normal distributions is confounded by skewness, an appropriate transformation on total testosterone levels is often sufficient to make the normal mixture distribution appropriate for the transformed data. As conventional transformations such as logarithms or square roots are not sufficient to provide an appropriate adjustment for skewness in the data, the Box-Cox power transformation is used to determine which transformation is most appropriate for correcting skewness and best suits normal mixture distributions. Mixture models of transformed normal components by the Box-Cox power transformation have been used in other studies (MacLean, Morton and Elston, 1976; Schork and Schork, 1988; Gray, 1994; Gutierrez *et al.*, 1995).

Let Y_i denote the total testosterone level obtained from subject i and $n = 404$ be the number of men in this study. Assume that there exists a real number λ such that the Box-Cox power transformation on Y_i denoted as $Y_i^{(\lambda)}$ is distributed as a two-component normal mixture distribution, the probability density function of Y_i is given as follows

$$f(y_i; \theta) = \{\pi f_1(y_i^{(\lambda)}; \mu_1, \sigma_1^2) + (1 - \pi) f_2(y_i^{(\lambda)}; \mu_2, \sigma_2^2)\} y_i^{\lambda-1}, \quad (2.1)$$

where

$$y_i^{(\lambda)} = \begin{cases} (y_i^\lambda - 1)/\lambda & \text{if } \lambda \neq 0, \\ \log y_i & \text{if } \lambda = 0, \end{cases}$$

λ is the unknown Box-Cox power transformation parameter, $\theta = (\pi, \mu_1, \mu_2, \sigma_1^2, \sigma_2^2)$ is a vector of unknown parameters, $f_i(\cdot)$ is the normal density function with mean μ_i and variance σ_i^2 , $\pi > 0$ is the mixing proportion, and the last term on the right-hand side of (2.1) corresponds to the Jacobian of the transformation from y_i to $y_i^{(\lambda)}$. To ensure identifiability of θ for a given value of λ , we assume without loss of generality that $\mu_1 < \mu_2$. Kiefer and Wolfowitz (1956) noted that the likelihood function of a normal mixture with unequal variances is unbounded above. Day (1969) argued that spurious maximizers may occur if some component distribution has a very small variance relative to others when the corresponding cluster contains few data points sufficient close together. Hathaway (1985) suggested

using the constraint that $\min_{i,j}(\sigma_i/\sigma_j) \geq c > 0$, $c \in (0, 1]$ to rule out spurious maximizers and showed that there exists a constrained global maximizer of the likelihood function on the constrained parameter space. To cope with unboundedness of the likelihood function and avoid spurious maximizers, we further assume that $\min(\sigma_1/\sigma_2, \sigma_2/\sigma_1) \geq c > 0$, for some c in the unit interval. The mixing proportion π represents that a subgroup π of these men have testosterone deficiency whereas a subgroup $1 - \pi$ of these men do not have testosterone deficiency.

The log-likelihood function of y_1, \dots, y_n is given by

$$L(\lambda, \theta; \mathbf{y}) = \sum_{i=1}^n [\log\{\pi f_1(y_i^{(\lambda)}; \mu_1, \sigma_1^2) + (1 - \pi)f_2(y_i^{(\lambda)}; \mu_2, \sigma_2^2)\} + (\lambda - 1) \log y_i]. \quad (2.2)$$

For a fixed λ , the last term on the right-hand side of (2.2), $\sum(\lambda - 1) \log y_i$, is just a constant and maximization of L can be achieved by maximizing the log-likelihood function of the transformed data, $y_1^{(\lambda)}, \dots, y_n^{(\lambda)}$ given in the first two terms on the right-hand side of (2.2) through application of the EM algorithm with some modification subject to the constraint $\min(\sigma_1/\sigma_2, \sigma_2/\sigma_1) \geq c$. Thus, the maximizer of L can be found through use of the EM algorithm and a grid search over a range of possible λ values. It can be shown that the maximum likelihood estimator $\hat{\lambda}$ is that value of λ that maximizes (2.2). Alternatively, the maximum likelihood estimator for λ can be found using Newton's method in the M step of the EM algorithm to maximize L .

