Journal of Modern Applied Statistical Methods

Volume 6 | Issue 2 Article 25

11-1-2007

Generalized Linear Mixed-Effects Models for the Analysis of Odor Detection Data

Sandra Hall
University of Kansas Medical Center

Matthew S. Mayo
University of Kansas Medical Center, mmayo@kumc.edu

Xu-Feng Niu
Florida State University

James C. Walker
Florida State Unviersity

Journal of Modern Applied Statistical Methods
November, 2007, Vol. 6, No. 2, 608-618

1538 – 9472/07/$95.00

Generalized Linear Mixed-Effects Models for the Analysis of Odor Detection Data

Sandra Hall Matthew S. Mayo
University of Kansas Medical Center

Xu-Feng Niu James C. Walker
Florida State University

Olfactory detection has become a science of interest. Seven individuals' odor detection abilities are explored and an attempt is made to characterize all subjects with one generalized linear mixed effects model. Two methods of fitting the models were used and simulations were conducted to discover which method yielded the best results.

Key words: olfactory, conditional distribution, Metropolis Algorithm**,** Monte Carlo Newton Raphson Method, random effects, detectability, odor, human, sensitivity.

Introduction

The quality of indoor air is one of the least understood health problems that industry faces today. A major problem that poor indoor air quality causes is Sick Building Syndrome (EPA, 1989). This occurs when a substantial proportion of a building's occupants experience discomfort and health effects that are relieved upon leaving the building. It has been reported that sick buildings cause an estimated loss of between ten and one hundred billion dollars a year for non-medical aspects of diminished indoor air quality, excluding medical events such as asthmatic attacks (Fisk & Rosenfeld, 1997). Human symptoms of Sick Building Syndrome range from repetitive office headaches and common cold-like symptoms to serious ailments such as respiratory infections, asthma and allergies. A 1996 Cornell University study found that**,** in each of 35 buildings surveyed, at least 20% of the occupants had experienced symptoms associated with Sick Building Syndrome (Mann, 1998). Odor threshold is the point at which the probability of odor detection becomes greater than chance. Threshold is the most basic measure of sensory function. To understand higher order capabilities (e.g. odor discrimination, odor identification, identification of target in mixtures, and perception of odor quality), it is necessary to take into account the sensitivity of each individual to each chemical. Thus, it is important to have a valid way to quantify sensitivity. One example of why odor threshold might be studied is to gain a better understanding of issues related to olfaction such as Sick Building Syndrome. Another is that it has been hypothesized that early stages of Alzheimer's disease can be detected by a loss of odor detectability (Devanand et. al., 2000).

To help researchers understand this concept of accurately quantifying odor detection ability, a study was conducted at the Florida State University Sensory Research Institute's (SRI). Subjects received stimuli via a facemask that covered the person's mouth and nose, although the stimuli were taken in through only the nose. The subject then responded using a computer mouse and monitor screen as to whether or not an odor was detected. By using this olfactometer, the subject was given a precise concentration of the chemical (Walker et. al., 2003).

Sandra Hall is an independent statistical consultant. Her research interests include applied statistical analysis, brain imaging data analysis and computation statistics. Email: sandy_b_h@hotmail.com. Xu-Feng Niu is Professor of Statistics. His research interests include time series analysis and spatial statistics, niu@stat.fsu.edu. Matthew Mayo is Professor and Chair of Biostatistics and his research interests include clinical trials and experimental design, MMAYO@kumc.edu. James Walker's interests are in quantifying capabilities of the olfactory systems, jwalker@psy.fsu.edu.

For that study, seven subjects were recruited. They were selected so as to have a variety of different ages as well as subjects of each and both genders. As the subjects responded to the posters via phone calls they were asked routine questions to determine if they had a prior history of nasal defect. The researchers desired both a male and female subject in each of the following age categories: 18-20, 21-30, 31-44 and > 45 years. After three weeks of recruitment, no male subject was found in the 31-44 year- old group and the experimenters elected to continue the study without a male subject from this age group. Each subject completed 12 to 14 sessions over the course of 3 to 4 months. Each session consisted of 75 trials (15 trials of clean air in addition to 15 trials at each of 4 different concentrations of amyl acetate) each lasting approximately 18 seconds and separated by 90-second intervals. Hence, a typical session ran for 2 hours and 15 minutes.

A trial consists of the subject being asked to come to the mask, where they breathed the stimulus. The subject then used a mouse to click whether they detected an odor or not. The method of stimulus presentation allowed for very precise control. Before and during stimulus presentation, breathing was measured. After several seconds of pre-stimulus sampling of respiratory behavior was stored, the next exhalation onset triggered operation of the flow valve that (unless a clean air trial is scheduled) sent odorant to the mask. This approach essentially eliminated the vexing problem of a stimulus rise time, because the concentration reached its asymptotic value during the interval from an exhalation onset to the next inhalation onset (Prah, Sears & Walker, 1995). The specific concentrations and corresponding yes's (y's) and no's (n's) from the subject for the session were recorded on the same computer that randomized the concentrations to be given.