A grid of $\lambda \in [-2, 2]$ were used in our attempt to determine which transformation is most appropriate for the normal mixture distribution. Since the likelihood of mixture distributions often has multiple modes, there is no guarantee that one or few starting values will find the correct root. Thode, Finch and Mendell (1988) suggested using several starting values to search for all possible modes. It is noted that problems associated with the likelihood function on the unconstrained parameter space can occur when c is small. How small can c be to ensure that the constrained parameter space contains the true values of the parameters. Lo (2005) reported based on a simulation study that $c = 0.25$ appears to work well for normal mixture distributions with unequal variances. In this study, we set the lower bound for the constraint on the component variances $c = 0.25$. For each λ value, one hundred sets of starting values for θ were used in our search for the global rather than local maximizer through application of the EM algorithm subject to the constraint that $\min(\sigma_1/\sigma_2, \sigma_2/\sigma_1) \geq 0.25$. For the selection of starting values, See, for example, Thode, Finch and Mendell (1988).

Fitting model (2.1) to our $n = 404$ measurements on total testosterone levels yielded the maximum likelihood estimates (\pm standard error) $\hat{\lambda} = 0.3$, $\hat{\pi} = 0.185 \pm 0.09$, $\hat{\mu}_1 = 13.19 \pm 0.97$, $\hat{\mu}_2 = 14.57 \pm 0.23$, $\hat{\sigma}_1 = 4.06 \pm 0.69$ and $\hat{\sigma}_2 =$

3.00 ± 0.24 . The approximate 95% confidence interval for the Box-Cox power transformation parameter is given by $0.05 \leq \lambda \leq 0.65$. The estimated mixing proportion $\hat{\pi} = 0.185$ gave an estimated 18.5% of men aged 50 to 59 with or at risk of acquiring HIV having testosterone deficiency, and an estimated 81.5% of these men with normal testosterone levels. Mean total testosterone levels were 207.67 ng/dl in men from the first component distribution and 271.32 ng/dl in men from the second component distribution after transforming Y_i back to its original scale, $Y_i = (1 + \lambda Y_i^{(\lambda)})^{1/\lambda}$. The histogram and density curve of $Y^{(\hat{\lambda}=0.3)}$ are given in Figure 2. The figure shows that the two-component normal mixture distribution fits the data well.

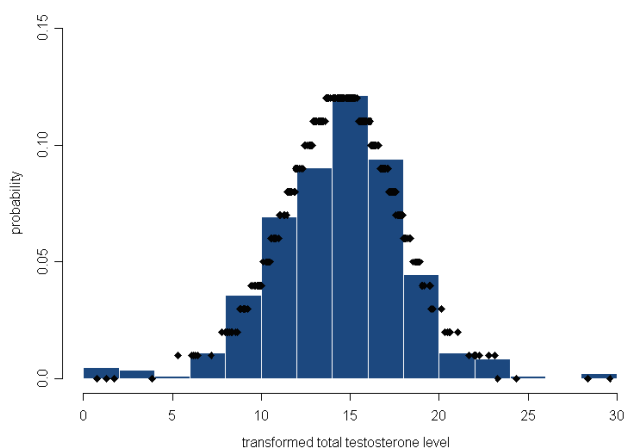


Figure 2: Histogram of $Y^{\hat{\lambda}=0.3}$ with a fitted density for the two-component normal mixture distribution

3. Testing for the Number of Components

The Box-Cox power transformation may remove skewness from the data such that a single normal distribution may be appropriate for the data. On the other hand, there may exist a Box-Cox power transformation such that a mixture of three normal distributions (for example, among men in the low testosterone group, some have very low testosterone levels) may give a better fit to the data. We wish to test whether a mixture of two transformed normal distributions fits the data significantly better than a single transformed normal distribution or to test whether a mixture of three transformed normal distributions fits the data better than a mixture of two transformed normal distributions. Note that the value of the Box-Cox power transformation λ for the null model is, in general,

different from that for the alternative model. As the null model is, in general, not nested within the alternative model, the likelihood ratio statistic for testing the number of components does not have the classic χ^2 reference distribution. Even if both null and alternative models have the same value of λ , the asymptotic χ^2 theory does not hold as under the null hypothesis, the mixing proportion is on the boundary of the parameter space and the parameters are not identifiable under the null model. For a thorough account of breakdown in regularity conditions under which classic asymptotic theory holds for the likelihood ratio test, refer to Titterton, Smith and Makov (1985) and McLachlan and Basford (1988).