Traditionally, longitudinal data might have been analyzed using a generalized linear model (GLM) for each subject. However, this method does not accommodate a population based model, which is ultimately desired. Thus, generalized linear mixed models (GLMMs) were used to address the problem.

The class of functions known as GLMMs extends GLMs by adding random effects to the linear predictor(s). The benefit of this model is that it allows for responses that are correlated and non-normally distributed, which can frequently occur in actual problems. By including the random effects, the GLMMs can model correlated errors, smooth regression relationships and model dependence among variables that occurs in repeated measure designs. Many problems involve multiple sources of variation such as analysis of data that has a hierarchical structure like clinical trial data. The GLMM can be used to model such data. In this particular study, the model needed to account for the randomness of the session. This random nature is not considered in the traditional generalized linear model which is initially used to describe the data analyzed in this study.

A natural alternative to this approach is to utilize generalized estimating equations (GEE). The GEE approach is attractive because it allows for a weighted estimate of the regression parameters and correctly adjusts for correlated data. The problems with GEE are that a) it provides only a population model of the data and b) it requires a large amount of subjects for the large sample distribution properties to provide correct standard errors for inference (hypothesis testing and confidence intervals). Since only seven subjects were available, the GEE approach would not be an appropriate choice.

GLMMs are useful as an alternative to GEE and might be an approach that is useful in small sample sizes. For example, SAS has a procedure called "GLIMMIX" that is promising. The problem is that GLIMMIX has not been completely assessed for its usefulness in small sample sizes.

GLMMs provide insight into the behavior, but accurately estimating the model can be quite difficult. Because GLMMs are an extension of GLMs, one might logically try to fit the model using maximum likelihood, the common method to fit GLMs. The maximum likelihood method will only work for very simple GLMMs due to the need to numerically evaluate high dimensional integrals that are

irreducible. Thus, statisticians have looked for other methods to fit these models that do not involve the difficulties of the numerically complicated integration. Many different methods have been proposed to fit generalized linear mixed models. The model and two specific previously proposed methods (one being the SAS- GLIMMIX approach) will be discussed. Shown next will be results of the simulation study comparing these two methods for the data, fit the model that was deemed best in the simulation study and summarize the work.

Methodology

Model

Let Y_{ij} be the jth response for subject i, with j = 1 to n_i and i = 1 to m where m is the number of subjects and n_i is the number of observations per subject. Let X_{ijk} be the j^{th} value of the k^{th} fixed effect for subject i, with k = 1 to p and i and j as described previously. Thus, the traditional generalized linear model is

$$g(\mu_{ij}) = \beta_0 + \sum_{k=1}^{p} \beta_k X_{ijk}$$

with $\mu_{ij} = E(Y_{ij})$, where g is the link function and p is the number of different fixed effects.

Upon including the random effects, the model becomes:

$$\eta_{ij} = g(\mu_{ij}) = \beta_0 + \sum_{k=1}^{p} \beta_k X_{ijk} + \sum_{l=1}^{c} \alpha_{il} Z_{ijl} \quad (1)$$

where **X** is still assumed to be the matrix for the fixed effects and Z_{ijl} is the j^{th} value for the l^{th} random effect for subject i where l = 1 to c with c being the number of random effects. Also, it is assumed that $\mu_{ij} = E(Y_{ij} \mid \boldsymbol{\alpha}_i, \boldsymbol{\beta})$ and $\text{var}(Y_{ij} \mid \boldsymbol{\alpha}_i, \boldsymbol{\beta}) = \phi a_i v(\mu_i)$, where ϕ is a dispersion parameter, $v(\cdot)$ is a specified variance function and a_i is a known constant. The random effects $(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_m)$ are assumed to be independent with mean **0** and $\text{cov}(\boldsymbol{\alpha}_i)$=**D**. It is assumed that the elements of **Y** conditional on $\boldsymbol{\alpha}$ are both independent and drawn from an exponential family distribution. Finally, $\boldsymbol{\alpha}$ is assumed to be distributed $f_{\alpha}(\boldsymbol{\alpha} \mid \mathbf{D})$. Let $\boldsymbol{\eta}_i = (\boldsymbol{\eta}_{i1}, \ldots, \boldsymbol{\eta}_{in_i})^T$. Then, the function becomes:

$$f_{Y_i|\alpha}(\mathbf{Y}_i \mid \boldsymbol{\alpha}, \boldsymbol{\beta}, \phi) = \exp\left\{ \frac{\mathbf{Y}_i \boldsymbol{\eta}_i - c(\boldsymbol{\eta}_i)}{a(\phi)} + d(\mathbf{Y}_i, \phi) \right\} \quad (2)$$

and the likelihood function is:

$$L(\beta, \phi, \mathbf{D} \mid \mathbf{Y}) = \int \prod_{i=1}^{n} f_{Y_i|\alpha}(\mathbf{Y}_i \mid \boldsymbol{\alpha}, \boldsymbol{\beta}, \phi) f_{\alpha}(\boldsymbol{\alpha} \mid \mathbf{D}) d\boldsymbol{\alpha} \quad (3)$$

(Breslow &Clayton, 1993; Clayton, 1993; Jiang, 1998; Lin & Breslow, 1996; Lindstrom & Bates, 1990; McCulloch, 1997; Vonesh, 1996).