Several studies have been conducted to investigate the asymptotic distribution of the likelihood ratio statistic for testing the number of components in a mixture model, see, for example, Dacunha-Castelle and Gassiat (1999), Lemdani and Pons (1999), Chen, Chen and Kalbfleisch (2001), and Lo, Mendell and Rubin (2001). These approaches can be used to determine the number of components when both null and alternative models have the same power transformation. When power transformation under the null model is different from that under the alternative model, these approaches may not be suitable for determining the number of mixture components without appropriate modification to account for Box-Cox power transformation parameters. Parametric bootstrap tests are hence used to test whether total testosterone levels after a Box-Cox power transformation have arisen from a single normal distribution, a mixture of two normal distributions, or a mixture of three normal distributions. The likelihood ratio statistic is defined as

$$LR = -2(L_0 - L_1)$$

where L_0 and L_1 are the log likelihood functions maximized under the null and alternative hypotheses, respectively. Inferences are drawn based on the observed bootstrap p-value that is defined as the proportion of replicates of LR that are as extreme as or more extreme than the value of LR obtained from the observed data. The null hypothesis will be rejected if the observed bootstrap p-value is less than or equal to a pre-specified significance level. For each bootstrap sample, the maximum likelihood estimates of the parameters under the null and alternative models are obtained in turn through a grid search of $\lambda \in [-1, 1]$ and the EM algorithm, using 100 sets of starting values, subject to the constraint on the component variances.

We first tested the null hypothesis that data have arisen from a single normal distribution after a Box-Cox power transformation against the alternative hypothesis that the data have arisen from a mixture of two normal distributions after another Box-Cox power transformation. One thousand bootstrap samples of size $n = 404$ were generated from a random variable Y such that

$$Y^{(\lambda)} \sim N(\mu, \sigma^2)$$

where λ , μ , and σ were estimated by the maximum likelihood estimates $\hat{\lambda} = 0.4$, $\hat{\mu} = 20.815$, and $\hat{\sigma} = 6.548$ obtained from the observed data. Among 1000 replicates of LR , there was none in excess of 19.24, the observed value of LR , so the observed bootstrap p-value was 0.0. Thus, there is strong evidence that there are at least two subpopulations. The histogram of 1000 bootstrap replicates of the likelihood ratio statistic is given in Figure 3(a).

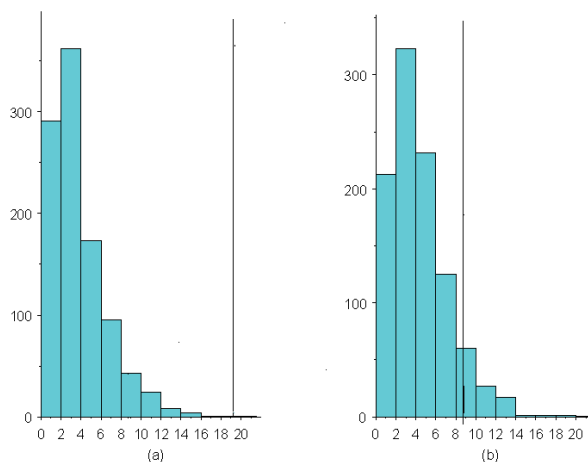


Figure 3: Histograms of 1000 bootstrap replicates of LR for testing (a) a single normal versus a two-component normal mixture and (b) a two-component normal mixture versus a three-component normal mixture. The vertical line is the observed value of LR .

We then tested the null hypothesis that data have been drawn from a mixture of two normal distributions after a Box-Cox power transformation against the alternative hypothesis that the data have been drawn from a mixture of three normal distributions after another Box-Cox power transformation. One thousand bootstrap samples of size $n = 404$ were generated from a random variable Y such that

$$Y^{(\lambda)} \sim \pi N(\mu_1, \sigma_1^2) + (1 - \pi)N(\mu_2, \sigma_2^2)$$

where λ , π , μ_1 , μ_2 , σ_1 , and σ_2 were estimated by the maximum likelihood estimates $\hat{\lambda} = 0.3$, $\hat{\pi} = 0.185$, $\hat{\mu}_1 = 13.19$, $\hat{\mu}_2 = 14.57$, $\hat{\sigma}_1 = 4.06$, and $\hat{\sigma}_2 = 3.00$ obtained from the observed data. Among 1000 replicates of the test statistic LR , there were 80 in excess of 8.72, the observed value of LR , so the observed bootstrap p-value was 0.08. Thus, there is no significant evidence that more than two subpopulations are required. The histogram of 1000 bootstrap replicates of the likelihood ratio statistic is given in Figure 3(b).

4. Adjustment for Covariates

Of 404 men aged 50 to 59, 231 were black, 224 had human immunodeficiency virus (HIV) infection, 295 had hepatitis C virus (HCV) infection, 168 had both HIV infection and HCV infection, 48 were injection drug users, 291 smoked cigarettes, and 128 used psychotropic medications. The median (range) body mass index (BMI) was 26.3 (16.7 - 56.7) kg/m^2 ; 157 men were overweight with a BMI between 25 and 30 kg/m^2 , and 87 men were obese with a BMI greater than 30 kg/m^2 . To adjust effects of these factors on total testosterone levels, the mixture model given in (2.1) can be easily extended to a mixture of two normal linear regression models by letting

$$\begin{aligned} Y^{(\lambda)} &= \mathbf{x}^T \boldsymbol{\beta}_1 + \epsilon_1 && \text{with probability } \pi, \\ Y^{(\lambda)} &= \mathbf{x}^T \boldsymbol{\beta}_2 + \epsilon_2 && \text{with probability } 1 - \pi \end{aligned}$$

where \mathbf{x} is a vector of predictors, $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ are vectors of unknown parameters, and the error terms $\epsilon_k \sim N(0, \sigma_k^2)$ for $k=1$ and 2. The log-likelihood function of y_1, \dots, y_n is then given as follows:

$$L(\lambda, \theta; \mathbf{y}) = \sum_{i=1}^n [\log\{\pi f_1(y_i^{(\lambda)}; \mu_{1i}, \sigma_1^2) + (1 - \pi) f_2(y_i^{(\lambda)}; \mu_{2i}, \sigma_2^2)\} + (\lambda - 1) \log y_i], \quad (4.1)$$

where $\mu_{1i} = \mathbf{x}^T \boldsymbol{\beta}_1$ and $\mu_{2i} = \mathbf{x}^T \boldsymbol{\beta}_2$. Similarly, for a fixed value of λ , the maximum likelihood estimators for π , $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$, σ_1^2 , and σ_2^2 can be obtained by maximizing the log-likelihood function in (4.1) through application of the EM algorithm subject to the constraint on the component variances. The covariance matrix of the model parameters can be estimated by the method proposed by Louis (1982). Details on derivation of the covariance matrix can be found in Thompson, Smith and Boyle (1998) and Turner (2000) when using the EM algorithm to find the maximum likelihood estimators.

We began our modeling with a model including all the covariates and an interaction between HIV infection and HCV infection, and an interaction between HIV infection and injection drug use in the first component distribution and in the second component distribution. A stepwise approach along with the Bayesian information criterion (BIC) and likelihood ratio tests was used for model selection. The parameter estimates and their corresponding standard errors for the final model are given in Table 1. The estimated mixing proportion $\hat{\pi} = 0.29$ was different from that in the model without adjustment for covariates given in (2.1). The regression model for μ_{1i} included terms for black, BMI, injection drug use, and use of psychotropic medications. This suggests that among men in the first

component population, African American men were likely to have higher total testosterone levels; increased BMI, injection drug use, and use of psychotropic medications were associated with lower total testosterone levels. The regression model for μ_{2i} included terms for BMI, cigarette smoking, HCV infection, HIV infection, and an interaction between HCV infection and HIV infection. Among men in the second component population, increased BMI and smoking cigarettes were associated with decreased total testosterone levels. Both HCV infection and HIV infection were associated with lower total testosterone levels. However, there was a significant interaction between HCV and HIV; the effect of HCV infection on testosterone was modified by HIV infection. Age was not significantly associated with total testosterone levels. Compared to the mixture regression model, the estimated means of two components in (1) appear to be very close when the covariates were not included.