Simulation methods

Several methods have been proposed to estimate the solution to the generalized linear mixed model. McCulloch (1997) proposed algorithms for Monte Carlo EM (MCEM) and Monte Carlo Newton-Raphson (MCNR). Lin and Breslow (1996) proposed using a penalized quasi-likelihood approach with bias correction to estimate the model.

The Monte Carlo EM algorithm considers the random effects $\boldsymbol{\alpha}$ to be missing data. Therefore, the complete data would be W=(**Y**, $\boldsymbol{\alpha}$) and the log likelihood for the complete data would be

$$\ell_W = \sum_i \ln f_{Y_i|\alpha}(\mathbf{Y}_i \mid \boldsymbol{\alpha}, \boldsymbol{\beta}, \phi) + \ln f_{\alpha}(\boldsymbol{\alpha} \mid \mathbf{D}) \quad (4)$$

Thus, the $\mathbf{Y}_i$'s become independent when the $\boldsymbol{\alpha}$'s are known. Note that $\boldsymbol{\beta}$ and ϕ enter into the above equation only in the first term, the maximization with respect to those two terms is similar to a standard GLM computational problem with the $\boldsymbol{\alpha}$'s known. Then maximizing with respect to **D** involves replacing the sufficient statistics with their conditional expected value and then performing maximum likelihood using the distribution of $\boldsymbol{\alpha}$. McCulloch's (1997) algorithm follows:

1. Choose starting values for $\boldsymbol{\beta}^{(0)}$, $\phi^{(0)}$, and $\mathbf{D}^{(0)}$. Set m=0.
2. Calculate (with expectations evaluated under $\boldsymbol{\beta}^{(m)}$, $\phi^{(m)}$, and $\mathbf{D}^{(m)}$):
   a. $\boldsymbol{\beta}^{(m+1)}$ and $\phi^{(m+1)}$ which maximize $E[\ln f_{Y|\alpha}(\mathbf{Y} \mid \boldsymbol{\alpha}, \boldsymbol{\beta}, \phi) \mid \mathbf{Y}]$
   b. $\mathbf{D}^{(m+1)}$ which maximizes $E[\ln f_{\alpha}(\boldsymbol{\alpha} \mid \mathbf{D}) \mid \mathbf{Y}]$
   c. Set m=m+1
3. If convergence is achieved, declare $\boldsymbol{\beta}^{(m+1)}$, $\phi^{(m+1)}$, and $\mathbf{D}^{(m+1)}$ to be maximum likelihood estimates. Otherwise repeat step two.

Neither expectation in step two can actually be found in closed form. It is, however, possible to produce random draws from the conditional distribution of $\boldsymbol{\alpha}|\mathbf{Y}$ by using the Metropolis algorithm (Vonesh, 1996), which does not require a specification of f_Y. Monte Carlo approximations may then be formed in order to estimate the two required expectations. For sufficiently large sample sizes, it was discovered that this method gains likelihood and would converge to a local maximum under appropriate regularity conditions (McCulloch, 1997). Although this holds promise, in variance component problems, such as with GLMMs, the likelihood surfaces are not necessarily unimodal; thus, this method may only converge to a local maximum and never to the global one. A second problem is that it is limited to the binary response with the probit link. Incorporating this Metropolis algorithm into the EM algorithm gives the MCEM algorithm below (McCulloch, 1997):

1. Choose starting values for $\boldsymbol{\beta}^{(0)}$, $\phi^{(0)}$, and $\mathbf{D}^{(0)}$. Set m=0.
2. Generate N values, $\boldsymbol{\alpha}^{(1)}$, $\boldsymbol{\alpha}^{(2)}$, … , $\boldsymbol{\alpha}^{(N)}$ from $f_{\alpha|Y}(\boldsymbol{\alpha} \mid \mathbf{Y}, \boldsymbol{\beta}^{(m)}, \phi^{(m)}, \mathbf{D}^{(m)})$ using the Metropolis algorithm:
   a. Choose $\boldsymbol{\beta}^{(m+1)}$ and $\phi^{(m+1)}$ to maximize a Monte Carlo estimate of $E[\ln f_{Y|\alpha}(\mathbf{Y} \mid \boldsymbol{\alpha}, \boldsymbol{\beta}, \phi) \mid \mathbf{Y}]$ that is maximize $\frac{1}{N}\sum_{k=1}^{N} \ln f_{Y|\alpha}(\mathbf{Y} \mid \boldsymbol{\alpha}^{(k)}, \boldsymbol{\beta}, \phi)$
   b. Choose $\mathbf{D}^{(m+1)}$ to maximize $\frac{1}{N}\sum_{k=1}^{N} \ln f_{\alpha}(\boldsymbol{\alpha}^{(k)} \mid \mathbf{D})$
   c. Set m=m+1
3. If convergence is achieved declare $\boldsymbol{\beta}^{(m+1)}$, $\phi^{(m+1)}$, and $\mathbf{D}^{(m+1)}$ to be maximum likelihood estimates. Otherwise repeat step two.