Table 1: Parameter estimates and standard errors for the mixture of two normal linear regression models fitted to the testosterone data with $\hat{\lambda} = 0.3$

Variable	Estimate	Estimated standard error	Z-value	p-value
π	0.29	0.06	4.56	
Intercept	13.76	0.75	18.45	
Black	2.34	0.73	3.21	0.001
BMI*	-0.51	0.25	-2.07	0.039
Injection drug use	-5.81	0.91	-6.40	< 0.001
Use of psychotropic medications	-2.11	0.69	-3.06	0.002
σ_1	1.73	0.21	8.19	
Intercept	18.36	0.78	23.44	
BMI*	-0.86	0.26	-3.34	< 0.001
Smoking	-1.26	0.56	-2.25	0.025
HCV	-3.95	0.78	-5.05	< 0.001
HIV	-2.23	0.96	-2.32	0.021
HCV x HIV	2.73	1.08	2.52	0.012
σ_2	3.74	0.19	19.92	

5. Discussion

We have demonstrated the use of a mixture distribution with two normal components for estimating age-specific prevalence of testosterone deficiency in men and the use of a mixture of two normal linear regression models for adjusting effects of covariates on testosterone levels. The age-specific prevalence of testosterone deficiency in men aged 50 to 59 with or at risk for HIV infection was estimated to be about 29% after adjusting for covariates. This estimated prevalence rate is lower than 61.9% estimated by using the cut-off point 325

ng/dl suggested by Harman *et al.* (2001), 55.4% using the cut-off point 300 ng/dl suggested by the recent endocrine society annual andropause consensus meeting, or 37.6% using the cut-off point 216 ng/dl suggested by Mohr *et al.* (2005) for men in their 50s as the lower limit of the normal range for testosterone levels in these HIV at risk men. Due to study design, only the age-specific prevalence in men aged 50 to 59 can be estimated in this study. However, the mixture model can be used to estimate age-specific prevalence in men, for example, in their 60s and 70s, respectively when there are a sufficient number of men in these two age groups.

In this study, we found that injection drug use and use of psychotropic medications were associated with decreased total testosterone levels. Low testosterone levels have been reported in opiate addicts (Cicero, 1980). Opiates may act by suppressing the hypothalamic-pituitary-gonadal axis, a reproductive hormonal axis in men that consists of hypothalamus, pituitary gland, and testis (Cicero, 1980). Use of psychotropic medications may lead to sexual dysfunction, including testosterone deficiency resulting from hyperprolactinemia (Dobs *et al.*, 1988). Early in the HIV/AIDS epidemic, hypogonadism was common in men with AIDS (Ferri, Bertozzi and Zignego, 2002). Low testosterone levels in men with HIV infection have been correlated with advanced immunodeficiency. We observed an association between lower testosterone levels and HIV infection in HCV uninfected men. Hepatitis C virus infection had the strongest association with lower total testosterone levels in the second component population. This may suggest that severe liver disease is associated with low testosterone levels. Ferri *et al.* (2002) reported that plasma levels of total and free testosterone were generally lower in 207 patients with hepatitis C than in 2010 Italian men previously evaluated for erectile dysfunction. Increased body mass was associated with lower total testosterone levels in this study. Obesity associated with low total testosterone levels has been reported by Gapstur *et al.* (2002) in their study of associations of age, obesity, and race with serum androgen concentrations in young men. Glass *et al.* (1977) reported that serum total testosterone levels decrease with increasing body mass because of decreasing sex hormone-binding globulin. Higher testosterone levels in African American men than in white men have been reported by Ellis and Nyborg (1992) and Winters *et al.* (2001), but participants were generally younger than the sample of men studies here. We found smoking cigarettes was significantly associated with decreased testosterone levels. A literature review on cigarette smoking and male reproduction found reports variously suggesting that testosterone may be unchanged, or significantly elevated or decreased in smokers (Vine, 1996). We did not find age significantly associated with testosterone levels. Perhaps, this is because most of change in male testosterone levels occurs before age 50 (Harman *et al.*, 2001). At only ten

years, the range is not great enough to provide power to detect associations with age. A further investigation of these seems warranted.

After estimating the unknown parameters and testing for the number of components, mixture models can be used to cluster individuals into one of two sub-populations. The classification can be achieved by making use of posterior probabilities of population membership of each observation as calculated in the E step of the EM algorithm. Individual observations will be assigned to the component distribution to which it has the highest estimated posterior probability of belonging. As these estimated posterior probabilities may have limited reliability in small samples, they may not give a satisfactory assignment of the data. When sample sizes are large, mixture models can also be used to aid in determination of cut-off points for identifying men with testosterone deficiency. A cut-off point can be chosen to attain, for example, a specificity rate in the range of 90% to 99% to limit the number of false positive results as treatment of older HIV and HCV infected men with testosterone may be associated with greater risk than treatment of younger patients. Both sensitivity and specificity are calculated based on the selected mixture model under the assumption that the model correctly identifies two distinct populations and the component population with lower testosterone levels represents men with testosterone deficiency; see Thompson, Smith, and Boyle (1998) for details and related references. For example, among injection drug and psychotropic medication users, respective cut-off points to have a specificity of 90% for detecting testosterone deficiency, are 104 ng/dl for black non-smoking men aged 50 to 59 with HIV infection and HCV infection, and 223 ng/dl for black non-smoking men aged 50 to 59 without HIV infection and HCV infection. The corresponding sensitivity rates are 87% and 99%, respectively. An average value was used for BMI.

We have used mixture models assuming that the mixing proportion is independent of covariates as we are interested in estimating the age-specific prevalence rate in a particular age group. If subject-specific probabilities of having testosterone deficiency are of interest instead, a logistic link function that allows mixing proportions depending on individual characteristics and some predictors can be incorporated into mixture models.

Acknowledgments

The author would like to thank the referee and the editor for their comments and suggestions that improved the presentation of this paper.

References

- Box, G. E. P. and Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society, Series B* **26**, 211-246.
- Chen, H., Chen, J. and Kalbfleisch, J. D. (2001). A modified likelihood ratio test for homogeneity in the finite mixture models. *Journal of the Royal Statistical Society, Series B* **63**, 19-29.
- Cicero, T. J. (1980). Effects of exogenous and endogenous opiates on the hypothalamic-pituitary-gonadal axis in the male. *Federation Proceedings* **39**, 2551-2554.
- Dacunha-Castelle, D. and Gassiat, E. (1999). Testing the order of a model using locally conic parameterization: population mixtures and stationary ARMA processes. *Annals of Statistics* **27**, 1178-1209.
- Day, N. E. (1969). Estimating the components of a mixture of normal distributions. *Biometrika* **56**, 463-474.
- Dempster, A., Laird, N. M. and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B* **39**, 1-38.
- Dobs, A. S. *et al.* (1988). Endocrine disorders in men infected with human immunodeficiency virus. *American Journal of Medicine* **84**, 611-616.
- Ellis, L. and Nyborg, H. (1992). Racial/ethnic variations in male testosterone levels: a probable contributor to group differences in health. *Steroids* **57**, 72-75.
- Everitt, B. S. and Hand, D. J. (1981). *Finite Mixture Distributions*. Chapman & Hall.
- Ferri, C., Bertozzi, M. A. and Zignego, A. L. (2002). Erectile dysfunction and hepatitis C virus infection. *Journal of the American Medical Association* **288**, 698-699.
- Gapstur, S. M. *et al.* (2002). Serum androgen concentrations in young men: a longitudinal analysis of associations with age, obesity, and race. The CARDIA male hormone study. *Cancer Epidemiology Biomarkers & Prevention* **11**, 1041-1047.
- Gay, N. J. (1996). Analysis of serological surveys using mixture models: application to a survey of parvovirus B19. *Statistics in Medicine* **15**, 1567-1573.
- Glass, A. R. *et al.* (1977). Low serum testosterone and sex-hormone-binding-globulin in massively obese men. *Journal of Clinical Endocrinological Metabolism* **45**, 1211-1219.
- Gray, G. (1994). Bias in Misspecified mixtures. *Biometrics* **50**, 457-470.
- Gutierrez, R. G. *et al.* (1995). Analysis of tomato root initiation using a normal mixture distribution. *Biometrics* **51**, 1461-1468.
- Harman, S. M. *et al.* (2001). Longitudinal effects of aging on serum total and free testosterone levels in healthy men. Baltimore Longitudinal Study of Aging. *Journal of Clinical Endocrinology and Metabolism* **86**, 724-731.
- Hathaway, R. J. (1985). A constrained formulation of maximum-likelihood estimation for normal mixture distributions. *Annals of Statistics* **13**, 795-800.