The next method that McCulloch (1997) used is the Monte Carlo Newton-Raphson method. This method also seemed robust to starting values. Again, since the likelihood surfaces are not unimodal, they are definitely not concave and thus this method may not converge at all, let alone to the global maximum. In practice it was discovered that this method generally got close to the correct answer. The algorithm appears below:

1. Choose starting values for $\boldsymbol{\beta}^{(0)}$, $\phi^{(0)}$, and $\mathbf{D}^{(0)}$. Set m=0.
2. Generate N values, $\boldsymbol{\alpha}^{(1)}$, $\boldsymbol{\alpha}^{(2)}$, … , $\boldsymbol{\alpha}^{(N)}$ from $f_{\alpha|Y}(\boldsymbol{\alpha} \mid \mathbf{Y}, \boldsymbol{\beta}^{(m)}, \phi^{(m)}, \mathbf{D}^{(m)})$ using the Metropolis algorithm and use them to form Monte Carlo estimates of the expectations (denoted as $\hat{E}[\cdot]$):
   a. Calculate

$$\boldsymbol{\beta}^{(m+1)} = \boldsymbol{\beta}^{(m)} + \hat{E}[\mathbf{X}^{\mathrm{T}}\mathbf{S}(\boldsymbol{\theta}^{(m)}, \boldsymbol{\alpha})\mathbf{X} \mid \mathbf{Y}]^{-1}\mathbf{X}^{\mathrm{T}}([\mathbf{S}(\boldsymbol{\theta}^{(m)}, \boldsymbol{\alpha})\frac{\partial \boldsymbol{\eta}}{\partial \boldsymbol{\mu}}\Big|_{\theta=\theta^{(m)}} (\mathbf{Y} - \boldsymbol{\mu}(\boldsymbol{\beta}^{(m)}, \boldsymbol{\alpha})) \mid \mathbf{Y}]) \quad (5)$$

   where $\boldsymbol{\mu}_{ij}(\boldsymbol{\theta}, \boldsymbol{\alpha}) = E[\mathbf{Y}_{ij} \mid \boldsymbol{\alpha}_i]$, $\frac{\partial \boldsymbol{\eta}}{\partial \boldsymbol{\mu}} = diag\left\{\frac{\partial \eta_{ij}}{\partial \mu_{ij}}\right\}$, and $\mathbf{S}(\boldsymbol{\theta}, \boldsymbol{\alpha})^{-1} = diag\left\{\left(\frac{\partial \eta_{ij}}{\partial \mu_{ij}}\right)^2 \mathrm{var}(Y_{ij} \mid \boldsymbol{\alpha}_i)\right\}$

a. Calculate $\phi^{(m+1)}$ to solve $E\left[\frac{\partial \ln f_{Y|\alpha}(\mathbf{Y} \mid \boldsymbol{\alpha}, \theta)}{\partial \phi} \mid \mathbf{Y}\right] = 0$ or a scoring equation.

c. Choose $\mathbf{D}^{(m+1)}$ to maximize $\frac{1}{N}\sum_{k=1}^{N} \ln f_{\alpha}(\boldsymbol{\alpha}^{(k)} \mid \mathbf{D})$

d. Set m=m+1

3. If convergence is achieved declare $\boldsymbol{\beta}^{(m+1)}$, $\phi^{(m+1)}$, and $\mathbf{D}^{(m+1)}$ to be maximum likelihood estimates. Otherwise repeat step two.

The MCEM and MCNR algorithms are very similar. In fact, the maximization to calculate the fixed effects coefficients in the MCEM algorithm cannot explicitly be carried out for binomial data, and thus, an estimation method is necessary, such as the Newton Raphson Method. Thus, for our purposes, the MCNR is equivalent to the MCEM algorithm.

Breslow and Clayton (1993) proposed performing a method known as penalized quasi-likelihood analysis (PQL) in order to approximate the maximum likelihood estimates. The key feature of this analysis is that it is easy to implement, especially since there exists a SAS macro for this method. The procedure is to repeatedly fit a linear mixed model to a modified dependent variable. They realized that a limitation of the PQL is that when assessing the uncertainty in both random and fixed effects it does not take into account the contribution of the estimated variance components. Lin and Breslow (1996) proposed a four-step procedure of bias correction for the PQL.

Lin and Breslow (1996) provided a four step algorithm to find the bias-corrected penalized quasi-likelihood estimates of the regression coefficients and variance components. They performed simulation studies and found that the bias correction procedure can improve asymptotic performance of the estimates for correlated binary data. They also discovered that this simple correction procedure would effectively reduce the bias of variance components of the PQL estimates and the associated mean square error as long as the sample size is reasonably large.