- Kiefer, J. and Wolfowitz, J. (1956). Consistency of the maximum-likelihood estimator in the presence of infinitely many incidental parameters. *Annals of Mathematical Statistics* **27**, 888-906.
- Klein, S. R. *et al.* (2005). Androgen levels in older men who have or who are at risk of acquiring HIV infection. *Clinical Infectious Diseases* **41**, 1794-1803.
- Lemdani, M. and Pons, O. (1999). Likelihood ratio tests in contamination models. *Bernoulli* **5**, 705-719.
- Lo, Y. (2005). Likelihood ratio tests of the number of components in a normal mixture with unequal variances. *Statistics & Probability Letters* **71**, 225-235.
- Lo, Y., Mendell, N. R. and Rubin, D. B. (2001). Testing the number of components in a normal mixture. *Biometrika* **88**, 767-778.
- Louis, T. A. (1982). Finding the observed information matrix when using the EM algorithm. *Journal of the Royal Statistical Society, Series B* **44**, 226-233.
- MacLean, C. J. *et al.* (1976). Skewness in commingled distributions. *Biometrics* **32**, 695-699.
- McLachlan, G. J. and Basford, K. E. (1988). *Mixture Models: Inference and Applications to Clustering*. Marcel Dekker.
- McLachlan, G. J. and Peel, D. (2000). *Finite Mixture Models*. Wiley.
- Misra, M., Papakostas, G. I. and Klibanski, A. (2004). Effects of psychiatric disorders and psychotropic medications on prolactin and bone metabolism. *Journal of Clinical Psychiatry* **65**, 1607-1618.
- Mohr, B. A. *et al.* (2005). Normal bound and nonbound testosterone levels in normally ageing men: results from the Massachusetts Male Ageing Study. *Clinical Endocrinology* **62**, 64-73.
- Pfeiffer, R. M., Gail, M. H. and Brown, L. (2000). A mixture model for the distribution of IgG antibodies to *Helicobacter pylori*: application to studying factors that affect prevalence. *Journal of Epidemiology and Biostatistics* **5**, 267-275.
- Pfeiffer, R. M. *et al.* (2008). Combining assays for estimating prevalence of human herpesvirus 8 infection using multivariate mixture models. *Biostatistics* **9**, 137-151.
- Schork, N. J. and Schork, M. A. (1988). Skewness and mixtures of normal distributions. *Communications in Statistics, Series A* **17**, 3951-3969.
- Thode, H. C. Jr., Finch, S. J. and Mendell, N. R. (1988). Simulated percentage points for the null distribution of the likelihood ratio test for a mixture of two normals. *Biometrics* **44**, 1195-1201.
- Titterton, D. M., Smith, A. F. M. and Makov, U. E. (1985). *Statistical Analysis of Finite Mixture Distributions*. Wiley.

-
- Thompson, T. J., Smith, P. J. and Boyle, J. P. (1998). Finite mixture models with concomitant information: assessing diagnostic criteria for diabetes. *Applied Statistics* **47**, 393-404.
- Turner, T. R. (2000). Estimating the propagation rate of a viral infection of potato plants via mixtures of regressions. *Applied Statistics* **49**, 371-384.
- Vine, M. F. (1996). Smoking and male reproduction: a review. *International Journal of Andrology* **19**, 323-337.
- Winters, S. J. *et al.* (2001). Testosterone, sex hormone-binding globulin, and body composition in young adult African American and Caucasian men. *Metabolism* **50**, 1242-1247.

Received August 23, 2007; accepted October 29, 2007.

Yungtai Lo
Department of Community and Preventive Medicine
Mount Sinai School of Medicine
One Gustave L. Levy Place, Box 1057
New York, NY 10029-6574
ungtai.lo@mssm.edu