Results

The two methods, Monte Carlo Newton-Raphson and penalized quasi-likelihood with bias correction, were used in a simulation study in order to determine which method better estimates the fixed affects as well as the random effects. The MCNR program was written in Matlab. The penalized quasi-likelihood program with bias correction (PQBC) was coded using SAS and the GLIMMIX macro available from SAS's website: http://ftp.sas.com/techsup/download/stat/.

The response vector for each program was generated using a binomial random generator. Binomial probabilities were calculated for each combination of subject, session and concentration.

It was then determined how many simulations of the program should be carried out in order to have results that converge. Thus, each of the programs was run a total of 100 and 1000 times, respectively. Each time, a new response vector was generated. The response vectors were based on the following model, using concentrations from four of the seven subjects,

$$p_{ij} = \frac{e^{-15-3.5*conc_{ij}}}{1+e^{-15-3.5*conc_{ij}}} \qquad (6)$$

where p_{ij} is the probability for the j^{th} concentration of subject i. The model gives the probability to be used for each concentration value. The binomial generator was then used along with the probabilities found in the model to generate fifteen binary responses for each concentration as it occurred. It can be seen, in Table 1, that both programs appear to have converging results with as few as 100 simulations.

Next, it is necessary to test the random effects portion of the programs. For this step, concentrations for four of the seven subjects were used. For each combination of subject, concentration, σ level (σ = 0.5, 1.5, 2.0, 2.5 and 3.0) and (α, β) pair [values of (α, β) used were as follows: (-10, -2), (-12.5, -2.75), (-15, -3.5),

Table 1. Simulation Size Necessary

Number of simulations		MCNR	PQBC
100	Intercept	-14.9606	-15.0603
	Slope	-3.4918	-3.5031
1000	Intercept	-15.0111	-15.0100
	Slope	-3.5028	-3.5066

(-17.5, -4.25), and (-20, -5)] the following process was used to generate simulation data sets:

Step 1: Generated a random number, γ, from the $N(0, \sigma^2)$ distribution.

Step 2: Generated a binomial probability using the following model:

$$p = \frac{e^{\alpha+\beta*conc+\gamma}}{1+e^{\alpha+\beta*conc+\gamma}} \qquad (7)$$

Step 3: Used this generated probability to randomly generate data from the binomial distribution with n equal to 15 and the value generated in step 2 for each time the subject/concentration combination occurred. This gave the ability to weight the different concentrations properly for each subject.

This process was repeated 100 times, so that 100 different data sets were generated for each individual combination of subject, concentration, σ level and (α, β) pair.

In Table 2, models for ten subjects with a standard deviation of 0.5, 1.5, 2.0, 2.5, and 3.0 are considered. The MCNR program tends to overestimate the slope and intercept, while the PQBC program tends to estimate the slope and intercept accurately. The PQBC program seems to underestimate the standard deviation, yet the MCNR program tends to estimate the standard deviation fairly close to the actual value.

In Table 3, models for twenty simulated subjects with a standard deviation of 0.5, 1.5, 2.0, 2.5, and 3.0 are considered. The MCNR program tends to come close to estimating the slope and the intercept or else slightly overestimate them, while the PQBC program tends to estimate the slope and intercept rather accurately. The PQBC program seems to underestimate the standard deviation only when it is equal to 0.5 and 1.0, otherwise it estimates the standard deviation fairly well. The MCNR program tends to estimate the standard deviation fairly close to the actual value.

Upon considering both of these tables, it is observed that the MCNR program better estimates the standard deviation then the PQBC program does. Both methods do a good job of estimating the slope and intercept; however, the PQBC program cannot accurately estimate the random effect term effectively when the number of subjects is small. It should also be noted that there does exist a procedure in SAS that has recently been developed to fit a general linear mixed effects model. The problem with this procedure is that it currently allows for only one random effect. Therefore, it will not be used here as it has the potential for two random effects, one for subject and one for session.

Based on these findings, it was decided that the MCNR program would be the best program to use to try to fit the actual data since the number of subjects that is present is seven.

Table 2. Simulation Results with Random Effects and 10 subjects

sigma=0.5		**MCNR**			**PQBC**		
True int	True slope	Int	Slope	S.D.	Int	Slope	S.D.
-10	-2	-9.9580	-1.9960	0.4441	-9.9587	-1.9895	0.1402
-12.5	-2.75	-12.6137	-2.7720	0.5213	-12.4233	-2.7599	0.1585
-15	-3.5	-14.9111	-3.4827	0.5129	-14.8682	-3.5199	0.1846
-17.5	-4.25	-17.7383	-4.3082	0.5748	-17.5716	-4.2465	0.1306
-20	-5	-20.0249	-4.9879	0.6037	-19.7835	-4.9974	0.1310
sigma=1.5							
True int	True slope	Int	Slope	S.D.	Int	Slope	S.D.
-10	-2	-9.6906	-1.9466	1.2642	-10.1027	-2.0090	1.4606
-12.5	-2.75	-11.4029	-2.5562	1.2313	-12.5513	-2.7654	1.3046
-15	-3.5	-13.8923	-3.2693	1.2745	-14.6752	-3.5230	1.6457
-17.5	-4.25	-15.9715	-3.9428	1.3471	-17.3032	-4.2515	0.9298
-20	-5	-18.6952	-4.6878	1.3556	-19.0461	-4.9920	0.6949
sigma=2.0							
True int	True slope	Int	Slope	S.D.	Int	Slope	S.D.
-10	-2	-8.5403	-1.7130	1.5299	-10.6636	-1.9815	1.1236
-12.5	-2.75	-10.4584	-2.3086	1.4860	-11.6325	-2.7387	2.7422
-15	-3.5	-12.9367	-3.0179	1.6964	-15.0677	-3.5473	1.9940
-17.5	-4.25	-15.5526	-3.7921	1.7620	-19.0472	-4.3063	1.8099
-20	-5	-17.7169	-4.4776	1.7517	-21.2486	-5.0573	1.8953
sigma=2.5							
True int	True slope	Int	Slope	S.D.	Int	Slope	S.D.
-10	-2	-8.3160	-1.7365	1.8244	-8.8011	-1.9977	3.0365
-12.5	-2.75	-10.1496	-2.2878	1.9534	-13.0257	-2.7455	4.0708
-15	-3.5	-12.6502	-3.0247	1.9906	-15.9416	-3.4950	2.6698
-17.5	-4.25	-15.6909	-3.7334	2.0197	-17.1718	-4.2794	2.2737
-20	-5	-16.3706	-4.0743	1.9568	-22.0580	-5.0519	3.5395
sigma=3.0							
True int	True slope	Int	Slope	S.D.	Int	Slope	S.D.
-10	-2	-8.4546	-10	-2	-8.4546	-10	-2
-12.5	-2.75	-10.0818	-12.5	-2.75	-10.0818	-12.5	-2.75
-15	-3.5	-12.3014	-15	-3.5	-12.3014	-15	-3.5
-17.5	-4.25	-15.6226	-17.5	-4.25	-15.6226	-17.5	-4.25
-20	-5	-16.2475	-20	-5	-16.2475	-20	-5

Table 3. Simulation Results with Random Effects and 20 subjects

sigma=0.5		**MCNR**			**PQBC**		
True int	True slope	Int	Slope	S.D.	Int	Slope	S.D.
-10	-2	-9.9054	-1.9816	0.4930	-9.4116	-1.9873	0.2649
-12.5	-2.75	-12.4214	-2.7306	0.4904	-12.5022	-2.7393	0.2649
-15	-3.5	-14.9141	-3.4801	0.4938	-15.0879	-3.4892	0.2626
-17.5	-4.25	-17.6045	-4.2722	0.5487	-17.4891	-4.2707	0.2662
-20	-5	-20.2569	-5.0673	0.6018	-19.9316	-4.9978	0.2381
sigma=1.5							
True int	True slope	Int	Slope	S.D.	Int	Slope	S.D.
-10	-2	-9.3067	-10	-2	-9.3067	-10	-2
-12.5	-2.75	-11.6219	-12.5	-2.75	-11.6219	-12.5	-2.75
-15	-3.5	-13.9817	-15	-3.5	-13.9817	-15	-3.5
-17.5	-4.25	-16.2716	-17.5	-4.25	-16.2716	-17.5	-4.25
-20	-5	-19.3602	-20	-5	-19.3602	-20	-5
sigma=2.0							
True int	True slope	Int	Slope	S.D.	Int	Slope	S.D.
-10	-2	-8.9972	-10	-2	-8.9972	-10	-2
-12.5	-2.75	-11.2885	-12.5	-2.75	-11.2885	-12.5	-2.75
-15	-3.5	-13.4287	-15	-3.5	-13.4287	-15	-3.5
-17.5	-4.25	-16.0496	-17.5	-4.25	-16.0496	-17.5	-4.25
-20	-5	-18.2798	-20	-5	-18.2798	-20	-5
sigma=2.5							
True int	True slope	Int	Slope	S.D.	Int	Slope	S.D.
-10	-2	-8.6322	-1.7685	2.1067	-10.0411	-1.9679	3.0753
-12.5	-2.75	-10.8068	-2.3196	2.2214	-13.0177	-2.7644	3.7780
-15	-3.5	-13.0848	-3.1166	2.1031	-15.2309	-3.5379	3.1254
-17.5	-4.25	-15.9064	-3.6013	2.2602	-17.6589	-4.2495	2.1665
-20	-5	-16.8736	-4.1663	2.1241	-19.0147	-5.0061	1.8221
sigma=3.0							
True int	True slope	Int	Slope	S.D.	Int	Slope	S.D.
-10	-2	-8.7646	-1.7309	2.6432	-9.2498	-1.9716	3.1413
-12.5	-2.75	-10.1691	-2.2667	2.4309	-12.1977	-2.7603	2.8307
-15	-3.5	-12.8800	-3.0994	2.5139	-14.7616	-3.5233	2.9032
-17.5	-4.25	-15.8659	-3.5443	2.3783	-17.6541	-4.3037	3.6437
-20	-5	-16.3182	-4.1412	2.4833	-20.0524	-5.0549	3.5551

Final model

Now the generalized linear mixed-effects models will be applied to the actual odor detection data. This begins by performing the MCNR analysis with both session and subject being random variables, while concentration remains the fixed variable. Subjects were chosen as a random effect because our previous analysis found that each subject did yield a different model. One of the main purposes of including random effects is to accommodate different subjects in one model with the random term. Session was chosen as a second possible random effect because it was somewhat significant in our early analysis of the data, and it is random in that the subjects may vary slightly from one session to another.

Five of the subjects had similar coefficients for their individual fixed-effects models,

$$p_{ij} = \frac{e^{\alpha+\beta*conc_{ij}}}{1+e^{\alpha+\beta*conc_{ij}}} \qquad (8)$$

For example, slope estimates for the five subjects were -0.600, -0.695, -0.487, -0.872, and -0.471 while intercept estimates were -2.881, -3.880, -1.468, -4.381, and -2.457. For the remaining two subjects, slope estimates were -3.090 and -2.287, while intercept estimates were -11.577 and -9.225;. It could be speculated that these two groupings indicate that there are two categories of smellers and that it might prove useful to split the group of seven into these two separate groups to lessen the variability of the data for modeling purposes. This began, however, by keeping all seven subjects together and estimating a model. The general form of the model is

$$p_{ij} = \frac{e^{\alpha+\beta*conc_{ij}+\gamma_{subj}+\gamma_{sess}}}{1+e^{\alpha+\beta*conc_{ij}+\gamma_{subj}+\gamma_{sess}}} \qquad (9)$$

where γ_{sess} could be zero.

From Table 4, it is evident that the variability was quite large when all seven subjects were together and hence the resulting model had extraordinarily small coefficients. Thus, the data was split into two groups and estimated a separate model for each group; the results appear in Table 4. Therefore, several models based on the split groups of subjects will

Table 4. Final Models

No. of subjects in model	Both subject and session as random effects	Only subject as random effect	No random effect
7	α = -145.8455, β = -40.2636, $\gamma_{subject}$=N(0,0.8069), $\gamma_{sess.}$=N(0,0.7281)	α = -368.0081, β = -80.3047, $\gamma_{subject}$=N(0,2.7066)	α = -453.2103, β = -98.5135
5	α = -3.4457, β = -0.7252, $\gamma_{subject}$=N(0,1.0605), $\gamma_{sess.}$=N(0,0.1730)	α = -3.5403, β = -1.0846, $\gamma_{subject}$=N(0,4.6057)	α = -2.9907, β = -0.6833
2	α = -3.5596, β = -0.6995, $\gamma_{subject}$=N(0,1.1177), $\gamma_{sess.}$=N(0,0.0013)	α = -2.7960, β = -0.8711, $\gamma_{subject}$=N(0,2.3808)	α = -3.1377, β = -0.6414

be considered to see which will yield a better fitting model.

Vonesh, Chinchilli and Pu (1996) observed that for a generalized linear mixed-effects model, a valid measure of the goodness of fit of the model is given by r_c. There are several advantages to using r_c over other methods. First, it does not require the specification of a null model. Second, it measures the level of concordance between y_i and $\hat{y}_i$. A higher value of r_c indicates a better fitting model. (Vonesh, Chinchilli & Pu, 1996)

For the three models that involved the 5 subjects, the r_c follows: for the one with two random effects the r_c was 0.418, for the model with only subject as a random effect the r_c was slightly better (higher) with a value of 0.445 and for the model with no random effects the r_c was only 0.342. Thus, for the 5-subject group, the best model is the one that includes only subject as the random effect. Upon looking at the three models that involved the 2 subjects the r_c follows: for the one with two random effects it was 0.371, for the model with only subject as a random effect the r_c was not quite as good with a value of 0.316 and the model with no random effects included yielded an r_c of 0.318. Thus, for the 2 subject group, the best model is the one that includes both subject and session as the random effects.

Therefore, the conclusion is drawn that the data for all seven subjects can be best represented using the following two models. For the group of five subjects the best model is

$$p_{ij} = \frac{e^{-3.5403-1.0846 * conc_{ij} + \gamma_{subject}}}{1 + e^{-3.5403-1.0846 * conc_{ij} + \gamma_{subject}}} \quad (10)$$

where γ=N(0, 4.6057) and for the group of two subjects the best model is

$$p_{ij} = \frac{e^{-3.5596-0.6995 * conc_{ij} + \gamma_{sub} + \gamma_{sess}}}{1 + e^{-3.5596-0.6995 * conc_{ij} + \gamma_{sub} + \gamma_{sess}}} \quad (11)$$

where γ_{sub}~N(0, 1.1177) and γ_{sess}~N(0, 0.0013). Thus, based on this small sample of individuals, it was found that two models will adequately represent the whole sample of seven individuals as opposed to the idea of finding a single model for each subject. It also would make it very difficult to adequately model the population as a whole if there had been individual models for each subject.

Conclusion

How accurate are people at detecting odors? In general terms, the question could also be stated as sensitivity – what is lowest concentration needed for reliable, if not perfect, detection? From there, there was an attempt to characterize all seven subjects with one generalized linear mixed effects model.

Two methods of fitting the generalized linear mixed effects models were used. Simulations were conducted to discover which method would yield the best results, in terms of stable estimates and a high r_c value, for the data. It was discovered that for this data, the method that would yield the best results was the MCNR method. Once this method was implemented, it was discovered the data was best fit by two models as opposed to just one model. The subjects were split into one group of five and one group of two based on the results discovered in the initial portion of the simulation study. For the group of five, it was necessary to have a random effects term for the subjects and for the group of two, it was necessary to have a random effects term for the subjects and another for the sessions.

Thus, all seven subjects' odor detection ability was able to be modeled for the one chemical tested through the use of two models. This is an improvement over the seven models that were initially investigated. The benefit of the smaller number of models is that it allows one to represent a population's ability to detect odors with just a few models instead of a different model for each individual in the population.

It would be instructive to perform a study with a larger sample in which the same task was asked of participants as in this study, namely: Do you detect an odor or not? An ideal situation would be to have many subjects of each gender and in each age group. This would

allow an expansion of these models to attempt to include a term for gender and also for age. Some researchers have hypothesized that as humans age, they begin to lose their sense of smell (Doty, 1994). Others (Hales, 1999) have wondered if sensitivity varies with gender. By expanding the model, it would begin to answer these questions.

References

Breslow, N. E. & Clayton, D. G. (1993). Approximate inference in generalized linear mixed models. *Journal of the American Statistical Association, 88*, 9-25.

Clayton, D. (1992). Generalized linear mixed models in biostatistics. *Statistician, 41*, 327-328.

Devanand, DP, Michaels-Marston, KS, Liu, X, Pelton, GH, Padilla, M, Marder, K, Bell, K, Stern, Y, & Mayeux, R (2000). Olfactory deficits in patients with mild cognitive impairment predict Alzheimer's disease at follow-up. *The American Journal of Psychiatry*, *15,* 1399-1405.

Doty RL. (1994). Olfactory dysfunction in the elderly and in Alzheimer's disease. In (Kurihara K, Suzuki N, Ogawa H, eds.) *Olfaction and taste XI: proceedings of the 11th International Symposium on Olfaction and Taste and of the 27th Japanense Symposium on Taste and Smell*. Joint meeting held at Kosei-nenkin Kaikan, Sapporo, Japan, July 12-16, 1993. Tokyo: Springer-Verlag, 597-601.

Environmental Protection Agency (1989). Report to congress on indoor air quality, Vol. ii. Assessment and Control of Indoor Air Pollution

Fisk W. J. & Rosenfeld, A. H. (1997). Estimates of improved productivity and health from better indoor environments *Indoor Air, 7*, 158-172.

Hales, D (1999). *Just Like a Woman: How Gender Science is Revealing What Makes Us Female*. New York, NY: Bantam Books.

Jiang, J. (1998). Consistent Estimators in Generalized Linear Mixed Models. *Journal of the American Statistical Association*, *93*, 720-729.

Lin, X & Breslow, N. E. (1996). Bias correction in generalized linear mixed models with multiple components of dispersion. *Journal of the American Statistical Association*, *91,* 1007-1016.

Lindstrom, M. J. & Bates, D. M. (1990). Nonlinear mixed effects models for repeated measures data. *Boimetrics*, *46*, 673-687.

Mann, Arnold. (1998, December 21). This Place Makes me Sick. *Time Magazine*, 152 (25).

McCulloch, C. E. (1997). Maximum likelihood algorithms for generalized linear mixed models. *Journal of the American Statistical Association, 92,* 162-170.

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., & Teller, E. (1953). Equations of state calculations by fast computing machine. *Journal of Chemical Physics, 21*, 1087-1091.

Prah, J. D., S. B. Sears & J. C. Walker (1995) Modern approaches to air-dilution olfactometry. In: *Handbook of Olfaction and Gustation*, R. L. Doty (Ed.) Marcel Dekker, New York, pp. 227-255.

Vonesh, E. F. (1996). A note on the use of laplace's approximation for nonlinear mixed-effects models. *Biometrika*, *83*, 447-452.

Vonesh, E. F., Chinchilli, V.M. & Pu, K. (1996). Goodness-of-fit in generalized nonlinear mixed-effects models. *Biometrics, 52,* 572-587.

Walker, JC, Hall, SB, Walker, DB, Kendal-Reed, MS, Hood, AF & Niu, X-F (2003). Human odor detectability: new methodology used to determine threshold and variation. *Chemical Senses*, *28,* 817-826.